THÈSE DE DOCTORAT

CARACTÉRISTIQUES DE LA PAROLE ET DU COMPORTEMENT OCULOMOTEUR DU TROUBLE COGNITIF INDUIT PAR UN CANCER ET SES TRAITEMENTS

Présentée en vue de l'obtention du grade de docteur
en **Sciences du Langage** de l'**Université de Montpellier Paul-Valéry**

Ecole Doctorale n°58 – LANGUES, LITTÉRATURES, CULTURES, CIVILISATIONS (LLCC)

Présentée par **Amélie RICHARD**

Dirigée par **Fabrice HIRSCH**
Co-dirigée par **Sophie JACQUIN-COURTOIS**

Soutenue publiquement le **05 juin 2025** devant le jury composé de

Robert HARTSUIKER Professeur, Ghent University (Belgium)	Rapporteur
Bénédicte GIFFARD Professeure, Université de Caen Normandie	Rapporteure
Robin LICKLEY Maître de Conférences, Queen Margaret University (UK)	Examinateur
Corine ASTESANO Professeure, Université Toulouse – Jean Jaurès	Examinatrice
Serge PINTO Directeur de Recherche CNRS, Aix-Marseille Université	Examinateur
Sylvie MORITZ-GASSER Professeure, Université de Montpellier	Présidente du jury
Sophie JACQUIN-COURTOIS Professeure, Université Claude Bernard Lyon 1	Codirectrice de thèse
Fabrice HIRSCH Professeur, Université de Montpellier Paul-Valéry	Directeur de thèse

Aux femmes —
À celles qui se battent, qu'importe l'issue,
À celles qui se sont battues et celles qui ne se battent pas
À celles qui continuent d'avancer,
Aux femmes qui croient, à celles qui pensent et celles qui savent,
À celles qui doutent,
Aux femmes inspirantes dont j'ai croisé le chemin et à celles que je n'ai pas encore rencontrées.

To all women —
To those who fight, no matter the outcome,
To those who have fought and those who do not fight,
To those who keep moving forward,
To the women who believe, to those who think and those who know,
To those who doubt,
To the inspiring women whose paths I have crossed and to those I have yet to meet.

Remerciements

Je remercie chaleureusement Pr. Fabrice Hirsch et Pr. Sophie Jacquin-Courtois de votre confiance et de votre aide à la réalisation de ce projet ambitieux. Merci Sophie pour votre accompagnement depuis mon master d'orthophonie. Vous m'avez transmis votre passion pour la compréhension des troubles subtils de la cognition. J'espère la communiquer à mon tour aux plus jeunes générations.

J'aimerais remercier Pr. Sylvie Moritz-Gasser, Pr. Bénédicte Giffard, Pr. Serge Pinto (DR CNRS) et Pr. Corine Astésano d'avoir accepté d'être membres de mon jury de thèse. I would like to thank Prof. Robert Hartsuiker and Dr. Robin Lickley for having accepted to be members of my dissertation committee.

J'adresse mes remerciements aux membres de l'École Doctorale 58 Langues, Littératures, Cultures, Civilisations de m'avoir attribué un financement par contrat doctoral.

Cette thèse doit son élégant caractère interdisciplinaire grâce à l'appui du laboratoire Praxiling et à celui du Centre de Recherche en Neurosciences de Lyon. Je leur suis très reconnaissante de leur accueil, et de leur apport humain et matériel qui m'ont permis d'effectuer mes recherches dans les meilleures conditions possibles. Je souhaite adresser mes remerciements à l'ensemble de la direction du Centre de Recherche en Neurosciences de Lyon pour son engagement à faire de ce lieu un environnement de travail agréable et épanouissant.

Je remercie vivement mes collaboratrices et collaborateurs, en particulier Maude Beaudoin, Anne Cheylus, Alexandre Foncelle et Laure Pisella, merci pour vos qualités de pédagogue et de votre rigueur. Manon Lelandais et Pauline Mouchès, merci pour votre esprit critique. Merci à Manon pour ta relecture minutieuse de mon manuscrit de thèse.

Merci infiniment aux participantes de mes études, spécialement aux patientes de l'hôpital Henry Gabrielle. Nos échanges ont nourri mes idées. Cette thèse est pour vous.

Je tiens à remercier la formidable équipe de gestionnaires et d'ingénieurs du CRNL, en particulier Gaëlle L., Hervé H., Justine E.-D., Pierre-Emmanuel A., Sonia A. et Thibaut W. Un immense merci à Lesly F. et Lydie M. pour leur aide dans mes démarches éthiques.

Merci à toutes les belles rencontres que j'ai faites au Neurocampus et au-delà. Merci Christina et Christelle de votre soutien, Clara et Maude de votre engagement auprès des chercheurs·euses non-permanent·e·s du CRNL, Pauline et Laura du club de *Machine Learning*, j'ai adoré nos réunions du

mercredi !, Linh d'avoir conçu et lancé l'aventure des séminaires *Mind Your Language* avec moi, Farha et Marion pour votre énergie.

Mes pensées vont à Ahamada, Lou, Elora, Cwiosna, mes collègues de Montpellier. Bon courage pour la fin de votre thèse.

Merci à mes amies et amis qui m'ont connue avant et pendant ma thèse : Thierry, Marion et Jess, Anaëlle, Silvan, Adeline et Nico, Lucie, Élise et Gabi, Alice et Kenan, Sabina, Aline et Nick.

Merci à Hélène Milan-Denys qui, il y a bien longtemps, m'a encouragée à écrire.

Thanks to Jaime & Cyril, Marika, Mark & Ann, Jen Yu & Régis, Quentin for the great talks we have. You are all amazingly generous !

Merci à Loïc, Francine, Maud et Manu, Marie et Victorien de m'avoir acceptée dans votre famille. Merci également à Maryse, nous nous sommes manquées, mais je te dois beaucoup.

Merci à ma famille Régis, Virginie, Mathilde et Pierre d'avoir cru en moi et d'être toujours là.

I'd like to thank my feet for keeping up the pace when my head refused to run.

Special thanks to Manon, my partner in life. As a great philospher said once : "When you know yourself and your strengths, there is always something you can do to raise the standard for everyone." (M. Rapinoe, One Life [2020]). You were a rock while I was under water, a lighthouse when I got lost, a mountain as I stood on a hill. I can't wait to move on to the next step with you.

What's next ?

Résumé

Cette thèse s'intéresse au trouble cognitif subtil induit par un cancer et ses traitements, ou *Cancer-Related Cognitive Impairment* (CRCI). Le CRCI désigne un ensemble de difficultés cognitives décrites par les patients en oncologie (hors cancers du système nerveux central). L'un de ses symptômes prédominants est le manque du mot. Bien qu'il puisse affecter jusqu'à 75% des patients atteints de cancer, le CRCI reste difficile à diagnostiquer au moyen des tests neuropsychologiques actuels, notamment en raison de l'absence d'outil conçu pour évaluer ce trouble. Cette thèse vise à identifier des marqueurs de la parole permettant de détecter le CRCI et à mieux comprendre les mécanismes cognitifs sous-jacents à ce trouble. Quarante-quatre survivantes d'un cancer du sein et 13 participantes témoins saines ont pris part à cette étude. Les participantes ont été invitées à raconter une histoire à partir d'une séquence d'images. Leurs productions orales ont été enregistrées et transcrites à l'aide de méthodes semi-automatisées. Les interruptions de parole (c.-à-d. les pauses silencieuses, les pauses remplies et les voyelles allongées) ont été annotées. Quatorze paramètres de parole ont été pré-sélectionnés et intégrés à des modèles d'apprentissage automatique. De plus, le comportement oculomoteur de 8 survivantes et 13 témoins a été enregistré à l'aide d'une technique de suivi oculaire. Des cartes de chaleur ont été générées à partir des coordonnées des points de fixation. Enfin, toutes les participantes ont passé une évaluation psychologique et cognitive. La relation entre les données de parole et les autres variables a été examinée à l'aide de statistiques bayésiennes. Les modèles d'apprentissage automatique ont identifié un sous-ensemble de trois marqueurs prosodiques du CRCI : le ratio parole/silence, la durée moyenne des pauses silencieuses et la durée moyenne des pauses remplies. La combinaison de ces trois paramètres a permis de classer les survivantes d'un cancer du sein présentant un CRCI, celles sans plainte cognitive et les témoins avec une précision de 73,7%. L'analyse des données oculomotrices a révélé un délai significatif entre le premier regard porté sur un objet et sa dénomination chez les survivantes. Ce résultat suggère que ces survivantes ont des difficultés à accéder à leur stock lexical. De plus, les cartes de fixation suggèrent que les mécanismes cognitifs impliqués dans le CRCI pourraient être davantage associés à un déficit du contrôle exécutif qu'à des problèmes de mémoire. Cependant, les analyses des cartes de chaleur n'écartent pas la possibilité de déficits attentionnels affectant la recherche lexicale. Cette thèse offre de nouvelles perspectives quant aux outils diagnostiques basés sur la parole. Cette recherche sera élargie par de futures études sur d'autres types de cancers et des pathologies présentant des déficits cognitifs subtils.

Abstract

This dissertation sheds light on Cancer-Related Cognitive Impairment (CRCI). This condition refers to a set of cognitive changes reported by cancer patients (excluding cancers of the central nervous system). One of its predominant symptoms is word-finding difficulties. Although it can affect up to 75% of cancer patients, CRCI remains challenging to diagnose using current neuropsychological tests. This dissertation aims to identify speech markers that can detect CRCI, and to better understand the cognitive mechanisms underlying this disorder. Forty-four breast cancer survivors and 13 healthy control participants participated in this study. The participants were instructed to tell a picture-based story. Speech productions were recorded and transcribed using semi-automated methods. Speech breaks (i.e., silent pauses, filled pauses, and sustained vowels) were annotated. Fourteen speech parameters were preselected and incorporated into machine learning models. In addition, the oculomotor behavior of 8 survivors and 13 controls was recorded using an eye-tracking technique. Heatmaps were generated based on fixation point coordinates. Finally, all participants underwent a psychological and cognitive evaluation. The relationship between speech data and other variables was examined using Bayesian statistics. Machine learning models identified a subset of three speech markers of CRCI: speech-to-silence ratio, silent pause mean duration, and filled pause mean duration. The combination of these three parameters enabled the classification of breast cancer survivors with CRCI, survivors without cognitive complaints, and controls with an accuracy of 73.7%. Analysis of the eye-tracking data showed a significant delay between the first gaze at an object and its naming in cancer survivors. This result suggests that survivors struggle to access their lexical stock. Moreover, the fixation maps suggest that cognitive deficits in CRCI may be more associated with a deficit in executive control than in memory problems. However, heatmap analyses do not rule out possible attentional deficits affecting lexical search. This dissertation offers new insights into speech-based diagnostic tools. This research will be extended through further studies on other cancer types and conditions with subtle cognitive impairment.

Chapter Speech and Oculomotor Features of Cancer-Related Cognitive Impairment: A Comprehensive English Summary includes a comprehensive summary of this dissertation in English.

Table des matières

Contribution des auteurs 2

Introduction **5**

1 Question de recherche et objectifs 6

2 Hypothèses 7

3 Structure 7

Cadre théorique **8**

1 Cancers du sein 8

1.1 Généralités 8

1.2 L'après-cancer 10

2 Trouble cognitif induit par un cancer et ses traitements 10

2.1 Facteurs de risque d'apparition du CRCI 12

2.2 Modifications cérébrales structurelles et fonctionnelles associées au CRCI . . 19

2.3 Description des difficultés cognitives 21

2.4 Évaluation subjective du CRCI 22

2.5 Évaluations neuropsychologiques 24

2.6 Discordance entre difficultés ressenties et tests neuropsychologiques 26

3 La parole : une ouverture sur la cognition 29

3.1 De la pensée aux mots : modèles de production de la parole 29

3.2 Substrats neuronaux du langage 30

3.3 Anomie : Comprendre la cognition par l'erreur 33

3.4 Interférences lexicales sur la prosodie 34

3.5 Interruptions du flux de parole 35

3.6 Limites de l'analyse de la parole 40

4 Investiguer les processus linguistiques par le regard 41

4.1 Principes fondamentaux du comportement oculomoteur 42

4.2 Interaction entre parole et recherche visuelle 44

5 Manifestations du CRCI à travers la parole et le regard 47

Méthodologie **48**

1 Démarches éthiques 48

2 Population 48

2.1 Conditions d'éligibilité 49

2.2 Cohorte 49

3 Méthode d'analyse de la parole 49

3.1 Tâches de narration . . . 50
3.2 Matériel et conditions d'enregistrement . . . 51
3.3 Paramètres linguistiques . . . 51
3.4 Analyse des pauses . . . 53
4 Étude du comportement oculomoteur . . . 56
4.1 Matériel d'oculométrie . . . 57
4.2 Tâche narrative sur écran . . . 58
4.3 Extraction et analyse des données oculomotrices . . . 58
5 Évaluation cognitive et physiologique, psychologique . . . 61
5.1 Évaluation cognitive . . . 61
5.2 Évaluation psychologique . . . 62
5.3 Évaluation de la fatigue . . . 62
6 Modèle d'apprentissage machine . . . 62
7 Analyses statistiques . . . 63
7.1 Variables démographiques et cofacteurs . . . 63
7.2 Variables de parole . . . 63

Investiguer les troubles subtils de la cognition **65**
1 Contribution de l'article à la thèse . . . 65
2 Résumé de l'article . . . 65
2.1 Objectifs . . . 66
2.2 Méthode . . . 66
2.3 Principaux résultats . . . 66
2.4 Discussion . . . 67
3 Apport de l'article à la recherche . . . 68
Article : **Linguistic markers of subtle cognitive impairment in connected speech: A Systematic Review** . . . 69

Dévoiler les marqueurs prosodiques du CRCI **99**
1 Contribution de l'article à la thèse . . . 99
2 Résumé de l'article . . . 99
2.1 Objectifs . . . 100
2.2 Méthode . . . 100
2.3 Principaux résultats . . . 101
2.4 Discussion . . . 102
3 Apport de l'article à la recherche . . . 103
Article : **Using machine learning to identify speech markers of reported cognitive impairment in breast cancer survivors** . . . 104

Comprendre l'anomie en contexte de CRCI grâce à l'oculométrie **121**
1 Contribution de l'article à la thèse . . . 121

2 Résumé de l'article . . . 121
2.1 Introduction . . . 121
2.2 Objectifs . . . 122
2.3 Méthode . . . 123
2.4 Principaux résultats . . . 124
2.5 Discussion . . . 125
3 Apport de l'article à la recherche . . . 125
Article : **Oculomotor behavior of breast cancer survivors reporting cognitive changes** 127

Discussion générale **146**
1 Indices prosodiques d'un CRCI . . . 146
1.1 Interruptions silencieuses . . . 146
1.2 Relation entre silence et facteurs confondants . . . 149
1.3 Continuum des changements cognitifs induits par un cancer . . . 150
2 Corrélats cognitifs de l'anomie en contexte de CRCI . . . 151
2.1 Délai de dénomination . . . 151
2.2 Parole continue vs. fluence verbale . . . 153
2.3 L'anomie révèle un trouble plus large . . . 156
3 Interventions cliniques pour le traitement du CRCI . . . 156
3.1 Thérapies actuelles . . . 157
3.2 Pistes d'intervention en orthophonie . . . 159
4 Limites et perspectives . . . 160

Conclusion **162**

Bibliographie **166**

Annexes **198**
Annexe 1 : **Assessing Cancer-Related Cognitive Impairment for breast cancer survivors with speech analysis** . . . 199
Annexe 2 : **Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment** . . . 223
Annexe 3 : **Speech markers of Cancer-Related Cognitive Impairment : A pilot study** . 240
Annexe 4 : **Self-interruptions in Breast Cancer Patients Who Complain of Anomia** . . 246

Comprehensive English Summary **252**

Introduction **256**

Theoretical Framework **259**
1 Introduction . . . 259
2 Breast Cancer . . . 259
3 Cancer-Related Cognitive Impairment . . . 261

4 Speech: A Gateway into Cognition 275
5 Investigating Linguistic Processes Through Eye Behavior 285

General Discussion **291**
1 Prosodic Signs of CRCI 291
2 Clinical Interventions for CRCI Treatment 298
3 Limitations and Perspectives 301

Conclusion **303**

Table des figures

Cadre théorique **8**

1 Nombre d'études référencées sur PubMed portant sur le CRCI 11
2 Estimation de la prévalence du CRCI en fonction du parcours de soin 12
3 Schéma des potentiels mécanismes du CRCI induits par la chimiothérapie 14
4 Schéma des potentiels mécanismes impliqués dans le développement du CRCI . . . 15
5 Schéma des facteurs prédictifs d'un CRCI 18
6 Schéma des mécanismes ayant un rôle dans le développement d'un CRCI 18
7 Synthèse des études montrant les régions cérébrales associées au CRCI 21
8 Schéma des processus linguistiques impliqués dans la production de la parole 30
9 Spectrogrammes de voyelles prolongées 39
10 Spectrogrammes de pauses 39
11 Classification des événements oculomoteurs 42
12 Schéma des mesures oculomotrices 43
13 Scène visuelle et cartes de chaleur correspondantes 46

Méthodologie **48**

1 Images présentées lors de la tâche narrative 50
2 Exemple d'erreur d'alignement 55
3 Configuration de l'expérience de l'étude DisCCo 57
4 Alignement de la parole et des mouvements oculaires dans ELAN 58
5 Calcul du délai *Eye-Voice Span* 59
6 Carte de fixation, extraits de l'image globale, carte de saillance et carte sémantique . 60

Article 1 **65**

1 Literature screening and selection flowchart following PRISMA guidelines 81
2 Number of studies per pathology per year 82
3 Number of participants and cognitive function assessed across studies 83
4 Linguistic feature frequency according to case population 84
5 Linguistic feature frequency according to type of discourse 85
6 Quality assessment ratings 87
7 Summary of the quality assessment 88

Article 2 **99**

1 Canonical spectrograms of silent pauses, filled pauses, and sustained vowels 109
2 Steps of the speech analysis 110
3 Graphical illustration of the machine learning pipeline 111
4 Heatmap of the results of the machine learning models 114
5 Correlation and classification errors, distribution and density of the three-group model's results 116

Article 3 **121**
1 Picture-based narrative task, scene patch, saliency map and meaning map 132
2 Non-verbal control task 134
3 Comparison of the time spent on pictures 136
4 Example of the time overlaid results for a typical control participant 138
5 Example of the spatial overlaid results for a typical control participant 139
6 Distribution of Eye-Voice Span duration 140

Annexe 1 **199**
1 Data ordering from four participants in RStudio 209
2 Overlapping distribution of mean silent pause duration 212
3 Distribution of mean duration of silent pauses 212
4 Boxplots showing mean silent pause duration for each group 213
5 Correlation matrix of the relationship between the mean silent pause duration and co-factors 215

1 Possible mechanisms underlying cognitive impairment associated with hormone, immune, and targeted therapy. 264
2 Schematic of factors predictive of CRCI 266
3 Synthesis of studies showing an association between brain regions and CRCI. Blue dots indicate cerebral hypoactivation. Red dots indicate cerebral hyperactivation. . . 268
4 Diagram of the linguistic processes involved in speech production 276
5 Visual scenes and corresponding heatmaps 288

Liste des tableaux

Cadre théorique **8**

Méthodologie **48**

1 Paramètres prosodiques 52
2 Exemples d'erreurs de transcription automatique 53

Article 1 **65**

1 Types of data extracted 75
2 Proposal for a harmonized terminology of linguistic markers 77

Article 2 **99**

1 Description of the speech variables characteristics and measures 107
2 Characteristics of the population and scores at neuropsychological tests 113
3 Results of the two-class machine learning model 114
4 Results of the two-class machine learning model 115

Article 3 **121**

1 Population characteristics 135
2 Summary of the overlap comparison 137
3 Overview of a qualitative analysis of lexical search events 143
4 List of narrative features 144

Annexe 1 **199**

1 Characteristics of the population 211
2 Mean pause duration 214
3 Examples of transcription errors using the Whisper large v3 model 217

Contribution des auteurs

Cette section déclare les contributions de la doctorante et celles ses collaboratrices et collaborateurs à l'élaboration de cette thèse. La taxonomie (*Contributor Role Taxonomy*) CRediT est utilisée pour identifier la contribution de chaque co-auteur (Cf. https://credit.niso.org/).

Contribution de la doctorante

Amélie Richard, pour l'ensemble de la thèse, des articles et des actes de communication : Conceptualisation, Méthodologie, Développement informatique, Analyse formelle, Fourniture de ressources, Curation des données, Rédaction – Version originelle, Rédaction – Révision et correction, Visualisation, Administration du projet.

Contribution des collaboratrices et collaborateurs

Maude Beaudoin, pour l'article « Oculomotor behavior of breast cancer survivors reporting cognitive changes » : Supervision.

Anne Cheylus, pour l'article « Oculomotor behavior of breast cancer survivors reporting cognitive changes » : Développement informatique.

Alexandre Foncelle, pour l'article « Using machine learning to identify speech markers of reported cognitive impairment in breast cancer survivors », ainsi que pour les actes de communication « Speech markers of Cancer-Related Cognitive Impairment: A pilot study » et « Self-interruptions in Breast Cancer Patients Who Complain of Anomia » : Analyse formelle, Curation des données, Rédaction – Révision et correction.

Fabrice Hirsch, pour l'article « Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment », ainsi que pour les actes de communication « Speech markers of Cancer-Related Cognitive Impairment: A pilot study » et « Self-interruptions in Breast Cancer Patients Who Complain of Anomia » : Rédaction – Révision et correction.

Sophie Jacquin-Courtois, pour les articles « Using machine learning to identify speech markers of reported cognitive impairment in breast cancer survivors », « Assessing Cancer-Related Cognitive Impairment for breast cancer survivors with speech analysis et « Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment », ainsi que pour les actes de communication « Speech markers of Cancer-Related Cognitive Impairment: A pilot study » et « Self-interruptions in Breast Cancer Patients Who Complain of Anomia » : Fourniture de ressources, Supervision. Pour l'article « Assessing Cancer-Related Cognitive Impairment for breast cancer survivors with speech analysis » : Conceptualisation, Rédaction – Révision et correction, Supervision. Pour l'article « Oculomotor behavior of breast cancer survivors reporting cognitive changes » : Supervision.

Manon Lelandais, pour l'article « Linguistic markers of subtle cognitive impairment in connected speech: A Systematic Review » : Conceptualisation, Méthodologie, Curation des données, Rédaction – Version originelle, Rédaction – Révision et correction, Supervision. Pour l'article « Assessing Cancer-Related Cognitive Impairment for breast cancer survivors with speech analysis », ainsi que pour l'acte de communication « Speech markers of Cancer-Related Cognitive Impairment: A pilot study » : Rédaction – Version originelle, Rédaction – Révision et correction.

Pauline Mouchès, pour l'article « Using machine learning to identify speech markers of reported cognitive impairment in breast cancer survivors » : Méthodologie, Analyse formelle, Curation des données, Rédaction – Révision et correction, Supervision.

Laure Pisella, pour l'article « Oculomotor behavior of breast cancer survivors reporting cognitive changes » : Rédaction – Révision et correction.

Karen T. Reilly pour l'article « Linguistic markers of subtle cognitive impairment in connected speech: A Systematic Review » et les actes de communication « Speech markers of Cancer-Related Cognitive Impairment: A pilot study » et « Self-interruptions in Breast Cancer Patients Who Complain of Anomia » : Conceptualisation, Rédaction – Révision et correction, Supervision.

LES 14 TYPES DE CONTRIBUTION CREDIT

- **Conceptualisation** : Idée, formulation ou évolution de la définition des objectifs de la recherche
- **Méthodologie** : Développement ou conception des méthodes ; création de modèles
- **Développement informatique** : Programmation, développement de logiciels ; conception de programmes informatiques ; mise en œuvre du code informatique et des algorithmes qui le sous-tendent ; test des éléments du code
- **Validation** : vérification, comme une partie de l'activité ou séparément, de la reproduction/reproductibilité globale des résultats et des expérimentations ainsi que des autres produits de recherche
- **Analyse formelle** : Application de techniques statistiques, mathématiques, informatiques ou d'autres techniques formelles pour analyser ou synthétiser les données de l'étude
- **Recherche** : Conduite d'un processus de recherche et d'enquête, spécifiquement en réalisant les expériences ou en recueillant des données ou des éléments de preuve
- **Fourniture de ressources** : Mise à disposition de matériels d'étude, de réactifs, de matériaux, de patients, d'échantillons de laboratoire, d'animaux, d'instruments, de ressources informatiques, ou d'autres outils d'analyse
- **Curation des données** : Activités de gestion des données pour annoter les données (produire des métadonnées), nettoyer les données et mettre à jour les données de recherche (y compris le code logiciel lorsqu'il est nécessaire pour interpréter les données elles-mêmes) à des fins de première utilisation et de réutilisation
- **Rédaction – Version originelle** : Préparation, création et/ou présentation du travail publié, spécifiquement rédaction de la version initiale (y compris la traduction substantielle)
- **Rédaction – Révision et correction** : Préparation, création et/ou présentation du travail publié par les membres de l'équipe originelle de recherche, spécifiquement examen critique, commentaire ou révision – incluant les étapes préalables ou postérieures à la publication
- **Visualisation** : Préparation, création et/ou présentation du travail publié, spécifiquement visualisation et présentation des données
- **Supervision** : Responsabilité de surveillance et de direction de la planification, et de l'exécution de la recherche, incluant le mentorat externe à l'équipe principale
- **Administration du projet** : Responsabilité de gestion et de coordination de la planification, et de l'exécution de la recherche
- **Obtention du financement** : Obtention du soutien financier pour le projet menant à la publication

Introduction

La parole regorge d'éléments dont la granularité permet de saisir le fonctionnement de la cognition humaine. Son analyse est une approche prometteuse pour le repérage de changements cognitifs subtils. En particulier, l'identification de variables prosodiques pourrait aider au diagnostic de troubles cognitifs, que les tests orthophoniques ou neuropsychologiques actuels peinent à révéler. Tel est le cas du trouble cognitif induit par un cancer et ses traitements (*Cancer-Related Cognitive Impairment*, CRCI), qui se manifeste par une symptomatologie discrète, malgré ses répercussions négatives sur la qualité de vie des patients. Les patients en oncologie et les survivants d'un cancer décrivent des difficultés de mémoire, de maintien attentionnel, des difficultés à effectuer des tâches routinières et à s'exprimer de manière fluide. Chez cette population, le manque du mot est l'une des inquiétudes les plus courantes, indépendamment du type de cancer ou de ses traitements. De manière générale, ces altérations fonctionnelles interfèrent significativement sur la vie courante des patients. Par conséquent, leur détection précoce est indispensable afin de limiter leurs effets à long terme et de proposer une prise en charge adaptée aux besoins le cas échéant.

La présence d'un CRCI peut être attestée par des questionnaires standardisés, qui recueillent les symptômes vécus quotidiennement par les patients. En revanche, les évaluations psychométriques révèlent des résultats inconsistants, dont les scores demeurent souvent dans les normes. Il en résulte une discordance entre les difficultés perçues et les scores obtenus qui limite l'interprétation clinique des tests neuropsychologiques. De fait, le caractère subtil de ces symptômes entraîne un manque de reconnaissance par le corps médical. Or, en l'absence de diagnostic clair, les patients ne peuvent bénéficier des soins appropriés. Les patients peuvent avoir tendance à compenser ou à minimiser les manifestations de ce handicap « invisible » (Lewis et al., 2016). Par ailleurs, la diversité des éventuels déficits cognitifs associés à ce trouble (qu'ils soient d'ordre mnésique, attentionnel, linguistique, etc.) complique la compréhension des mécanismes sous-jacents au CRCI. En d'autres termes, bien que les patients décrivent une vaste palette de troubles cognitifs, allant de la mémoire au langage, en passant par les fonctions exécutives et l'attention, il reste difficile de cerner précisément la nature et l'étendue de ces altérations. Les connaissances lacunaires des mécanismes du CRCI freinent le développement d'interventions thérapeutiques adaptées.

Le langage interagit avec les autres fonctions cognitives, telles que la mémoire, les processus attentionnels, le contrôle exécutif et les fonctions motrices. Dès lors, son analyse offre la possibilité d'investiguer l'activité des processus cognitifs. Plus précisément, étudier la parole de survivants d'un cancer pourrait permettre de mieux comprendre les difficultés qu'ils décrivent. En effet, parler est un acte quotidien, dont les variations, même légères, peuvent renseigner sur le statut cognitif

des locuteurs. Cependant, les productions orales sont sujettes à des variabilités interindividuelles importantes. Leur analyse présente une complexité inhérente aux propriétés du langage. À elle seule, la parole ne permet pas toujours d'interpréter avec certitude les processus cognitifs qui sous-tendent sa production. Par conséquent, l'ajout d'une approche complémentaire telle que l'oculométrie renforce les connaissances sur la production de la parole.

Les perspectives qu'apporte l'analyse de la parole pour révéler la présence d'un CRCI sont encourageantes. Toutefois, sa mise en application soulève plusieurs questionnements d'ordre méthodologique.

Tout d'abord, seules deux études ont réalisé une analyse psycholinguistique dans le cadre d'un CRCI (Myers et al., 2022 ; K. Williams et al., 2021). Ces travaux se sont focalisés sur la complexité syntaxique des discours de patients guéris d'un cancer de la prostate. S'ils soulignent l'intérêt de l'analyse des paramètres linguistiques pour la détection du CRCI, l'angle linguistique choisi, à savoir l'étude de la syntaxe, semble toutefois introduire certains biais inter-juges. Or, ces facteurs compromettent à la fois la faisabilité et la réplicabilité de telles recherches.

A contrario, l'analyse de la prosodie repose sur la mesure du signal acoustique. Les nombreux outils informatisés et automatisés consacrés au traitement de signal acoustique, désormais disponibles, font de l'analyse de la prosodie une méthode de choix quant à la détection du CRCI. Cependant, la littérature sur les paramètres linguistiques susceptibles de détecter des troubles de la cognition est disparate. La variété des pathologies étudiées et la diversité des variables examinées empêchent la sélection de marqueurs potentiels du CRCI.

1 Question de recherche et objectifs

Les difficultés vécues par les patients sont bien documentées à travers les études qualitatives. Les entretiens menés auprès de survivants d'un cancer révèlent que le manque du mot (ou anomie) est une source d'inquiétude répandue à travers cette population. De plus, cette difficulté linguistique est souvent associée à la description de problèmes mnésiques. À ce jour, aucune étude spécifique n'a été conduite sur ce phénomène dans un contexte du CRCI. Par conséquent, l'origine cognitive de ce manque du mot demeure inconnue. Puisque les patients se plaignent de trouble de la mémoire, il serait tentant de penser que leur anomie découlerait d'une dégradation des connaissances sémantiques. Néanmoins, les observations cliniques ainsi que les études utilisant des techniques de neuro-imagerie ne soutiennent pas cette hypothèse. Au contraire, ces recherches suggèrent plutôt la présence de déficits attentionnels ou exécutifs. Étant donné que la prise en charge d'un manque du mot est déterminée par ses substrats cognitifs, un éclairage sur ce phénomène très fréquent semble requis. En outre, les effets du manque du mot sur la production discursive des patients constituent un point d'entrée intéressant pour déchiffrer les mécanismes sous-jacents au CRCI.

L'objectif de cette thèse est double : (i) Identifier les marqueurs prosodiques pouvant indiquer la présence du CRCI chez des survivantes d'un cancer du sein, et (ii) examiner les processus cognitifs sous-jacents à ce trouble à travers l'étude du manque du mot.

2 Hypothèses

La parole des survivantes d'un cancer du sein pourrait contenir des marqueurs prosodiques distinctifs des difficultés cognitives auxquelles elles sont confrontées. Pour tester cette hypothèse, une présélection de paramètres sera incorporée dans des modèles d'apprentissage automatique. Ces modèles seront capables de classer des survivantes rapportant des difficultés cognitives, des survivantes n'ayant pas de plainte, ainsi que des sujets témoins.

Par ailleurs, le manque du mot rapporté par les survivantes du cancer du sein pourrait résulter d'un déficit attentionnel, exécutif ou mnésique. Cette hypothèse fera l'objet d'une investigation du comportement oculomoteur lors d'une tâche narrative. Cette approche visera à déterminer dans quelle mesure les mécanismes cognitifs précités influencent les difficultés de recherche lexicale observées chez ces patientes.

3 Structure

Le Cadre théorique présente d'abord un état de l'art sur le trouble cognitif induit par un cancer et ses traitements, avant d'offrir un aperçu du lien qu'entretient le langage avec la cognition et de poursuivre l'investigation des processus de la parole par une technique d'oculométrie. La Méthodologie détaille le cadre expérimental de cette thèse, et justifie le choix des modèles statistiques et d'apprentissage automatique utilisés pour examiner les variables prosodiques. Le chapitre « Investiguer les troubles subtils de la cognition » passe en revue les paramètres linguistiques couramment utilisés pour l'évaluation de troubles cognitifs subtils. Le chapitre « Dévoiler les marqueurs prosodiques du CRCI » propose un ensemble de marqueurs prosodiques capables de déceler la présence d'un CRCI grâce à des modèles d'apprentissage automatique. Le chapitre « Comprendre l'anomie en contexte de CRCI grâce à l'oculométrie » éclaire le phénomène du manque du mot en contexte de CRCI en explorant le comportement oculomoteur lors de la production d'un discours. Enfin, les résultats et perspectives de ces chapitres expérimentaux seront abordés dans un contexte plus général au sein de la Discussion générale.

Cadre théorique

Les cancers sont des pathologies hétérogènes répandues à travers le monde. Selon le Centre International de Recherche sur le Cancer, environ vingt millions de nouveaux cas ont été diagnostiqués en 2022 (Ferlay et al., 2024). En France, l'Institut National du Cancer estime l'incidence des cancers diagnostiqués à 443 136 cas par an et sa prévalence (c.-à-d. le nombre de personnes vivant après un cancer) à 3,8 millions (Institut National du Cancer, 2024). Grâce à l'avancée des connaissances et des technologies dans la recherche contre les cancers, certaines tumeurs malignes telles que les cancers du sein sont de mieux en mieux diagnostiquées et traitées. Ainsi, le nombre de survivants augmente et la qualité de vie après-cancer se positionne désormais parmi les enjeux majeurs de santé publique (Institut National du Cancer, 2021). Aux côtés de l'alopécie, la fatigue ou d'autres effets secondaires bien documentés, le trouble cognitif induit par un cancer et ses traitements est une séquelle des cancers fréquente et invalidante. Ce trouble aussi appelé, *Cancer-Related Cognitive Impairment* (CRCI) ou parfois *Cancer-Related Cognitive Changes*, fait référence à un panel de difficultés cognitives qui affectent le fonctionnement quotidien des survivantes. Le CRCI est un syndrome polyfactoriel dont les mécanismes cognitifs sous-jacents sont encore mal compris.

Ce cadre théorique se divise en quatre sections : La section « Cancers du sein » présente les connaissances actuelles sur les cancers du sein. La section « Trouble cognitif induit par un cancer et ses traitements » est dédiée à un état de l'art sur les séquelles cognitives liées aux cancers. La section « De la pensée aux mots : modèles de production de la parole » met en évidence les avantages de l'analyse de la parole pour étudier la cognition et ses troubles. Enfin, la section « Investiguer les processus linguistiques par le regard » aborde l'intérêt de l'oculométrie comme méthode de mesure comportementale complémentaire à l'analyse de la parole.

1 Cancers du sein

1.1 Généralités

D'après les données du Centre International de Recherche sur le Cancer, le cancer du sein était le cancer le plus diagnostiqué au monde en 2020, comptant plus de 2,26 millions de nouveaux cas par an (Ferlay et al., 2024). À l'échelle nationale, il reste également le cancer le plus fréquent chez les femmes et demeure la deuxième cause de décès par morbidité chez cette population (Institut

National du Cancer, 2024), juste derrière les maladies de l'appareil cardio-vasculaire. Pour autant, l'amélioration du dépistage, du diagnostic et des traitements font de ce cancer l'un des mieux soignés de France. En effet, les chances de survie d'un cancer du sein cinq ans après son diagnostic s'élèvent à 88% (Institut National du Cancer, 2018, 2024).

Une cellule cancéreuse se caractérise par son indépendance au regard des signaux qui contrôlent la croissance et la division cellulaire, sa capacité à résister au processus d'apoptose (c.-à-d. de mort cellulaire), sa faculté à migrer (c.-à-d. à se déplacer d'un environnement à un autre) et envahir les tissus adjacents (c.-à-d. de coloniser les cellules des autres environnements) (Brown et al., 2023). La classification moléculaire de référence des cancers du sein fait état d'au moins quatre sous-types principaux de cancer du sein (Roy et al. [2023], voir Organisation mondiale de la santé [2019] pour une classification complète) : les cancers luminaux A et B, les cancers HER2 surexprimé et les cancers triple négatif. Les cancers du sein luminal A et luminal B sont des cancers hormono-sensibles (RH+). Ils expriment des récepteurs à l'œstrogène (RE+) et à la progestérone (RP+) et concernent environ 70% des cancer du sein. Les cancers HER2 surexprimé (HER2+) ou HER2-like, quant à eux, possèdent un récepteur qui favorise la multiplication des cellules cancéreuses. Ils concernent environ 15% des cancers du sein. Enfin, la catégorie des cancers Basal-Like et triple négatif fait référence à l'absence des récepteurs hormonaux (RE- et RP-) ou du récepteur HER2 (HER2-). Ils représentent environ 15% des cancers du sein. Le choix des traitements chirurgicaux et pharmacologiques est guidé par cette classification, de même que la localisation de la tumeur, son stade de sévérité (c.-à-d. sa taille et l'éventuelle présence de métastases) et l'état de santé des patientes.

En fonction de la taille de la tumeur, l'intervention chirurgicale peut consister en une tumorectomie ou, dans certains cas, une mastectomie totale, comprenant l'ablation du mamelon. La technique dite du « ganglion sentinelle » consiste à analyser le premier ganglion lymphatique du sein affecté (Giammarile et al., 2022). Si celui-ci contient des cellules cancéreuses, un curage ganglionnaire sera alors envisagé. La chirurgie est généralement proposée en première ligne, suivie de traitements pharmacologiques adjuvants. Parmi ces traitements, la chimiothérapie, la radiothérapie, l'hormonothérapie, les thérapies ciblées et l'immunothérapie sont les plus couramment prescrits dans la prise en charge du cancer du sein (Burguin et al., 2021 ; Kerr et al., 2022).

Administrée en néo-adjuvant (avant la chirurgie), la chimiothérapie permet de réduire la taille de la tumeur avant l'intervention. En traitement adjuvant (après la chirurgie), elle empêche la prolifération des cellules cancéreuses. Elle est le plus souvent délivrée par voie intraveineuse sur une période de trois à six mois. Diverses molécules et traitements, tels que les anthracyclines (p. ex., doxorubicine, épirubicine), les taxanes (p. ex., paclitaxel [Taxol]) et le 5-fluoro-uracile (ou 5-FU), sont prescrits en fonction du sous-type de cancer et de son degré de sévérité (Waks & Winer, 2019).

La radiothérapie externe est utilisée pour détruire les cellules cancéreuses. Les séances, qui sont de courte durée, se déroulent en ambulatoire sur une période d'un à deux mois. L'hormonothérapie est souvent indiquée comme traitement adjuvant (en complément de la chirurgie, selon l'analyse des récepteurs hormonaux), afin de diminuer les risques de rechute. Ce traitement, généralement d'une

durée de cinq ans, peut parfois être prolongé jusqu'à dix ans après la prise initiale. L'hormonothérapie, prise sous forme de comprimés, bloque les récepteurs aux œstrogènes. Parmi les traitements hormonaux disponibles, les anti-œstrogènes (p. ex., tamoxifène), les inhibiteurs de l'aromatase (p. ex., anastrazole, letrozole, exemestane) et les analogues de la LH-RH sont des options souvent recommandées (Kerr et al., 2022).

Les thérapies ciblées ont pour objectif d'inhiber ou de ralentir l'action des protéines responsables de la croissance des cellules cancéreuses. Elles sont administrées par voie intraveineuse, généralement pendant une période de six à douze mois. L'immunothérapie, forme de thérapie ciblée, renforce le système immunitaire afin d'améliorer sa capacité à lutter contre les cellules cancéreuses (Debien et al., 2023). Les traitements sont délivrés par voie intraveineuse sur une période de trois à six semaines. Des anticorps monoclonaux, tel que le trastuzumab font partie des traitements disponibles (Kerr et al., 2022).

1.2 L'après-cancer

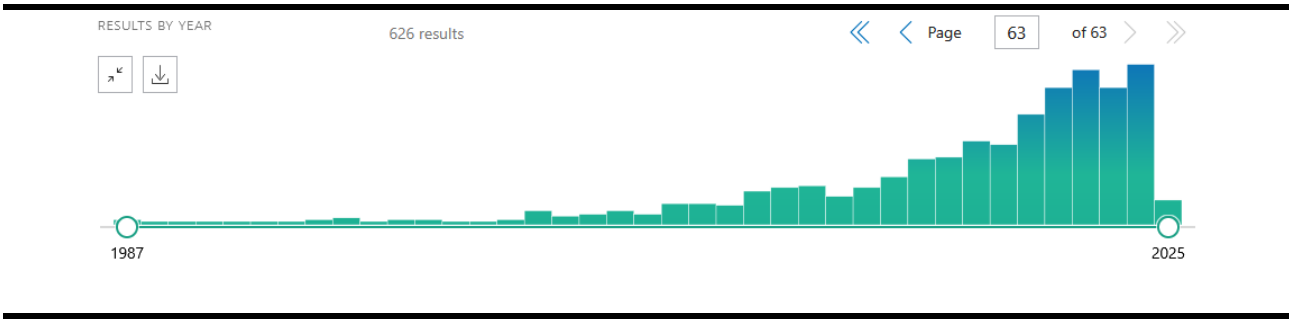
Les effets secondaires des traitements anticancéreux peuvent avoir un impact à court et à long terme sur la qualité de vie des survivantes et survivants. Le rapport VICAN-5, qui évalue la vie et le statut fonctionnel des patients en France cinq ans après le diagnostic d'un cancer, met ainsi en évidence un certain nombre de séquelles fréquemment rencontrées par les patientes guéries d'un cancer du sein (Institut National du Cancer, 2018). Parmi celles-ci, figurent des séquelles d'ordre esthétique, une fatigue persistante, des douleurs chroniques, des troubles psychologiques, ainsi que des perturbations gynécologiques et fonctionnelles (Delrieu et al., 2021). En plus de ces effets, des difficultés cognitives peuvent survenir au cours du parcours de soin en oncologie et perdurer après la rémission ou guérison d'un cancer.

2 Trouble cognitif induit par un cancer et ses traitements

Étudié depuis les années 1980, le trouble cognitif induit par un cancer et ses traitements, ou *Cancer-Related Cognitive Impairment* (désormais CRCI) gagne un intérêt croissant depuis une dizaine d'années. La figure 1 illustre la hausse constante d'études sur le sujet référencées sur la base de données PubMed. L'appellation CRCI est généralement attribuée aux tumeurs solides et hématopoïétiques (c.-à-d. cancers du sang) hors cancers du Système Nerveux Central. En effet, les tumeurs du cerveau (p. ex., glioblastomes, oligodendrogliomes) sont très susceptibles d'entraîner des dysfonctionnements cognitifs par leur localisation même (Meskal et al., 2016 ; Patchell, 2003). En outre, le CRCI est principalement documenté chez les femmes ayant un cancer du sein du fait de son incidence élevée et de son importante prévalence, c'est-à-dire, ses meilleures chances de guérison. Pour autant, ce trouble cognitif n'est pas spécifique à cette population. Des patients ayant d'autres

types de cancer, tels qu'un cancer du testicule, de la prostate, du poumon, du colon, etc., peuvent également faire face à des difficultés cognitives (Demos-Davies et al., 2024).

Figure 1. Nombre d'études référencées sur PubMed portant sur le CRCI. Résultat de l'équation de recherche : (((cancer[Title/Abstract]) OR (neoplasm[Title/Abstract]) OR (tumor[Title/Abstract])) AND ((cognitive dysfunction[Title/Abstract]) OR (cognitive impairment[Title/Abstract]) OR (chemobrain[Title/Abstract]) OR (chemofog[Title/Abstract]) OR (cognitive change[Title/Abstract])))

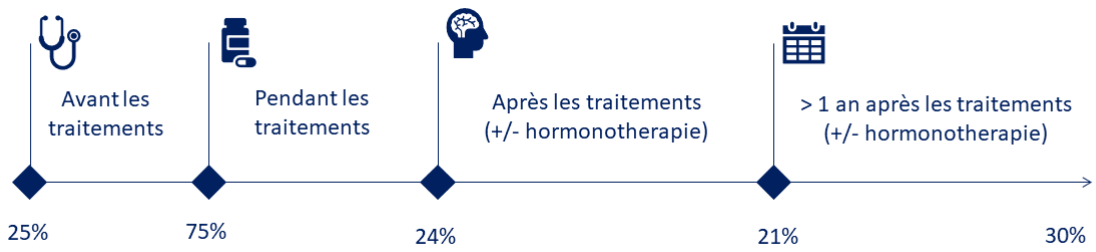


Le CRCI fait référence à un ensemble de difficultés cognitives décrites par les patients en oncologie dont les symptômes dégradent leur qualité de vie. L'autonomie des survivants est généralement préservée. L'expression clinique du CRCI peut être discrète, tandis que ses répercussions fonctionnelles sur leur vie quotidienne sont importantes. Les survivants d'un cancer rapportent souvent, par exemple, une plainte de mémoire à court-terme, des problèmes de concentration et des oublis de mots. Ces difficultés, qui perturbent leurs interactions sociales et gênent la réalisation de tâches pourtant routinières (voir section Description des difficultés cognitives) concerneraient jusqu'à 75% des patientes ayant un cancer du sein (Lange et al., 2019). Cependant, il est possible que ces chiffres soient sous-estimés en raison du manque d'outils diagnostics suffisamment sensibles pour détecter un CRCI (voir section Évaluations neuropsychologiques).

La temporalité du CRCI diffère selon les patients, leur parcours de soin, leur type de cancer et les traitements administrés (voir p. ex., l'étude longitudinale de Yusuf et al. [2022] sur le CRCI chez des patients ayant un myélome multiple [cancer du sang]). En ce qui concerne la prévalence du CRCI chez les femmes ayant un cancer du sein, une revue systématique montre qu'un quart des patientes présentent un trouble cognitif avant l'administration des traitements adjuvants, 24% immédiatement après la fin des traitements et 21% six mois après (Dijkshoorn et al., 2021). Si de plus en plus d'études mettent en évidence la présence d'un changement cognitif avant la prise des traitements adjuvants (Ahles et al., 2008), l'expression du trouble est, dans la majorité des cas, la plus élevée six mois après l'arrêt des médicaments anticancéreux (Janelsins et al., 2018). Cette fenêtre temporelle correspond à une période à laquelle les patients reprennent leurs activités et se trouvent alors confrontés de manière plus tangible à leurs difficultés cognitives. Pour la plupart des patients, la plainte cognitive a tendance à décroître au cours des deux ans suivant l'arrêt des traitements adjuvants, soit du fait de la régression spontanée du CRCI, soit en raison des stratégies de *coping*, c.-à-d. d'adaptation face au changement cognitif, adoptées par les patients. En revanche, pour une partie d'entre eux, le CRCI est un symptôme qui perdure au-delà de deux ans après l'arrêt des traitements (Whittaker et al., 2022), voire jusqu'à dix ans après (De Ruiter et al., 2011). Ces patients décrivent l'impression de ne jamais avoir récupéré

leurs capacités cognitives d'avant cancer. À l'heure actuelle, il n'existe pas d'outil clinique permettant d'identifier précocement le sous-groupe de patients à risque de développer des symptômes persistants ou de déterminer les facteurs de vulnérabilités préalables. La figure 2 schématise la prévalence estimée du CRCI à différents moments du parcours de soin.

Figure 2. Estimation de la prévalence du CRCI en fonction du parcours de soin.



Que ce soit à court-terme comme à long-terme, le CRCI interfère avec la vie des patients (Boykoff et al., 2009). Une étude menée en France révèle que 21% des survivantes d'un cancer du sein n'avaient pas repris le travail deux ans après guérison, bien que plus de la majorité des survivantes soit en âge de travailler (Dumas et al., 2020 ; Institut National du Cancer, 2018). Parmi les freins au retour à l'emploi, les effets psychologiques et neuropsychologiques secondaires au cancer et à ses traitements sont souvent mis au premier plan. En effet, les patientes émettent des craintes quant à leur reprise professionnelle au regard de la détérioration de leurs capacités cognitives (Munir et al., 2010). Par ailleurs, les patientes qui avaient repris leur activité ont rapporté ne plus être aussi efficace à leur travail qu'avant le cancer (Lewis & Mackenzie, 2022). De la même manière, le CRCI entraîne des répercussions sur leur vie privée et favorise leur risque d'isolement social à long-terme (Haywood et al., 2023 ; Player et al., 2014).

2.1 Facteurs de risque d'apparition du CRCI

Pour l'heure, les origines et mécanismes du trouble cognitif induit par un cancer et ses traitements demeurent mal compris. Cette section décrit les principaux facteurs impliqués dans l'apparition et/ou l'aggravation du CRCI.

Biologie du cancer

Le CRCI a longtemps été étiqueté sous l'appellation *chemofog* ou *chemobrain* en raison de l'association faite de ses symptômes à la chimiothérapie (Hurria et al., 2007). Cependant l'avancée de la recherche a démontré que le CRCI pourrait concerner près de 40% des patients avant même le début des traitements (Demos-Davies et al., 2024). En effet, d'autres facteurs, tels que les autres

traitements oncologiques ou le cancer en lui-même, joueraient un rôle dans son développement (Olson & Marks, 2019). Wefel et al. (2011) ont par exemple mis en avant l'existence d'une plainte et d'un dysfonctionnement cognitif chez des hommes diagnostiqués pour un cancer du testicule avant l'administration des traitements. Près de la moitié de leurs soixante-neuf patients présentaient un déficit cognitif dans les domaines des fonctions motrices, de l'apprentissage verbal et des fonctions exécutives (c.-à-d. inhibition, mise à jour et flexibilité mentale [Miyake & Friedman, 2012]) dès l'établissement de la ligne de base (Wefel et al., 2011). Des résultats similaires avaient été relevés par Quesnel et al. (2009) chez des femmes soignées pour un cancer du sein. Dans cette étude longitudinale, les patientes témoignaient de performances cognitives plus faibles que les contrôles en amont de la chimiothérapie et de la radiothérapie (Quesnel et al., 2009). Ces observations ont été confirmées par une étude ultérieure, qui a intégré une technique d'imagerie cérébrale en complément des évaluations neuropsychologiques (Kesler et al., 2017). La dynamique des réseaux cérébraux activés était significativement plus faible chez les patientes notamment lors des tâches de mémoire, de fonctions exécutives et de fluence verbale (Kesler et al., 2017). Selon les auteurs, ces résultats mettraient en lumière l'impact des cancers, hors tumeurs cérébrales, sur le fonctionnement du cerveau *via* par exemple des mécanismes d'inflammation. En effet, le cancer pourrait provoquer une réponse inflammatoire qui activerait des cytokines neurotoxiques, dont les interleukines et le facteur de nécrose tumorale (*Tumor Necrosis Factor*, TNF) sont souvent associés à un CRCI (Lyon et al., 2016). Les cytokines sont des protéines qui participent au fonctionnement du système immunitaire. Parmi elles, les interleukines modulent la croissance, la différenciation et l'activation cellulaires au cours des réponses inflammatoires et immunitaires, tandis que le TNF régule l'inflammation et envoie un signal aux cellules immunitaires pour détruire les cellules tumorales. Dans le cas de tumeurs malignes, ces cytokines affaiblissent et franchissent la barrière hémato-encéphalique, c'est-à-dire le tissu cellulaire qui sépare le sang du liquide encéphalique et protège le cerveau contre les agents pathogènes et toxines présents dans le sang. Les cytokines stimuleraient ainsi l'inflammation cérébrale en causant une dérégulation de l'activité des neurotransmetteurs impliqués dans le bon fonctionnement cognitif (Mayo et al., 2021). De plus, une concentration élevée de cytokines pro-inflammatoires dégraderait les performances cognitives (Henneghan et al., 2020) telles que les capacités de mémoire à court-terme (A. M. Williams et al., 2018).

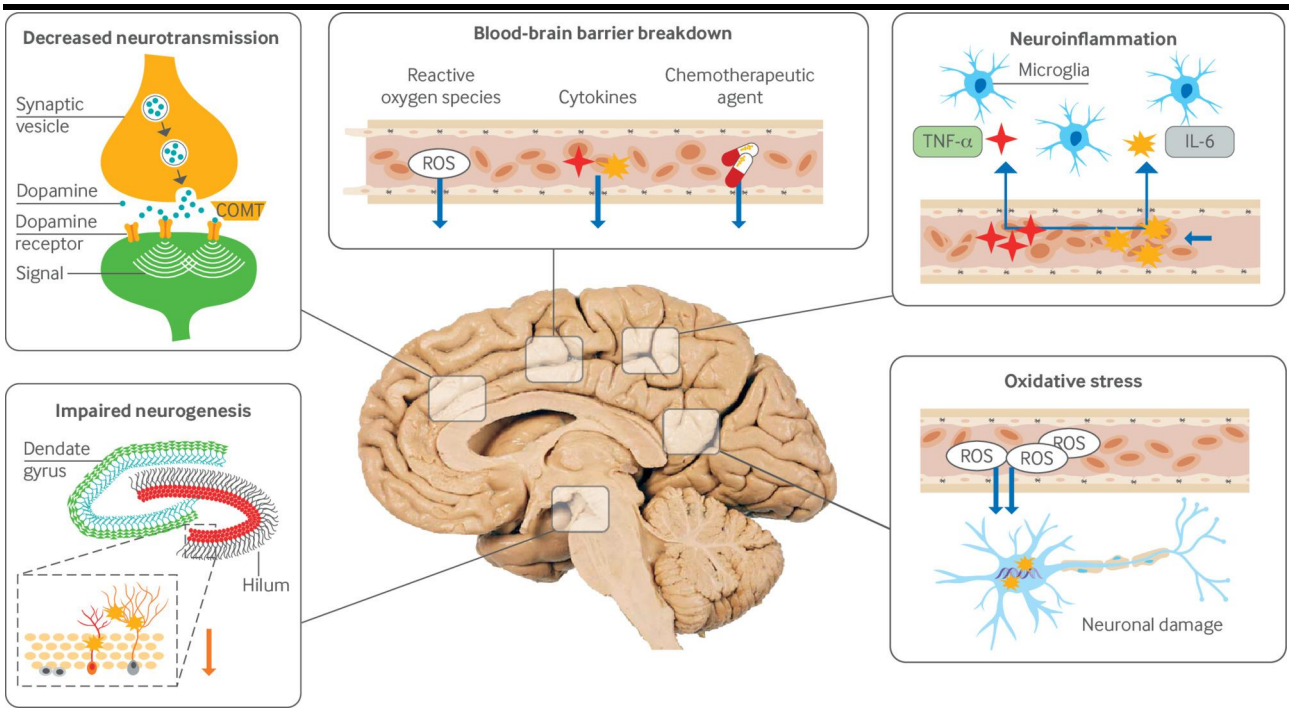
Traitements oncologiques

Les difficultés cognitives se manifestent de manière particulièrement prononcée durant la période des traitements. En effet, certains agents médicamenteux traversent la barrière hémato-encéphalique (Deeken & Löscher, 2007), et leur toxicité pourrait induire des dysfonctionnements cognitifs (Hurria & Lachs, 2007 ; Winocur et al., 2006 ; Winocur et al., 2018). Bien que la chimiothérapie soit fréquemment identifiée comme la principale cause du CRCI, il est toutefois possible que l'ensemble des traitements anticancéreux influence non seulement l'apparition, mais aussi la sévérité du CRCI (Demos-Davies et al., 2024 ; Fleming et al., 2023).

La chimiothérapie génèrerait un stress oxydatif (connu pour fragiliser la barrière hémato-encéphalique) et pourrait accroître le niveau de cytokines pro-inflammatoires. En effet, au cours des cycles de chimiothérapie, une élévation des taux de cytokines a été associée à des déficits dans divers domaines cognitifs, tels que la vitesse psychomotrice, l'attention complexe, les fonctions exécutives, la mémoire verbale et visuelle, ainsi que la flexibilité verbale, bien que les domaines touchés varient selon les études (Jim et al., 2012 ; Li & Caeyenberghs, 2018). Par ailleurs, les effets neurotoxiques des molécules chimiothérapeutiques sur le système nerveux périphérique sont largement reconnus (Verstappen et al., 2003). Ces substances pourraient également avoir des effets délétères sur le cerveau en lui-même. À titre d'exemple, des traces de 5-fluorouracile ont été détectées en faibles quantités dans le cerveau de modèles primates non-humains (Bourke et al., 1973). La présence de ces toxines cérébrales déclencherait une cascade de processus perturbant la réparation de l'ADN endommagé, entraînant ainsi la mort des neurones (Ren et al., 2019). En outre, l'intensité des doses de chimiothérapie semble exacerber la gravité du CRCI (Van Dam et al., 1998).

La figure 3, tirée de la revue de Fleming et al. (2023) sur le CRCI, schématise les différents mécanismes de la chimiothérapie favorisant l'apparition d'un trouble cognitif.

Figure 3. Schéma des potentiels mécanismes du CRCI induits par la chimiothérapie.



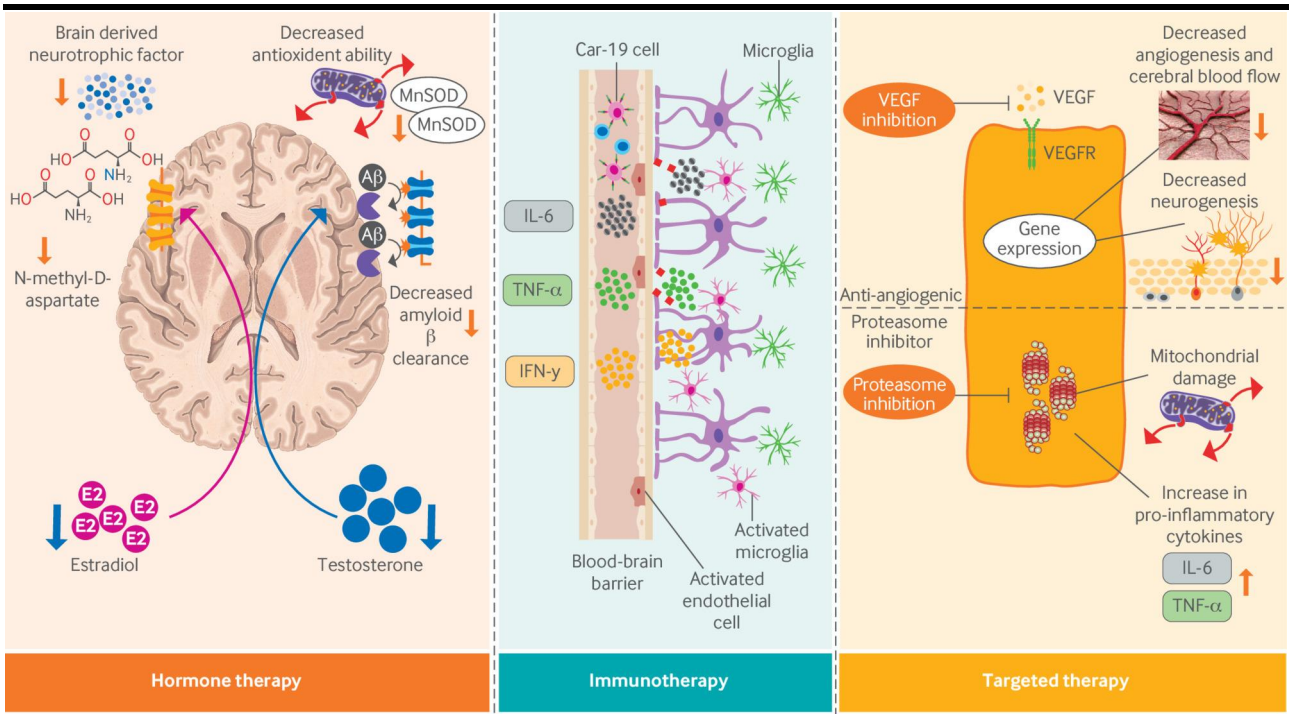
Note : Reproduit, avec permission, de Fleming et al. (2023). 2023 BMJ Publishing Group Ltd.

L'hormonothérapie a également été identifiée comme un facteur potentiel dans le développement du CRCI. L'impact de ce traitement sur la cognition a été mis en évidence à travers des questionnaires administrés à des patients atteints de cancer de la prostate (Boué et al., 2024). Une autre étude souligne les effets du tamoxifène sur la diminution des performances en fluence verbale et en mémoire verbale chez des patientes ayant un cancer du sein (Boele et al., 2015). Toutefois, ces résultats ne

sont pas confirmés par Van Dyk et al. (2019). Cette étude ne rapporte pas d'effet significatif de l'hormonothérapie sur les performances cognitives, tant évaluées par des questionnaires que par des tests neuropsychologiques, chez des patientes atteintes d'un cancer du sein. Cette divergence de résultats pourrait néanmoins être liée à la méthodologie employée pour l'évaluation des capacités cognitives (Van Dyk et al., 2019).

La figure 4, issue de Fleming et al. (2023), illustre les mécanismes neurologiques potentiellement impliqués dans le développement du CRCI selon les traitements.

Figure 4. Schéma des potentiels mécanismes impliqués dans le développement du CRCI selon les traitements.



Note : Reproduit, avec permission, de Fleming et al. (2023). 2023 BMJ Publishing Group Ltd.

Facteurs génétiques

Le rôle de trois prédispositions génétiques a désormais été clairement établi dans le développement et la sévérité du CRCI. La présence de l'allèle $\varepsilon 4$ de l'Apolipoprotéine E (APOE4), un gène également impliqué dans la Maladie d'Alzheimer, serait associée à une diminution des performances cognitives chez des patients traités par chimiothérapie notamment dans les domaines de la mémoire visuelle et des capacités visuo-spatiales (Ahles et al., 2003). D'autre part, le polymorphisme nucléotidique (c.-à-d. une variation génétique) de l'enzyme catéchol-O-méthyltransférase (COMT), laquelle intervient dans le fonctionnement du cortex préfrontal, entraînerait une altération des performances cognitives chez des patientes traitées par chimiothérapie pour un cancer du sein (Small et al., 2011). Enfin, le facteur neurotrophique dérivé du cerveau (*Brain-Derived Neurotrophic Factor*, BDNF) est une protéine active notamment dans l'hippocampe et jouant un rôle clé dans la mémoire à long terme. Il est codé par le gène BDNF, dont le polymorphisme semble réduire les risques de développement

du CRCI. Une étude révèle que les patientes atteintes d'un cancer du sein et présentant un polymorphisme du gène BDNF rapportent moins de déficits dans les capacités verbales et la gestion de la double tâche après chimiothérapie, par rapport aux autres patientes (Ng et al., 2016).

Facteurs psychologiques, physiologiques et socio-démographiques

Outre les facteurs biologiques, les facteurs psychologiques, physiologiques et socio-démographiques s'entrecroisent et influencent l'expression du CRCI. Bien que ces facteurs confondants ne permettent pas d'expliquer de manière directe l'apparition du CRCI, leur évaluation offre un éclairage précieux sur la trajectoire cognitive des patients.

Pour certains, l'annonce du diagnostic de cancer constitue un événement traumatique susceptible de générer un stress psychologique intense, voire un trouble de stress post-traumatique dans les cas les plus graves. Le stress pourrait ainsi être à la fois un facteur déclencheur et un facteur aggravant du CRCI (Hermelink et al., 2015 ; Hermelink et al., 2017). Cependant, la relation entre stress et développement du CRCI demeure encore incertaine. Hermelink et al. (2017) ont étudié l'impact du stress post-traumatique sur les capacités cognitives chez des femmes atteintes d'un cancer du sein. L'étude a comparé un groupe de patientes avant tout traitement, un groupe de patientes ayant reçu une chimiothérapie, ainsi qu'un groupe de témoins saines, qui ont toutes passées une évaluation clinique du stress post-traumatique et une batterie de tests neuropsychologiques. Les résultats ont montré qu'environ la moitié des participantes, toutes catégories confondues, ne présentaient pas de signes de stress post-traumatique. De plus, le chevauchement des scores entre les trois groupes, avec une tendance pour le groupe traité par chimiothérapie à afficher des performances cognitives moins bonnes, suggère que le stress n'a pas d'impact direct sur les performances cognitives. Les auteurs concluent que, bien que les femmes atteintes de cancer du sein, indépendamment des traitements, rencontrent des difficultés cognitives subtiles, ces dernières pourraient être en partie modulées par le stress post-traumatique. Pour autant, ce dernier ne serait pas explicatif de l'apparition du CRCI. Ces résultats sont corroborés par une étude plus récente qui indique que les patientes n'ayant pas encore reçu de traitement rapportent davantage de difficultés cognitives que les contrôles, sans toutefois présenter de différences notables en termes de performance cognitive globale (Aspelund et al., 2024).

En outre, la prévalence du stress post-traumatique chez les patientes atteintes de cancer reste relativement faible, avoisinant les 10% (Arnaboldi et al., 2017), et les symptômes de stress diminuent avec le temps. À ce jour, peu de preuves soutiennent l'existence d'un syndrome de stress post-traumatique à long terme chez les survivants d'un cancer. Une autre étude révèle qu'à trente jours de l'annonce du diagnostic, environ 19% des patientes présentent un syndrome de stress post-traumatique, tandis que 70% manifestent des symptômes anxieux. Cependant, l'intensité de ces symptômes diminue de manière significative au cours des deux années suivant le diagnostic (Arnaboldi et al., 2014). De la même manière, l'évaluation psychologique des patients en oncologie met souvent en lumière la présence de symptômes dépressifs et anxieux. Une revue systématique

a ainsi révélé que de 9,4% à 66,1% des survivantes d'un cancer souffrent de dépression, tandis que de 17,9% à 33,3% présentent des symptômes anxieux. Cependant, ces chiffres ne sont pas significativement plus élevés que ceux observés dans la population féminine générale (Maass et al., 2015).

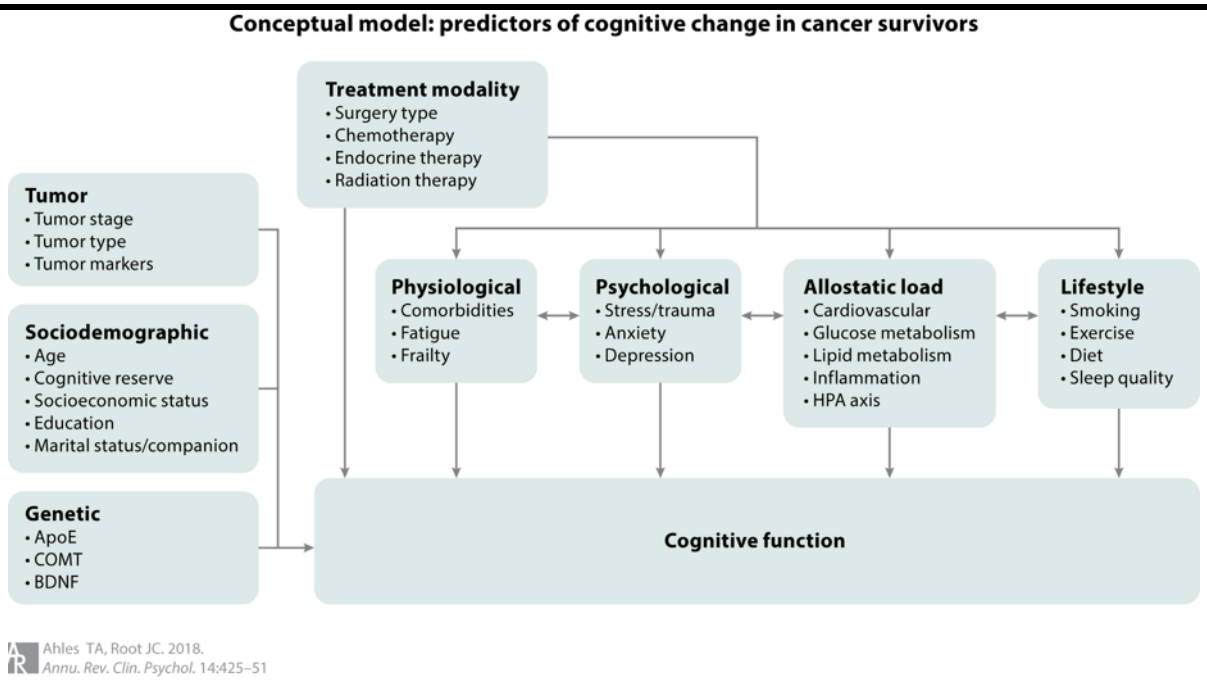
Un autre facteur souvent associé au cancer et ses traitements concerne la fatigue induite par le cancer et ses traitements, ou (*Cancer-Related Fatigue*). Ce symptôme, qui persiste souvent bien après l'arrêt des traitements, se caractérise par une sensation de fatigue généralisée non soulagée par le sommeil (Hofman et al., 2007 ; Kang et al., 2023 ; Rosas et al., 2023 ; Ruiz-Casado et al., 2021). Il n'est donc guère surprenant que cette fatigue puisse entraîner une baisse des capacités attentionnelles (Kohler et al., 2020). Dès lors, certains patients pourraient associer leurs plaintes cognitives à la fatigue, plutôt qu'à un véritable déficit cognitif.

La réserve cérébrale et cognitive, c'est-à-dire la capacité du cerveau à compenser les lésions cérébrales en établissant de nouvelles connexions neuronales ou en réactivant des réseaux existants, ainsi que l'âge biologique (la dégradation du corps au niveau cellulaire), jouent un rôle modérateur dans la sévérité du CRCI (Ahles et al., 2010 ; Carroll et al., 2019). En d'autres termes, une faible réserve cognitive ou un âge avancé augmenteraient le risque de développer un trouble persistant (Kesler, 2011 ; Lange et al., 2016). Cependant, il est important de souligner que même les patients jeunes (moins de quarante-cinq ans) peuvent être affectés par des difficultés cognitives à long terme (Von Ah et al., 2022).

Enfin, le mode de vie influence également les capacités cognitives des survivants. L'obésité, l'activité physique et la qualité du sommeil influenceraient notamment les scores des tests neuropsychologiques, bien que ceux-ci demeurent supérieurs aux seuils pathologiques (Hartman et al., 2015).

Pour résumer, le CRCI est un trouble polyfactoriel qui résulte probablement d'une interaction entre la tumeur elle-même et les traitements administrés, ainsi que les autres facteurs confondants. Ahles et Root (2018) proposent un modèle global des facteurs prédicteurs du CRCI chez les survivants d'un cancer (figure 5).

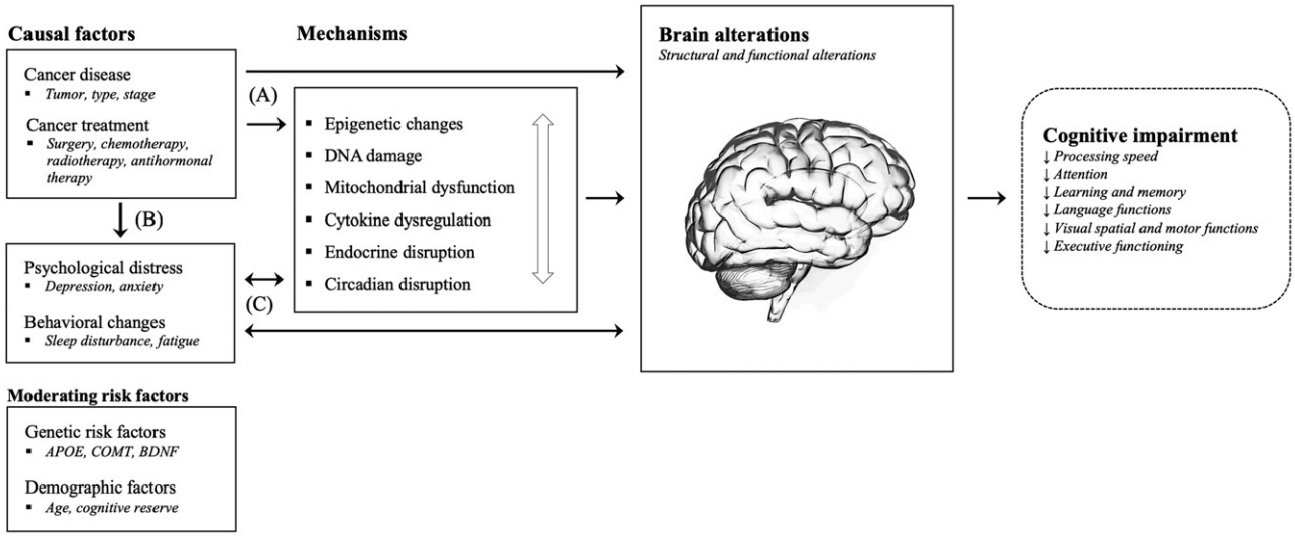
Figure 5. Schéma des facteurs prédictifs d'un changement cognitif chez des survivants d'un cancer.



Note : Reproduit, avec permission, de Ahles et Root (2018). 2018 Annual Reviews.

Amidi et Wu (2019), quant à eux, offrent un aperçu des mécanismes potentiellement impliqués dans le développement d'un CRCI en ajoutant une étape qui illustre l'impact de ces facteurs sur les modifications cérébrales (figure 6).

Figure 6. Schéma des mécanismes ayant un rôle dans le développement d'un CRCI.



Note : Reproduit, avec permission, de Amidi et Wu (2019). 2019 Acta Oncologica.

2.2 Modifications cérébrales structurelles et fonctionnelles associées au CRCI

De nombreuses études utilisant l'imagerie cérébrale ont révélé des modifications structurelles et fonctionnelles du cerveau chez les patients en oncologie. Ces recherches sont principalement conduites chez les femmes atteintes d'un cancer du sein et mettent en évidence les effets de la chimiothérapie sur le fonctionnement cérébral ou sur les structures du cerveau (Feng et al., 2019).

Effets de la chimiothérapie

Les modifications structurelles et fonctionnelles du cerveau engendrées par la chimiothérapie sont de mieux en mieux identifiées (Whittaker et al., 2022). Une dégradation de l'intégrité de la matière blanche est par exemple observée dans le faisceau longitudinal supérieur (associé à l'attention visuo-spatiale et aux comportements moteurs complexes) et dans le faisceau cortico-spinal (responsable des commandes motrices) (Menning et al., 2018). Deprez et al. (2012) ont analysé l'intégrité de la matière blanche avant et après chimiothérapie *via* une technique d'imagerie par résonance magnétique (IRM) de diffusion associée à une évaluation cognitive exhaustive. Leurs résultats ont montré une réduction de l'activation des faisceaux frontaux, pariétaux et occipitaux après chimiothérapie, en lien avec une altération des capacités attentionnelles et de la mémoire verbale (Deprez et al., 2012).

D'autres recherches ont noté une diminution de la matière grise dans les gyri préfrontal et parahippocampique ainsi que dans le précuneus chez les survivantes d'un cancer du sein ayant reçu une chimiothérapie (Inagaki et al., 2007). Cependant, cet effet n'a pas été retrouvé chez des patientes évaluées plusieurs années après la chimiothérapie, suggérant ainsi une possible récupération à long terme. Ces résultats restent toutefois à confirmer par d'autres études. En effet, les altérations de l'activité cérébrale perdurent parfois des années après l'arrêt de la chimiothérapie. Par exemple, des déficits ont été identifiés dans les régions fronto-corticales, les ganglions de la base et le cervelet cinq à dix ans après la fin de ce traitement (Silverman et al., 2007). De même, une étude a comparé les performances cognitives de patientes ayant reçu une chimiothérapie à haute dose dix ans après le traitement à celles de patientes non traitées (De Ruiter et al., 2011). Lors de cette étude, les participantes ont réalisé des tâches impliquant des fonctions exécutives et des processus attentionnels pendant que leur activité cérébrale était enregistrée par une IRMf. Les participantes ayant reçu de la chimiothérapie présentaient des troubles cognitifs à long terme caractérisés par une hyporéactivité du cortex préfrontal dorsolatéral, du gyrus parahippocampique et du cortex pariétal latéral postérieur (De Ruiter et al., 2011). Enfin, Deprez et al. (2014) ont mis en évidence des différences fonctionnelles chez des patientes ayant un cancer du sein traitées par chimiothérapie au moyen de tâches impliquant les fonctions exécutives et les ressources attentionnelles (Deprez et al., 2014). Ces tâches ont activé différentes régions cérébrales, telles que le réseau fronto-pariétal bilatéral, le cervelet, le cingulaire antérieur et le cortex occipito-temporal, dans le groupe de patientes ayant reçu de la chimiothérapie ainsi que dans les groupes contrôles. Toutefois, seule les patientes traitées par chimiothérapie ont montré une baisse générale de l'activité cérébrale. Cette réduction était corrélée à des plaintes

cognitives plus marquées chez ces patientes (Deprez et al., 2014).

Impact sur les régions fronto-pariétales et le contrôle exécutif

Les recherches ont révélé la présence de lésions diffuses et fines dans diverses régions cérébrales. Pour autant il est difficile de conclure à une localisation systématique de ces lésions (Amidi & Wu, 2019). Les études utilisant l'imagerie cérébrale mettent en lumière des altérations principalement situées dans les régions pariétales et frontales. Ces régions sont impliquées dans le fonctionnement attentionnel et exécutif (Cristofori et al., 2019 ; Knudsen, 2007). Une activation significativement réduite du cortex préfrontal dorsolatéral gauche et du cortex prémoteur est observée chez des patientes atteintes d'un cancer du sein, comparativement aux témoins sains. Ce groupe, traité par chimiothérapie, a montré des déficits en matière d'inhibition et une baisse de la vitesse de traitement (Kesler, 2011). Comparé à des patientes non traitées par chimiothérapie et à des témoins sains, les patientes ayant reçu de la chimiothérapie ont montré une réduction de l'hyperactivation frontale bilatérale lors de tâche mobilisant les fonctions exécutives (B. C. McDonald et al., 2012). Dans une autre étude, l'activité cérébrale de patientes atteintes d'un cancer du sein métastatique a été enregistrée par IRMf lors de tâches de mémoire déclarative verbale (Kesler et al., 2009). Les résultats de cette étude révèlent une activation plus faible du cortex préfrontal pendant l'encodage de la mémoire, par rapport aux témoins (Kesler et al., 2009). Quelques études ont également mis en avant des altérations des lobes temporaux partiellement recouvrées à distance des traitements (p. ex., Conroy et al. [2013] et Lepage et al. [2014]). Toutefois, les résultats de ces études sont disparates et ne suffisent pas à démontrer l'existence d'un foyer clairement défini (Amidi & Wu, 2019).

Amidi et Wu (2019) ont conduit une revue systématique portant sur les altérations cérébrales après un traitement oncologique. Cette revue montrent que des lésions diffuses et une réduction du volume de matière grise chez les patients en oncologie ont été identifiées par différentes techniques d'imagerie cérébrales (Amidi & Wu, 2019). La majorité des études analysées révélaient une diminution des volumes de matière grise, une altération de la matière blanche, ainsi qu'une connectivité amoindrie entre les régions cérébrales (Amidi & Wu, 2019). Bernstein et al. (2021), quant à eux, ont mené une méta-analyse sur les modifications structurelles du cerveau après chimiothérapie et leurs effets sur les fonctions exécutives et les processus attentionnels (Bernstein et al., 2021). Cette étude souligne que, malgré l'hétérogénéité des méthodes utilisées, deux régions cérébrales sont souvent associées au CRCI : le cortex pariétal supérieur gauche et les régions préfrontales supérieures droites. En particulier, l'activation du réseau fronto-pariétal, impliqué dans les processus attentionnels, serait plus faible après chimiothérapie. Par conséquent, la chimiothérapie aurait un impact sur les régions responsables de la régulation de l'attention et des fonctions exécutives. Les résultats de cette méta-analyse sont synthétisés dans la figure 7 (Bernstein et al., 2021).

Figure 7. Synthèse des études montrant les régions cérébrales associées au CRCI. Les points bleus indiquent une hypoactivation cérébrale. Les points rouges indiquent une hyperactivation cérébrale.

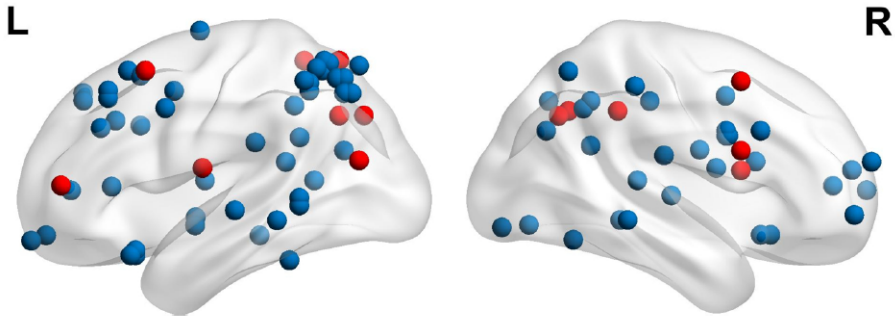


Fig. 2. Foci from studies showing greater activation (Red) and from studies reporting lower activation (Blue) associated with chemotherapy. L = Left; R = Right.

Note : Reproduit, avec permission, de Bernstein et al. (2021). 2021 Elsevier Ltd.

Un point intéressant qui émerge de ces recherches est que les différences cérébrales sont généralement plus marquées lorsque les groupes oncologiques sont comparés à des témoins sains, mais plus subtiles lorsque ces groupes sont comparés à des sujets également atteints de cancer (Amidi & Wu, 2019). En d'autres termes, les altérations sont davantage mises en évidence entre des patients et des sujets contrôles sains qu'entre des patients bénéficiant une chimiothérapie et des patients n'en recevant pas. Ce phénomène suggère que le cancer lui-même pourrait déjà induire des changements cérébraux avant tout traitement, et que la chimiothérapie aggraverait ces troubles.

2.3 Description des difficultés cognitives

Le CRCI peut affecter les patients atteints de n'importe quel type de cancer, indépendamment du stade de la tumeur, des traitements reçus et de la temporalité du cancer. En dépit de la diversité des sous-populations, les difficultés décrites par les patients et mesurées par les tests sont étonnamment similaires (Demos-Davies et al., 2024). Les patients en oncologie font état de symptômes cognitifs affectant plusieurs domaines, tels que le langage, la mémoire, la concentration, l'attention, la résolution de problèmes et l'établissement de priorités (Hutchinson et al., 2012).

Les difficultés linguistiques décrites par les patients sont nombreuses. Elles concernent l'oubli de mots, lesquels peuvent être substitués par un autre (Bolton & Isaacs, 2018 ; Player et al., 2014 ; Skoogh et al., 2012). Les patients éprouvent des difficultés à terminer leurs phrases et il n'est pas rare qu'ils inversent l'ordre des mots (Skoogh et al., 2012 ; Von Ah et al., 2013). En situation conversationnelle, ils peinent à suivre les tours de parole (c.-à-d. les prises de parole successives des locuteurs au sein d'une conversation, [Sacks et al., 1974]) et à interagir avec un groupe (Bolton & Isaacs, 2018 ; Skoogh et al., 2012). Plus rarement, les patients se plaignent d'oublier l'orthographe de certains mots (Bolton & Isaacs, 2018).

Le manque du mot, ou anomie (voir section Anomie : Comprendre la cognition par l'erreur), est

l'un des symptômes du CRCI les plus régulièrement décrits par les patients. La prévalence de ce phénomène s'étend d'un tiers (Von Ah et al., 2013) à l'ensemble des patients (Myers, 2012). Cette variabilité pourrait être expliquée par les différents moments du parcours de soin où les patients sont interrogés à propos de leurs difficultés cognitives. Par exemple, Downie et al. (2006) ont mené des entretiens avec des patientes atteintes d'un cancer du sein pendant la phase curative du traitement et ont trouvé que 78% des vingt-et-unes patientes ont rapporté des difficultés à nommer des mots (Downie et al., 2006). Myers (2012), à l'occasion d'entretiens avec des patientes six à douze mois après la phase curative, a observé que toutes les patientes mentionnaient des difficultés à retrouver les mots. Cependant, le pourcentage de plaintes liées aux difficultés linguistiques semble diminuer avec le temps écoulé après la fin des traitements (Myers, 2012). Dans une étude de Von Ah et al. (2013), menée auprès de femmes un à douze ans après la phase curative, 36% d'entre elles continuaient à signaler des problèmes de dénomination lexicale (Von Ah et al., 2013). Ce phénomène est loin d'être l'apanage de femmes ayant un cancer du sein. Chez les survivants d'un cancer du testicule en moyenne onze ans après le diagnostic, 16% d'entre eux déclarent « dire des mots similaires mais incorrects » au moins une fois par semaine, et entre 17% et 36% rapportent des difficultés pour « trouver des mots » ou « rencontrer d'autres problèmes de langage » (Skoogh et al., 2012, p. 191).

Les dysfonctionnements mnésiques représentent également une source d'inquiétude importante chez les patients en oncologie, qui éprouvent des difficultés à se rappeler des noms (Bolton & Isaacs, 2018 ; Von Ah et al., 2013), des visages familiers (Bolton & Isaacs, 2018) et des dates importantes (Von Ah et al., 2013). L'attention n'est pas non plus épargnée. Les patients rapportent des problèmes de concentration gênants, comme être incapables de se focaliser sur une tâche, être facilement distraits, regarder un film sans se souvenir de l'intrigue ni des noms des personnages (Bolton & Isaacs, 2018), ou devoir lire plusieurs fois la même page d'un livre sans en saisir le sens (Player et al., 2014 ; Von Ah et al., 2013). Certains patients déclarent ne pas se souvenir de ce qu'on leur a dit (Tenda et al., 2022), se répéter, ou encore oublier l'objectif d'une tâche qu'ils sont en train de réaliser (Von Ah et al., 2013). En ce qui concerne les fonctions exécutives, les patients indiquent des difficultés à se souvenir de tâches routinières, comme nourrir des animaux ou suivre un itinéraire familier (Bolton & Isaacs, 2018 ; Player et al., 2014 ; Von Ah et al., 2013), peiner à alterner entre deux tâches simultanées (Bolton & Isaacs, 2018 ; Von Ah et al., 2013), et être incapables de traiter de nouvelles informations ou d'acquérir de nouvelles compétences.

2.4 Évaluation subjective du CRCI

Le recueil de la plainte cognitive peut s'effectuer à l'aide d'évaluations subjectives, notamment par des questionnaires autoadministrés. L'évaluation des difficultés cognitives au moyen de questionnaires standardisés présente l'avantage d'être proche du quotidien des patients. Parmi les questionnaires existants, quatre ont été conçus pour une patientèle en oncologie (FACT-Cog, EORTC-QLQ-C30, PAOFI, AFI) et sont le plus souvent utilisés dans les études cliniques.

FACT-Cog

Le questionnaire *Functional Assessment of Cancer Therapy – Cognitive Function* (FACT-Cog v3 [Costa et al., 2018; Wagner et al., 2009]) a été spécifiquement développé pour évaluer la présence de troubles cognitifs chez une population oncologique. Il est considéré comme l'outil de référence pour détecter la présence de difficultés cognitives chez les patients ayant un cancer (Henneghan et al., 2021). Originellement écrit en anglais, il a fait l'objet d'une traduction française (Joly et al., 2012) validée sur une population normale (Lange et al., 2016). Sa version validée actuelle se compose de quatre sous-échelles (Déficiences Cognitives Perçues, *Perceived Cognitive Impairments*; Commentaires d'Autres Personnes, *Comments from Others*; Aptitudes Cognitives Perçues, *Perceived Cognitive Abilities*; Impact sur la qualité de Vie, *Impact on Quality of Life*) et évalue six dimensions cognitives (mémoire, acuité verbale, concentration/attention, acuité mentale, interférence fonctionnelle et multitâche) à l'aide de trente-sept questions. Quatre de ces items sont consacrés au langage :

- « J'ai eu du mal à me rappeler du nom d'un objet alors que j'étais en train de parler à quelqu'un »,
- « J'ai eu du mal à trouver le(s) mot(s) juste(s) pour m'exprimer »,
- « J'ai utilisé un mauvais mot pour désigner un objet »,
- « J'ai eu du mal à exprimer ce que je voulais dire dans mes conversations avec d'autres personnes ».

Costa et al. (2018) ont examiné les réponses de trois groupes de participants (des jeunes adultes sains, des adultes âgés sains et des patients en oncologie) en fonction des domaines cognitifs présents dans le questionnaire. Les scores des patients montraient des déficits pour l'ensemble des domaines cognitifs, tandis que ceux des jeunes adultes et ceux des adultes plus âgés étaient plus faibles dans certains domaines uniquement. Ce résultat suggère que l'ensemble des items du FACT-Cog doit être analysé comme un tout. Autrement dit, la structure du questionnaire ne permet pas d'identifier des domaines cognitifs les plus affectés. Par conséquent, bien que quatre des items du FACT-Cog soient dédiés aux capacités verbales, leur score isolé n'est pas capable de signaler d'éventuelles difficultés linguistiques. De plus, ces items ne permettent pas de comprendre précisément la nature de la plainte de langage.

EORTC QLQ-C30

Le questionnaire QLQ-C30 développé par la *European Organisation for Research and Treatment of Cancer* (EORTC) est un questionnaire autoadministré dédié à l'évaluation de la qualité de vie des patients en oncologie (Aaronson et al., 1993). Il possède cinq échelles fonctionnelles (physique, rôle,

cognitif, émotion et social) et trois échelles de symptômes (fatigue, douleur et nausées). La modularité de ce questionnaire permet d'obtenir une vision globale de la qualité de vie des patients. Cependant que seuls deux items sur trente seulement font explicitement référence à la cognition :

- « Avez-vous eu des difficultés à vous concentrer sur certaines choses, par exemple, pour lire le journal ou regarder la télévision ? »
- « Avez-vous eu des difficultés à vous souvenir de certaines choses ? »

Pour cette raison, il permet difficilement de rendre compte des troubles cognitifs vécus par les patients.

Patient's Assessment of Own Functioning Inventory

Le *Patient's Assessment of Own Functioning Inventory* (PAOFI) est un questionnaire conçu pour évaluer les difficultés cognitives perçues par les patients, toutes étiologies confondues (Chelune et al., 1986), qui a été adapté pour patients en oncologie (Van Dyk et al., 2016). Sa fiabilité et sa validité chez des femmes ayant un cancer du sein avant même la prise des traitements a été démontrée (Bell et al., 2013). Il contient des questions ouvertes regroupées par domaines cognitifs (mémoire, langage, perception sensorimotrice, fonctions cognitives de haut niveau) correspondant à la plainte des patients. Les items sont évalués par une échelle de Likert à 6 points. Six d'entre eux portent sur le langage, dont deux sur le manque du mot. Malheureusement, ce questionnaire ne dispose pas d'une traduction validée en français.

Attentional Function Index

L'*Attentional Function Index* (Cimprich et al., 2011) est un questionnaire autoadministré conçu pour estimer l'efficience perçue lors d'activités soutenues par l'attention et la mémoire de travail chez des patientes ayant survécu à un cancer du sein. Il présente l'avantage d'être rapide à compléter et d'interroger spécifiquement le retentissement des difficultés attentionnelles dans la vie quotidienne. En revanche, il ne prend pas en compte l'une des plaintes les plus fréquemment rapportées par ces patientes, à savoir l'anomie. De plus, ce questionnaire ne bénéficie pas non plus d'une traduction française.

2.5 Évaluations neuropsychologiques

De manière logique, les domaines cognitifs les plus testés sont ceux qui font écho à la plainte des patients. Il s'agit de la mémoire, des capacités verbales et de l'attention (Ahles et al., 2010 ; Kohler et al., 2020). Les tests couramment employés à l'évaluation du CRCI sont le *California Verbal*

Learning Test (Delis et al., 1987), le *Hopkins Verbal Learning Test - Revised* (Benedict et al., 1991), le *Rey Auditory Verbal Learning Test* (Rey, 1958), le *Rey–Osterrieth Complex Figure* (Osterrieth, 1944 ; Rey, 1941), l'échelle de mémoire de Wechsler (Wechsler, 1945), et le *Trail Making Test* (Armitage, 1946) (Cf. Jung et al. [2023] pour une revue).

La littérature sur l'évaluation du CRCI rapporte des résultats inconsistants concernant les domaines cognitifs touchés. Par exemple, une diminution des capacités verbales et visuo-spatiales peut être observée après un traitement par chimiothérapie (Jim et al., 2012), quand d'autres études montrent des résultats plus mitigés (Demos-Davies et al., 2024). Le décalage entre les articles vient principalement du fait des méthodes d'évaluation qui varie d'une étude à l'autre. Pour tenter d'harmoniser les résultats, le groupe de travail *International Cancer and Cognition Task Force* et le Groupe de Réflexion sur les Évaluations Cognitives en Oncologie ont proposé des batteries de tests pour aider au diagnostic du CRCI.

ICCTF

L'*International Cancer and Cognition Task Force* (ICCTF) est un groupe de travail multidisciplinaire réunissant des chercheurs, cliniciens et patients autour des questions relatives au CRCI. En 2011, ce groupe a formulé des recommandations pour l'évaluation du CRCI (Wefel et al., 2011), préconisant l'utilisation du *Hopkins Verbal Learning Test-Revised*, du *Trail Making Test* et du *Controlled Oral Word Association* issu du *Multilingual Aphasia Examination* (Benton et al., 1994). Ces tests mesurent différentes les principales capacités cognitives affectées par le CRCI, à savoir l'apprentissage, la mémoire, l'attention, la vitesse de traitement et les fonctions exécutives (Demos-Davies et al., 2024). Afin de compenser la faible sensibilité des tests à détecter le CRCI, le groupe a redéfini et adapté leurs seuils pathologiques à la population en oncologie. Ainsi, pour poser un diagnostic de CRCI, le patient doit obtenir un score inférieur à -1,5 écart-type à la moyenne normative à un minimum deux des tests mentionnés, ou bien un score inférieur à -2 écart-type à un seul de ces tests. Cette batterie présente l'avantage d'être facile et rapide à administrer. Toutefois, elle ne comprend pas de questionnaires, puisque l'ICCTF a estimé que les mesures des évaluations subjectives (ou *patient-reported outcomes*) seraient plutôt associées à des facteurs confondants, tels que la fatigue, qu'aux tests neuropsychologiques. En effet, les scores des questionnaires de qualité de vie et ceux des échelles de mesures psychologiques (p. ex., anxiété et dépression) se sont avérés fortement corrélés, mais n'ont montré aucune correspondance entre eux et les scores aux tests neuropsychologiques (Hermelink et al., 2010). Une revue systématique sur l'application des recommandations émises par l'ICCTF a montré que, parmi 93 études sur l'évaluation psychométrique sur CRCI, moins de la moitié suivaient ces lignes directrices (Ferguson et al., 2024). L'adoption des recommandations a néanmoins progressé entre 2011 et 2022 (56,2% en 2022), avant d'observer une diminution en 2023. Malgré l'établissement de seuils spécifiques, l'écart entre les évaluations subjectives et les résultats des tests persiste, et de ce fait, le CRCI demeure difficile à détecter.

GREC-Onco

En France, le Groupe de Réflexion sur les Évaluations Cognitives en Oncologie (GREC-ONCO) a proposé en 2015 une batterie de tests plus exhaustive comprenant un dépistage cognitif (fNART, Mattis, MoCA) et un examen plus approfondi de la mémoire (RL-RI 16, HVLT, empan de chiffres endroit et envers, copie de la figure de Rey à 3 minutes, empan visuels endroit et envers), du langage (BNT et Token Test), des fonctions visuo-constructives (copie de la figure de Rey), des fonctions exécutives et de l'attention (Stroop, TMT A/B, fluences verbales phonémiques et catégorielles, codes de la WAIS). À ces tests ont été ajoutés une évaluation comportementale, des questionnaires d'anxiété, de dépression, de fatigue et de qualité de vie (Taillia et al., 2015). Plutôt que de tenter d'établir des scores seuils, l'objectif de cette batterie est de détecter l'évolution cognitive chez le patient au fil du temps. Elle permet ainsi de prendre en compte la dynamique du CRCI, en proposant trois évaluations chronologiques : avant les traitements, à six mois après la batterie initiale, puis en suivi annuel. Toutefois, la multiplication des tests est chronophage et peut entraîner un biais de surdiagnostic (Rick et al., 2024). De plus, elle ne résout pas le problème du décalage entre les plaintes subjectives et les résultats des tests.

2.6 Discordance entre difficultés ressenties et tests neuropsychologiques

À ce jour, l'évaluation objective du CRCI reste un défi clinique. Il s'agit d'un trouble « subtil » en ce qu'il est difficilement détectable à travers les tests psychométriques existants, bien que son impact sur la vie des patients soit tangible. En effet, les scores des patients se situent souvent dans l'intervalle normal des scores aux évaluations neuropsychologiques (Stewart et al., 2008). Du défaut de tests adaptés résulte un décalage entre les difficultés décrites et leur évaluation objective. Cette discordance peut influencer de manière erronée la proposition de réponse thérapeutique et de prise en charge. Dans une méta-analyse, Whittaker et al. (2022) rapportent qu'un tiers des patients ayant reçu une chimiothérapie présentent une dégradation significative de la cognition, attestée par les outils neuropsychologiques. Cependant, ce nombre est nettement plus élevé lors des évaluations subjectives. En outre, l'inconsistance des résultats entre les études complique l'identification précise du CRCI. Par exemple, dans une étude de 2013, Ganz et al. (2013) ont mis en évidence la corrélation entre les difficultés mnésiques subjectives et la diminution des capacités verbales chez les patientes atteintes de cancer du sein. La même équipe ne constate pas de différence en termes de performances cognitives entre les évaluations effectuées avant et douze mois après avoir terminé les traitements de première ligne (chimiothérapie et radiothérapie) (Ganz et al., 2013). Pourtant, les patientes de cette cohorte exprimaient des difficultés de langage et la communication (Ganz et al., 2014). Plus récemment, Veal et al. (2023) ont examiné la relation entre les troubles rapportés par les patients et leur objectivation par des tests neuropsychologiques. Les scores du questionnaire FACT-Cog n'étaient pas corrélés à ceux des tests de vitesse de traitement, de traitement visuo-spatial et de mémoire verbale. En revanche, la description des trous de mémoire (*via* une liste de symptômes fournie) correspondait aux scores

obtenus aux tests neuropsychologiques.

Le décalage entre les réponses aux questionnaires et les scores obtenus pourrait s'expliquer par plusieurs facteurs. Tout d'abord, les tests neuropsychologiques ont généralement été développés et validés pour des populations présentant des troubles neurologiques plus sévères, chez des populations cérébrolésées. Bien que les seuils pathologiques aient été redéfinis, ces tests peinent à révéler les troubles cognitifs dont l'expression est plus subtile. Une autre explication réside dans le fait que les patients sont évalués dans des conditions cliniques optimales, c'est-à-dire en entretien avec un clinicien dans un environnement calme et contrôlé. Il est donc possible que les patients ayant un CRCI parviennent à compenser leurs difficultés cognitives, non sans effort, dans ces conditions idéales et sur une période relativement courte. De ce fait, ces tests ne réussissent pas à rendre pleinement compte de l'impact fonctionnel du CRCI sur la vie quotidienne des patients, où les stimulations sont multimodales et moins contrôlées. De plus, ils n'intègrent pas la présence de facteurs confondants, tels que la fatigue ou l'anxiété, qui peuvent induire une variabilité au sein des résultats et entraîner des incohérences dans le suivi des patients. Enfin, ces examens psychométriques sont conçus pour évaluer une fonction cognitive isolée de manière à tester son efficience indépendamment des autres composantes. Toutefois, cette approche ne correspond pas aux difficultés rencontrées par les patients en oncologie. Par conséquent, il est probable que les tests neuropsychologiques et les questionnaires ne mesurent pas les mêmes dimensions (Costa & Fardell, 2019). Cette hypothèse est étayée par des études qui ont trouvé une association entre les difficultés cognitives perçues et des changements neuroanatomiques, sans que ces derniers ne soient corrélés aux résultats des tests neuropsychologiques. Les questionnaires seraient ainsi plus sensibles aux changements cognitifs induits par les expériences de tous les jours.

À cet égard, la recherche sur le CRCI devrait à présent concevoir des outils de diagnostic plus efficaces. En plus de critères de sensibilité et de spécificité, ces outils devront également intégrer une dimension écologique, qui permettra d'évaluer avec précision les capacités cognitives des patients dans leur environnement quotidien. Une approche holistique prenant en compte l'existence de co-facteurs s'avère indispensable.

MESSAGES CLÉS — TROUBLE COGNITIF INDUIT PAR UN CANCER ET SES TRAITEMENTS

1. ***Cancer-Related Cognitive Impairment***

    - Les cancers et leurs traitements peuvent induire des changements cognitifs. Ces difficultés sont subtiles, mais néanmoins invalidantes. Elles dégradent la qualité de vie des patients ayant ou ayant eu un cancer.

    - Le cancer en lui-même provoquerait des inflammations cérébrales responsables de ces difficultés cognitives. Les traitements, tels que la chimiothérapie, ainsi que d'autres facteurs confondants (p. ex., fatigue, dépression, etc.) aggravent l'expression de ces difficultés.

    - La temporalité de ce syndrome cognitif est étendue : il peut apparaître avant le début des traitements, est généralement très élevé pendant et immédiatement à la fin des traitements, et peut perdurer au-delà de l'arrêt des traitements. Il est essentiel de repérer les patients à risque de développer des symptômes durables.

2. **Discordance entre les difficultés décrites et les performances objectivées**

    - Les patients en oncologie décrivent des difficultés concernant leur mémoire défaillante, leur fenêtre attentionnelle réduite, leur incapacité à jongler entre les tâches et une gêne persistante à trouver leurs mots.

    - Cet ensemble de difficultés est aisément mis en évidence par les questionnaires. En revanche, les tests neuropsychologiques actuels peinent à révéler l'amplitude des déficits. Ce décalage entre évaluations subjectives et objectives empêche certains patients d'avoir accès à des soins adaptés. De plus, le manque d'outil psychométrique fiable ralentit l'état des connaissances sur le fonctionnement du CRCI.

    - Le développement de nouvelles approches est nécessaire afin d'identifier les patients en oncologie souffrant d'un trouble cognitif induit par un cancer. En outre, ces méthodes apporteront un éclairage quant aux mécanismes sous-jacents au CRCI.

3 La parole : une ouverture sur la cognition

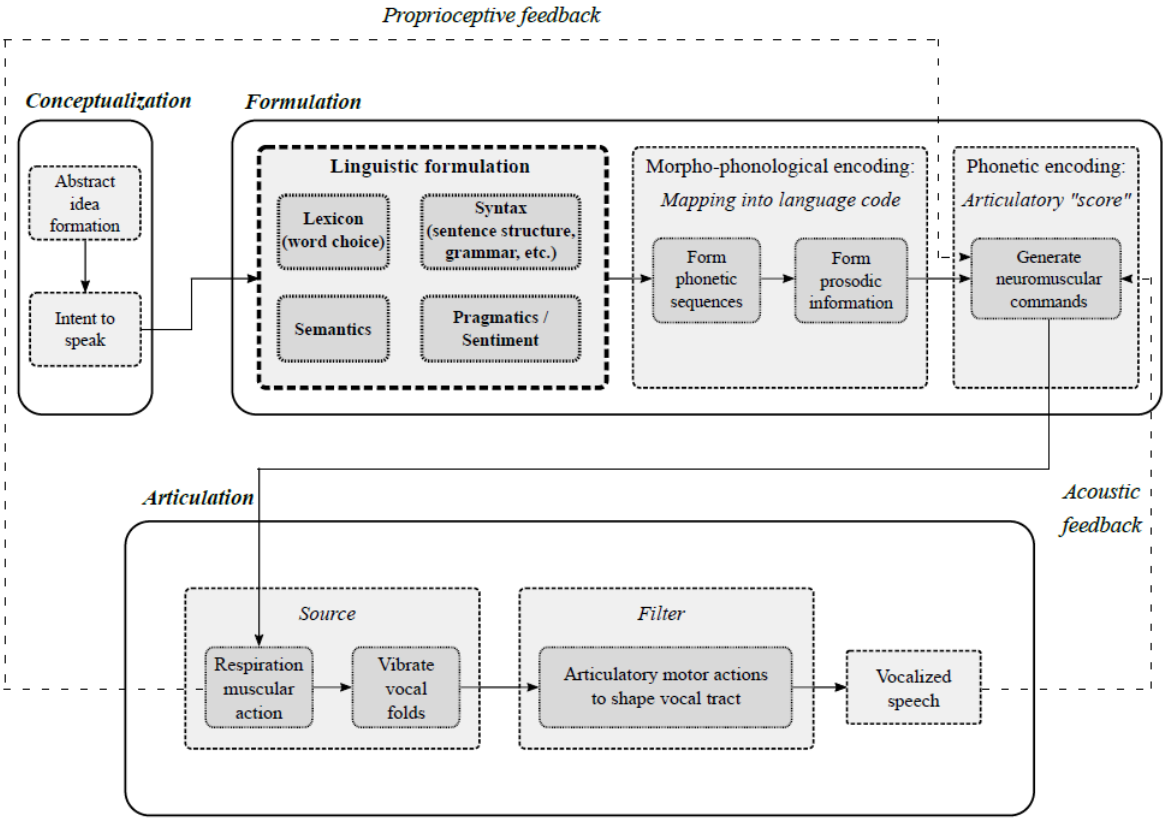
La parole, production audible du langage, est une faculté aussi riche que complexe employée lors d'interactions quotidiennes. Elle se traduit sous forme de signal acoustique interprétable grâce à des unités segmentales, c.-à-d. les phonèmes d'un système linguistique, et grâce à des éléments supra-segmentaux, c.-à-d. des caractéristiques prosodiques qui unissent un ensemble de phonèmes. La prosodie fait référence à trois types de paramètres perceptibles de la parole : la hauteur intonative, le volume et le rythme, relatifs à la fréquence en Hz, à l'intensité en dB, et le débit généralement mesuré en syllabes par seconde (De Jong et Wempe [2009], mais voir aussi Tilsen et Tiede [2023] pour des mesures plus fines). Elle véhicule des informations sur la structure du message oral, ainsi que sur les émotions et intentions réelles ou simulées d'un locuteur. La prosodie se trouve ainsi à l'interface d'autres domaines linguistiques comme la sémantique (Ben-David et al., 2016) ou la syntaxe (Degano et al., 2024).

3.1 De la pensée aux mots : modèles de production de la parole

Depuis plus d'un siècle, les théories de production de la parole ont contribué à la compréhension du fonctionnement du langage humain. Le modèle sériel de Levelt (1983) est le premier à inclure trois étapes principales lors de la production de la parole : la conceptualisation du message en pensée, sa formulation mentale, puis sa mise en articulation. Par contraste, le modèle connexionniste proposé par Dell (1986) met en avant les interactions multidirectionnelles entre la sémantique, la sélection lexicale et la phonologie. Cette vision dynamique du langage permet d'expliquer les erreurs phonémiques et sémantiques. Aujourd'hui, les théories du langage ont été enrichies et améliorées grâce à l'apport des neurosciences comportementales (voir p. ex., Kemmerer [2022] pour une revue complète sur le sujet).

La production d'un message oral est générée par un enchaînement de processus linguistiques et non-linguistiques, parfois simultanés, qui interagissent entre eux. Elle implique de récupérer des concepts en mémoire sémantique puis de sélectionner les items lexicaux parmi le stock disponible (Hagoort, 2017). Les items lexicaux choisis font alors l'objet d'un encodage phonologique, c'est-à-dire l'association d'un concept à la séquence de phonèmes qui lui est attribuée, puis d'un encodage phonétique, c'est-à-dire la transformation des sons de la parole en gestes moteurs articulatoires (Hagoort, 2017). Enfin, lors de la phonation, l'activité pulmonaire se coordonne à la mobilité des structures laryngées et articulatoires. Presque simultanément, le retour acoustique au locuteur de sa propre voix lui permet de détecter ses erreurs de production. La figure 8, issue de Voleti et al. (2020), illustre les processus linguistiques impliqués dans la production de la parole, de la conceptualisation d'une idée à sa réalisation vocale.

Figure 8. Schéma des processus linguistiques impliqués dans la production de la parole.



Note : Reproduit, avec permission, de Voleti et al. (2020). 2019 IEEE.

3.2 Substrats neuronaux du langage

Il est aujourd'hui incontestable que le langage ne se réduit pas à une fonction isolée et limitée à deux régions cérébrales, comme l'avaient suggéré les neurologues des siècles passés. Le modèle Broca-Wernicke, ainsi que ses variantes, a été largement dépassé, en particulier en raison de son manque de précision spatiale et de sa localisation restreinte (Tremblay & Dick, 2016). À l'inverse, le langage, dont la parole constitue la manifestation sonore, interagit de manière complexe avec d'autres fonctions cognitives telles que la mémoire, les fonctions exécutives, les processus attentionnels et les fonctions motrices. Ce postulat est renforcé par des études utilisant l'IRMf, qui ont démontré que le langage active des régions cérébrales qui ne sont pas exclusivement associées à des fonctions linguistiques (Duffau, 2003). Par exemple, la programmation d'un message en vue de son articulation engage des fonctions motrices (Tate et al., 2014). De même, la sélection lexicale active l'aire prémotrice supplémentaire. Plus précisément, le contrôle moteur et exécutif du langage s'effectuerait par l'intermédiaire du faisceau frontal oblique, qui relie le gyrus frontal latéral inférieur à l'aire prémotrice supplémentaire (Tremblay & Gracco, 2010).

Système sémantique et sélection lexicale

La sémantique lexicale réfère à des formes linguistiques (une séquence de phonèmes dans le cas de la parole) arbitrairement associées à des informations conceptuelles (Reilly et al., 2025). Le système sémantique, composante essentielle de la mémoire à long terme, serait responsable du stockage des connaissances relatives aux concepts. Néanmoins, plusieurs études remettent en question cette hypothèse : les connaissances sont-elles effectivement localisées dans une région cérébrale unique ou bien leur localisation varie-t-elle en fonction de la nature des connaissances elles-mêmes ? Il apparaît que les concepts d'actions seraient davantage associés aux régions frontales, tandis que le traitement lexical et sémantique des objets solliciterait principalement les régions temporales inférieures de l'hémisphère gauche (Antonucci & Reilly, 2008). Plus récemment, Borghesani et al. (2016) ont exploré la représentation multidimensionnelle des mots concrets, c'est-à-dire des concepts ayant un référent physique perceptible. Lors d'une expérience en IRMf, les participants étaient invités à évaluer le degré de similarité entre deux mots écrits. La principale distinction entre ces mots résidait dans la taille réelle des objets référencés (par exemple, « souris » vs. « girafe »). Lors de la lecture silencieuse, la voie ventrale du cortex visuel s'activait en premier, suivie par l'activation des régions temporales antérieures, impliquées dans le traitement des informations sémantiques plus abstraites (Borghesani et al., 2016). Autrement dit, le traitement sémantique des mots écrits s'effectuait initialement de manière perceptuelle par les régions visuelles, indépendamment de la longueur des mots, avant de mobiliser les régions temporales antérieures pour un traitement plus conceptuel.

La sélection lexicale lors de la formulation du message est sans doute l'un des processus les plus étudiés dans le domaine de la psycholinguistique. Les travaux de Levelt et de son équipe (notamment Levelt et al. [1999]) ont introduit le concept de compétition lexicale, qui désigne l'activation par un lexème (c.-à-d. le radical d'un « mot ») d'un réseau sémantique de lemmes en concurrence. La sélection du lemme (c.-à-d. une unité sémantique formant un « mot ») adéquat implique l'existence d'une représentation conceptuelle unifiée pour chaque mot du lexique (ou concept lexical). Ces représentations seraient organisées au sein d'un réseau sémantique, et l'activation simultanée de lemmes alternatifs engendrerait une compétition ralentissant ainsi le processus de sélection. En d'autres termes, plus un concept lexical possède de voisins sémantiques, plus la durée de sa sélection serait prolongée.

Cependant, de nombreuses études ultérieures remettent en question la validité de la théorie de la sélection par compétition lexicale. À travers une série d'expériences, Mahon et al. (2007) démontrent que les voisins sémantiques d'un mot cible (p. ex., « cheval » et « zèbre ») accélèrent la reconnaissance de ce dernier, contrairement aux mots sémantiquement éloignés (p. ex., « cheval » et « baleine »). Ainsi, les interférences sémantiques ne semblent pas intervenir au niveau de la recherche lexicale. Par ailleurs, l'analyse des mouvements oculaires concomitants à la production de parole apporte également des preuves allant à l'encontre de l'idée de compétition lexicale (Mahon et al., 2007). Huettig et Hartsuiker (2008) ont mené une expérience où les participants devaient nommer un objet cible parmi un distracteur sémantique (co-hyponyme), un distracteur visuel et un distracteur sans

lien avec la cible. Les chercheurs ont observé que les distracteurs sémantiques et visuels étaient plus fixés que les distracteurs sans lien, ce qui met en évidence l'implication du traitement visuel dans la production d'un mot. Du reste, aucune latence supplémentaire n'a été relevée en présence de distracteurs visuels ou sémantiques, suggérant l'absence d'interférence sémantique dans le processus de sélection lexicale (Huettig & Hartsuiker, 2008).

La récupération sémantico-lexicale désigne l'accès à une information spécifique au langage stockée en mémoire. La récupération d'un concept mobiliserait plusieurs régions cérébrales, telles que le cortex frontal inférieur gauche, les cortex temporaux supérieurs et moyens gauches, ainsi que le cortex pariétal inférieur gauche (Hagoort, 2017). Cette activation de multiples zones cérébrales témoigne de l'interaction des fonctions linguistiques avec d'autres processus cognitifs. En conséquence, l'accès au sens des mots impliquerait non seulement leur stockage dans la mémoire sémantique, mais également le contrôle exécutif des connaissances sémantiques (Antonucci & Reilly, 2008 ; Thompson et al., 2018).

Aspects prosodiques

Les recherches portant sur les bases neuronales de la prosodie se sont principalement orientées vers ses fonctions pragmatiques, notamment en ce qui concerne l'expression des émotions. Pour cette raison, elle est fréquemment associée à l'hémisphère droit. Par exemple, l'étude de Sammler et al. (2015) s'est intéressée aux régions cérébrales sollicitées lors de la perception de l'intonation. Les participants ont écouté des mots isolés prononcés avec un ton interrogatif ou déclaratif, afin de limiter l'activation des régions liées au traitement syntaxique. Les analyses d'IRMf ont révélé que le traitement de l'intonation entraînait une activation des voies dorsales et ventrales de l'hémisphère droit, ainsi que celle du cortex prémoteur. Néanmoins, la notion de latéralisation de la prosodie à l'hémisphère droit est remise en cause par d'autres études qui suggèrent plutôt une répartition interhémisphérique. En effet, certains patients souffrant de lésions cérébrales dans l'hémisphère droit parviennent à accomplir des tâches d'identification de la prosodie (Baum & Pell, 1999). L'hémisphère gauche semblerait ainsi impliqué dans le contrôle de la durée en tant qu'indice prosodique (Baum & Pell, 1999). La divergence entre ces recherches peut résider dans le fait que la première étude assimile la prosodie à l'intonation, tandis que la seconde inclut également le paramètre de la durée. Cependant, davantage de recherches sont requises afin de mieux cerner les processus sous-jacents à la production et à la perception de la prosodie linguistique, notamment au sein des langues accentuelles, où les paramètres prosodiques jouent un rôle structurant dans la transmission du message (Grice & Kügler, 2021).

Le langage est donc une fonction cognitive composite, qui, selon les besoins des locuteurs, sollicite divers aspects de la cognition et recrute différentes régions et faisceaux cérébraux à des niveaux d'implication variables. Si les connaissances sur le fonctionnement du langage sont rendues possibles grâce à celles sur la cognition, l'inverse est également acceptable. Prise sous un certain

angle, l'étude de la parole, qui est une production audible de l'activité cérébrale pourrait donc renseigner sur les processus qu'elle mobilise.

3.3 Anomie : Comprendre la cognition par l'erreur

La production verbale est souvent ponctuée d'erreurs sémantiques, phonétiques ou morphosyntaxiques, c.-à-d. la réalisation involontaire de lemmes, de phonèmes ou de morphèmes qui ne correspondent pas au contexte attendu par le locuteur ou l'auditeur. Leur fréquence et leur nature dépendent de divers facteurs, tels que les aptitudes du locuteur, le contexte communicationnel ou encore les caractéristiques spécifiques au genre discursif. Une analyse rigoureuse de ces erreurs linguistiques permet de distinguer celles qui relèvent d'une variation normale de celles qui témoignent d'une pathologie. Étant donné l'interaction du langage avec d'autres fonctions cognitives, l'altération de la parole est souvent le premier indice d'un changement cognitif sous-jacent.

L'étude des erreurs sémantiques et lexicales suscite depuis longtemps un intérêt particulier, notamment dans le but d'élucider les mécanismes de fonctionnement du langage. Le phénomène de manque du mot, ou anomie, désigne une perturbation de la production lexicale, affectant principalement des éléments lexicaux appartenant à des classes ouvertes. L'expression « avoir le mot sur le bout de la langue » incarne un phénomène bien connu de tous, dont l'intensité est souvent amplifiée par des facteurs tels que la fatigue et le stress. Toutefois, il est considéré pathologique lorsqu'il survient de manière récurrente sur une longue période et occasionne une gêne perceptible chez l'individu affecté. Il n'est donc guère surprenant que le manque du mot soit un symptôme commun à de multiples troubles et syndromes cognitifs. Ce dernier est emblématique des aphasies non fluentes secondaires à, par exemple, un accident vasculaire cérébral ou une maladie d'Alzheimer. Cependant, ces anomies trouvent leurs racines dans des processus physiopathologiques distincts et leur étude approfondie offre un éclairage précieux sur les mécanismes cognitifs sous-jacents perturbés.

Dans les années 1990, Caramazza et Hillis (1990) ont proposé un modèle théorique de la production d'erreurs sémantiques fondé sur l'analyse de cas cliniques d'aphasie. Selon ce modèle, les erreurs langagières en production orale peuvent résulter soit d'une altération des connaissances sémantiques d'un mot, soit d'un dysfonctionnement dans le processus d'accès à sa forme phonologique. En cas de déficit sémantique, les informations associées au concept sont « perdues », ce qui conduit à des erreurs systématiques : l'individu produit toujours le même type d'erreur et confond sans en avoir conscience ce mot avec un autre dont la forme visuelle ou phonologique est similaire. Ce phénomène est fréquemment accompagné de l'impression d'un allongement du délai entre le choix lexical et la production orale, bien qu'en réalité, il n'y ait pas de véritable sélection lexicale et que le mot attendu ne soit pas produit. En revanche, lorsqu'un manque du mot est induit par un déficit exécutif, les connaissances sémantiques liées au mot sont préservées, mais le locuteur éprouve des difficultés à accéder à sa forme phonologique. Ce dernier, en ayant

conscience de ses difficultés, peut alors choisir délibérément un mot de sens ou de forme proche, ou recourir à une description ou un synonyme qui s'adapte relativement bien au contexte. Les erreurs observées dans ce cas sont plus variables, et le mot recherché peut réapparaître après un certain temps. Une augmentation du délai entre la sélection lexicale et la production orale est également constatée (Caramazza & Hillis, 1990, 1991). Par ailleurs, des recherches ont mis en évidence l'impact de l'âge sur le phénomène du manque du mot, en particulier pour les mots à faible fréquence lexicale et les noms propres (Burke et al., 1991).

L'examen des erreurs sémantiques chez les patients atteints de la maladie d'Alzheimer et de démence sémantique permet de différencier ces deux pathologies (Reilly et al., 2011). En effet, leurs symptômes linguistiques paraîssent similaires, bien que leurs physiopathologies diffèrent. La maladie d'Alzheimer est une pathologie neurodégénérative qui affecte principalement la mémoire épisodique, c'est-à-dire la mémoire des événements et expériences vécus. La démence sémantique, quant à elle, est souvent précurseur d'une dégénérescence lobaire fronto-temporale et touche la mémoire sémantique, c'est-à-dire la connaissance des concepts. Lorsque ces populations sont comparées, l'anomie est nettement plus marquée dans les cas de démence sémantique. Elle est notamment caractérisée par une capacité de récupération lexicale moins efficace et un nombre d'erreurs plus élevé que dans la maladie d'Alzheimer (Reilly et al., 2011). En outre, les patients atteints de la maladie d'Alzheimer commettent davantage d'erreurs liées à des mots co-hyponymes (p. ex., « citron » et « fraise » sont deux co-hyponymes de « fruit »), tandis que ceux souffrant de démence sémantique tendent à utiliser plus fréquemment des circonlocutions pour décrire la fonction de l'objet (Reilly et al., 2011). Cette étude souligne que les bases neuronales du phénomène du manque du mot varient en fonction de l'étiologie sous-jacente. Dans le cadre d'un syndrome cognitif pour lequel le diagnostic précis de l'étiologie n'a pas encore été posé, l'analyse du manque du mot peut offrir des pistes pour formuler des hypothèses sur les mécanismes cognitifs impliqués. En ce qui concerne les patients en oncologie, les données de neuro-anatomie ne mettent pas en évidence la présence de lésions cérébrales pouvant causer une anomie (Bernstein et al., 2021). Le manque du mot étant une plainte couramment décrite par cette population, la relation entre le CRCI et anomie reste donc encore à explorer.

3.4 Interférences lexicales sur la prosodie

Lors de l'articulation d'un message, le temps de récupération d'un concept ou de sa forme phonémique en mémoire peut se manifester par une ou plusieurs ruptures de la chaîne parlée (Don & Lickley, 2015 ; Hartsuiker & Notebaert, 2010). Ces ruptures affectent directement le rythme de la parole, c'est-à-dire la « régularité » temporelle des accents qu'elle porte (Astésano, 2016 ; Tilsen, 2019). Depuis les années 1950, la prosodie fait l'objet de nombreuses recherches pour comprendre les mécanismes concomitants à la sélection lexicale. Pour Goldman-Eisler (1958), par exemple, l'acte de sélection entre des termes partageant des propriétés sémantiques proches se traduirait par l'apparition d'une pause remplie. D'autres travaux avancent que la recherche de mots rares est fréquemment associée à des hésitations, lesquelles se manifestent par des pauses silencieuses

(Beattie & Butterworth, 1979). Une étude portant sur l'interférence sémantique, c'est-à-dire le délai de dénomination induit par la présence d'un distracteur sémantique, révèle que les difficultés d'accès lexical chez un locuteur sain génèrent davantage de disfluences, telles que des réparations ou des pauses silencieuses, sans toutefois augmenter le nombre d'erreurs sémantiques (Rapoeye et al., 2023). Néanmoins, les résultats de cette étude reposent sur des données de parole manipulée, ce qui en limite la validité écologique. De même, la présentation des stimuli différait selon la cible, présentée sous forme de mots écrits, et le distracteur, présenté sous forme d'image. Il est donc également possible que les disfluences observées soient en partie dues au traitement visuel des stimuli plutôt qu'à une interférence sémantique.

Ces études, à l'instar de nombreuses recherches récentes, mettent en évidence l'influence de la recherche lexicale sur la prosodie. Toutefois, une critique fondamentale pouvant leur être adressée réside dans l'absence de rigueur quant à la définition et à la mesure des phénomènes de rupture. En effet, les concepts de disfluences, de pauses et d'hésitations demeurent vagues au sein de la littérature. Ce point sera exploré en détail dans la section suivante.

3.5 Interruptions du flux de parole

La parole continue, par contraste avec les mots isolés, se caractérise par une succession de phonèmes adjacents perçus et traités comme une séquence régulière (Ash & Grossman, 2015). Pour autant, la parole n'est pas intrinsèquement rythmique, puisque le signal acoustique qui la compose demeure un signal apériodique (Turk & Shattuck-Hufnagel, 2013). Cependant, ce dernier manifeste des « régularités temporelles » qui sont perçues par l'oreille humaine comme étant « quasi-rythmiques » (Poeppel & Assaneo, 2020, p. 323), et traitées de manière similaire par le cerveau. La production de la parole est rarement linéaire. Elle consiste en des éléments rythmiques et intonatifs qui, à des degrés divers, contribuent à sa structuration. L'expression « message oral » (ici envisagé comme une séquence d'éléments linguistiques, indépendamment de l'échelle de l'unité linguistique — phonème, syllabe, etc.) remplace l'expression abstraite d'« énoncé » souvent utilisée dans ce contexte. Un message oral se compose de segments prosodiques dont la longueur dépend des capacités de mémoire de travail et de motricité articulatoire du locuteur. Ces segments jouent également un rôle essentiel dans la compréhension du message par l'auditeur. En segmentant les unités syllabiques à des niveaux lexicaux et syntaxiques, les éléments prosodiques facilitent ainsi le traitement sémantique du message.

Disfluences

Certains éléments prosodiques marquent des ruptures dans le flux de la parole, signalant ainsi un changement perceptible de rythme et d'intonation. Une proposition de classification distingue deux types d'interruptions : les interruptions suspensives, qui retardent la progression du message, et les interruptions disfluentes, qui rompent le message et peuvent potentiellement en initier un nouveau

(Pallaud et al., 2019). Dans ce dernier cas, la rupture du message est interprétée par l'auditeur comme un signe d'inachèvement.

Le sujet des disfluences a été largement étudié, mais il n'existe pas de consensus clair quant à leur terminologie. Classiquement, l'appellation « disfluence » inclut les pauses remplies, les pauses silencieuses, les allongements vocaliques et parfois consonantiques, ainsi que les faux départs, les répétitions, les abandons et les réparations (Pallaud et al., 2019). Toutefois, la définition même de disfluence varie considérablement selon les auteurs. En résumé, une disfluence est l'interprétation d'une interruption du flux verbal perçue comme inattendue, soit par l'auditeur (Lickley & Bard, 1998), soit par le locuteur lui-même, qui pourrait alors discerner une rupture dans sa boucle phonologique (c.-à-d. le retour auditif de sa propre parole) (Huettig & Hartsuiker, 2010). Les disfluences se distinguent par leur nature non prévisible, bien que certaines régularités puissent être observées. En d'autres termes, les disfluences perturbent la rythmicité attendue de la parole. Les disfluences sont dites « atypiques » lorsqu'elles se manifestent dans le cadre de troubles des sons de la parole, comme le bégaiement (Lu et al., 2022). Dans cette condition, elles présentent des spécificités prosodiques et phonétiques liées à la réalisation articulatoire des phonèmes. À l'inverse, les disfluences « typiques » surviennent chez tous les locuteurs et sont généralement attribué à des processus cognitifs de haut niveau, tels que la sélection lexicale ou la planification syntaxique (Lickley, 2017). Leur fréquence élevée pourrait indiquer la présence de difficultés cognitives sous-jacentes. Selon les études, le taux de disfluences varierait de 6 à 10 disfluences pour 100 mots chez un locuteur sain (Bortfeld et al., 2001 ; Shriberg, 2001). Bortfeld et al. (2001) ont analysé les disfluences (répétitions, faux départs et *fillers*) en situation conversationnelle suscitée par un jeu de rôle. Dans leur analyse, les auteurs ont omis pauses silencieuses n'ont pas été ainsi que les énoncés de modalisation (c.-à-d. commentaires sur la production du message), car ces derniers étaient trop rares dans le corpus. Lorsque plusieurs types de disfluences étaient adjacents, chaque occurrence était comptabilisée séparément. Le taux de disfluences était calculé sur la base de cent mots, bien que la définition exacte de ce que représentait un mot n'ait pas été précisée. L'étude met en évidence un effet de l'âge, du genre (les hommes ayant produit plus de *fillers*) et de la familiarité du sujet (c.-à-d. description de photographies d'enfants vs. photographies de tangram) sur la production de disfluences. Cette étude présente néanmoins une limite majeure. En effet, elle repose sur une transcription orthographique des productions, bien que la parole soit avant tout un phénomène acoustique.

Dans le cadre de mesures objectives de la parole, la notion de « disfluence » pourrait être problématique à plusieurs niveaux. D'une part, les disfluences constituent une catégorie subjective et leur observation dépend fortement de leur définition. La variabilité inter-juge liée cette taxonomie empêche la reproductibilité des expériences. D'autre part, la majorité des recherches sur les disfluences ont été menées sur des corpus en anglais, et dans une moindre mesure en français. Or, leur fréquence et leur durée sont susceptibles de varier selon les systèmes linguistiques dans lesquels elles sont produites (Crible et al., 2017). Par ailleurs, le contenu et la forme d'un message verbal sont contexte-dépendants, au sens qu'ils sont fonction de la situation interactionnelle (Kachkovskaia & Kocharov, 2024). De plus, ils sont propres au locuteur, selon ses habitudes linguistiques et les facteurs

psychologiques qui l'influencent (p. ex., Cummins et al. [2015]). Enfin, l'idée que les disfluences soient distinguables du fait de leur caractère inopiné est discutable, car certaines disfluences sont prédictibles. Un interlocuteur peut, par exemple, anticiper une interruption pour prendre le tour de parole. Par conséquent, plutôt que de parler de disfluences, il semble plus approprié de se référer à ces ruptures prosodiques comme des interruptions du flux de parole. En effet, il s'agit de phénomènes physiques mesurables au moyen d'instruments d'analyse acoustique.

Pauses

Parmi les interruptions de parole, les pauses tiennent une place particulière. Sous l'étiquette « pauses » sont regroupées les pauses silencieuses et les pauses remplies par des éléments voisés non porteurs de sens (p. ex., « euh »).

La question du seuil des pauses silencieuses reste largement débattue à travers la littérature sur la parole. En effet, tout silence est-il une « pause » ? Les travaux de Goldman-Eisler (1958, 1961) ont été parmi les premiers à s'intéresser à la durée des pauses. Goldman-Eisler suggère qu'une valeur minimale de 250 ms reflèterait des processus linguistiques en activité. Cependant, l'établissement de ce seuil repose sur de la parole lue, et la méthodologie employée dans cette étude a été critiquée a posteriori (Boomer, 1970). Néanmoins, ces travaux posent les jalons de la recherche sur le lien entre pauses silencieuses et processus cognitifs sous-jacents. Comme le remarque Lickley (2015), ce seuil arbitraire était probablement lié aux limitations technologiques de l'époque pour l'analyse de la parole. Depuis, différents seuils ont été testés dans des configurations syntactiques variées (p. ex., pauses entre syntagmes, pauses au sein des mots, etc.). Pour l'heure, le seuil minimal de 200 ms fixé par Boomer (1965) a été repris par de nombreuses études ultérieures. De même, Goldman et al. (2010) proposent que les pauses au-delà de 200-250 ms signaleraient un effort cognitif, tandis que les pauses en dessous seraient plutôt liées à des contraintes physiologiques (déglutition, respiration, articulation). Toutefois, il paraît délicat de séparer l'activité mécanique des processus cognitifs. Dès lors, certaines recherches estiment qu'il n'est pas nécessaire de définir un seuil minimal (Campione & Véronis, 2002). Toutefois, l'absence de seuil minimal augmente le risque d'inclure des pauses qui seraient en réalité liées à des caractéristiques phonétiques, comme les *burst* de phonèmes consonantiques (Campione & Véronis, 2002). Quelques études ont se sont penchées sur un seuil maximal des pauses (p. ex., Candea [2000] et Grosjean et Deschamps [1972]). Les très longues pauses sont rares et sont surtout produites en parole spontanée. Les supprimer d'un corpus reviendrait à modifier la distribution générale des pauses et risque de rendre leur interprétation hasardeuse (Campione & Véronis, 2002). En outre, ces pauses longues peuvent se révéler particulièrement intéressantes en contexte pathologique.

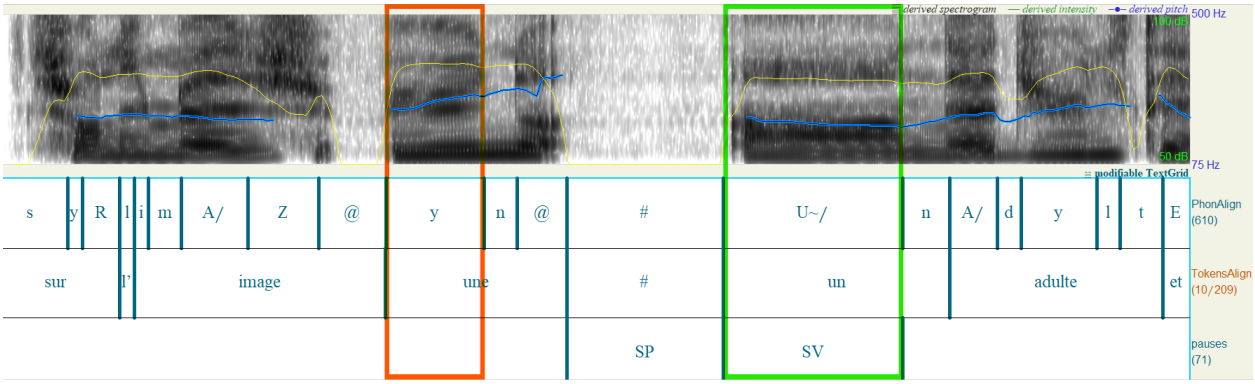
D'autre part, plusieurs recherches suggèrent que les pauses silencieuses auraient une fonction grammaticale ou pragmatique (Crible et al., 2022). Certaines pauses auraient une fonction démarcative dès lors qu'elles se situent entre deux propositions, tandis que d'autres pauses seraient

interprétées comme une marque d'hésitation (Campione & Véronis, 2005). Une pause silencieuse serait perçue « disfluente » lorsqu'elle rompt la syntaxe d'un segment verbal (Pallaud et al., 2019). Cependant, l'attribution de fonctions aux silences interroge. D'une part, le fait que chaque pause bénéficie d'une fonction spécifique écarte la possibilité que plusieurs fonctions puissent coexister simultanément. En effet, il est tout à fait possible qu'un locuteur éprouve des difficultés d'accès au lexique à une frontière prosodique majeure. Dans cette situation, le locuteur peinerait à trouver un concept, ou la forme phonologique qui lui est associée, commençant le syntagme. Cette pause serait alors étiquetée comme pause démarcative, et non comme pause d'hésitation. D'autre part, l'identification des pauses disfluentes repose principalement sur leur position au sein d'un groupe rythmique. Or, s'il existe un lien fort entre la prosodie et la syntaxe, la première montre une relative indépendance vis-à-vis de la seconde (Ferreira, 1993). De plus, les éléments prosodiques, paramètres temporels, échappent à la localisation spatiale de la syntaxe. Pour finir, notons que tout silence n'est pas lié à une pause cognitive. En effet, les silences interviennent régulièrement dans un discours fluide, entre les tours de parole, au sein des syntagmes ou encore en fin d'émission d'un son consonantique (Lickley, 2017). Sur un plan perceptif, les silences aident à la segmentation de la parole en accentuant la proéminence des structures syntaxiques (Duez, 1985). De fait, il n'existe pas de différence physique entre ces silences, hormis leur durée (Cole, 2015).

Les pauses remplies, quant à elles, auraient principalement une fonction pragmatique. En émettant un « euh » ou équivalent, le locuteur exprimerait le souhait de conserver son tour de parole (Clark, 2002). De, les pauses remplies seraient la marque voisée d'un effort cognitif (Clark, 2002 ; Simon & Christodoulides, 2016). Elles peuvent également contribuer aux effets de stylistique. En outre, il est parfois admis que, en tant que « *filler* », les pauses remplies faciliteraient le traitement du message. Corley et Hartsuiker (2011) ont testé cette hypothèse en comparant le temps de réaction lors d'une tâche de reconnaissance lexicale dans trois expériences. Les mots cibles étaient immédiatement précédés d'une pause remplie, d'un silence ou d'une tonalité. Le temps de réaction des participants pour détecter les mots cibles était similaire dans les trois expériences. Selon les auteurs, ce résultat suggère que tout délai, quelle que soit sa nature, aide à la reconnaissance lexicale (Corley & Hartsuiker, 2011).

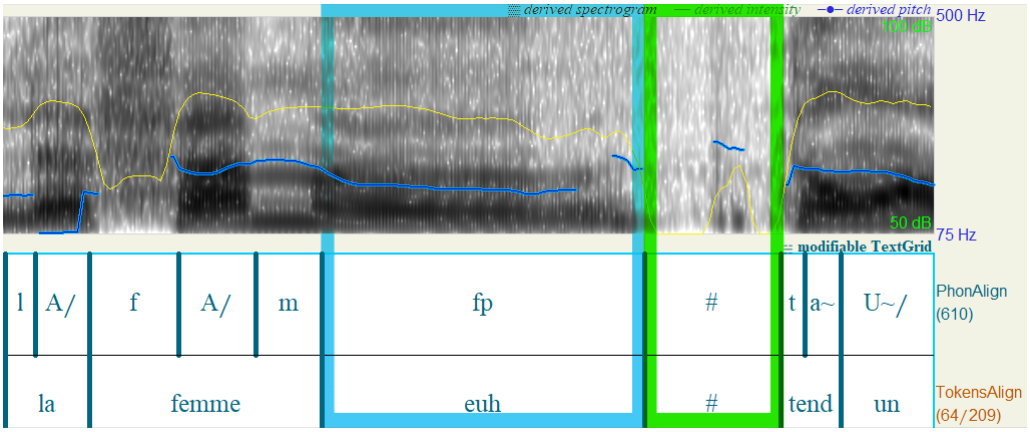
Enfin, les prolongations, ou voyelles allongées, sont des voyelles dont la durée d'émission dépasse celle intrinsèquement nécessaire à leur réalisation phonétique. En français, les voyelles allongées qui portent un accent deviennent un signal perceptible de frontière prosodique et sont appelées prolongations finales lorsqu'elles se situent en fin de groupe rythmique. Leur accentuation se manifeste par un contour intonatif caractéristique, c'est-à-dire par des variations de la fréquence fondamentale (f0) qui peut être montante, descendante ou montant-descendante. Lorsque les voyelles prolongées marquent un contour intonatif plat, elles sont, à l'instar de pauses remplies, perçues comme des indicateurs d'hésitation (Belz, 2023). La figure 9 représente la différence entre une voyelle prolongée accentuée (encadré orange) et une voyelle interprétée comme une pause remplie (encadré vert).

Figure 9. Spectrogrammes d'une voyelle prolongée accentuée (encadré orange) et d'une voyelle prolongée interprétée comme une pause.



Pour conclure cette section, les pauses sont définies dans cette thèse comme des interruptions du flux de la parole à un niveau abstrait. De même que la perception des phonèmes est continue (McMurray, 2022), nous proposons que, d'un point de vue acoustique, les pauses constituent un continuum allant de la chute brutale de la fréquence fondamentale à l'absence de réalisation phonémique. Les pauses remplies concernent donc des voyelles isolées perçues comme des « *fillers* » (« euh » ou « hum ») ou des syllabes dont la durée du noyau est prolongée. Le contour intonatif de ces phonèmes est maintenu plat tout au long de l'émission (Campione & Véronis, 2004). À l'extrémité du spectre des interruptions, les pauses silencieuses réfèrent à des situations où aucune émission de phonème n'est produite. Cette situation s'étend de l'inspiration audible au silence total en passant par les bruits bucco-pharyngés. La figure 10 illustre le continuum des pauses par une pause remplie (encadré bleu) et une pause silencieuse (encadré vert).

Figure 10. Spectrogrammes d'une pause remplie (encadré bleu) et d'une pause silencieuse (encadré vert).



3.6 Limites de l'analyse de la parole

La recherche lexicale, et par extension le manque du mot, ralentit non seulement le débit de parole, mais entraîne également une diminution de l'intonation et de l'intensité des syllabes produites lors de la recherche lexicale. De fait, l'examen du signal acoustique rend possible la formulation d'hypothèses quant aux corrélats cognitifs des interruptions de parole. Cependant, l'analyse de la parole soulève plusieurs questions d'ordre méthodologique.

D'une part, la prosodie est sujette à de fortes variations, tant entre les individus qu'au sein de la parole d'un même locuteur. Les caractéristiques propres à chaque système linguistique (p. ex., les différences entre une langue tonale et une langue à accentuation) ou des variables sociolinguistiques (p. ex., l'appartenance d'un locuteur à des groupes sociaux et ethniques, Holliday [2021]), ainsi que des aspects physiologiques (sexe, âge, intégrité des structures oro-laryngées) modulent la production de la prosodie. Ces variations s'expriment d'autant plus à travers la parole continue, qui, par sa nature dynamique, est difficilement contrôlable en laboratoire.

D'autre part, les recherches sur la prosodie sont nombreuses, mais peu d'entre elles se fondent sur des données acoustiques empiriques. Ces théories sont difficilement généralisables à des populations cliniques. De plus, pour gagner en robustesse, l'analyse du langage requiert des techniques d'investigation complémentaires, telles que l'oculométrie. Ces techniques ont le potentiel de valider ou d'infirmer les hypothèses issues de l'observation des phénomènes prosodiques. Leur éclairage permet de mieux comprendre les mécanismes cognitifs sous-jacents aux interruptions de parole.

MESSAGES CLÉS — LA PAROLE : UNE OUVERTURE SUR LA COGNITION

1. **Production d'un message verbal**

   - La production de la parole est une fonction dynamique de la cognition. Elle entraîne une cascade de processus linguistiques et non linguistiques, tels que la recherche lexicale, l'encodage phonologique, etc. Ces processus, nécessaires à la conceptualisation, à la formulation et à l'articulation d'un message verbal, recrutent des faisceaux cérébraux au-delà des régions associées au langage.

   - L'accès au sens des mots nécessiterait non seulement leur stockage dans la mémoire sémantique, mais également le contrôle exécutif des connaissances. Le déficit de l'un de ces mécanismes peut susciter des difficultés de recherche lexicale (manque du mot).

   - L'altération de la parole est souvent le premier indice d'un changement cognitif sous-jacent. L'analyse de la parole est un outil pertinent pour identifier et comprendre des troubles subtils de la cognition. Dans le cas des survivants d'un cancer, le manque du mot est le symptôme prépondérant du CRCI.

2. **Paramètres prosodiques de changements cognitifs**

   - Le manque du mot affecte la temporalité de la parole. Des variations prosodiques sont alors observées, tels que l'allongement de la durée des pauses.

   - Les pauses sont des interruptions de parole souvent examinées pour investiguer la cognition. Elles constituent un continuum s'échelonnant de la chute brutale de la fréquence fondamentale (voyelle prolongée et pause remplie) à l'absence de réalisation phonémique (pause silencieuse).

   - La multiplicité des paramètres de parole utilisés en recherche brouille la sélection de marqueurs potentiels de CRCI. La réalisation d'un inventaire des paramètres existants est nécessaire pour choisir un panel d'indices prosodiques susceptibles de repérer un CRCI.

4 Investiguer les processus linguistiques par le regard

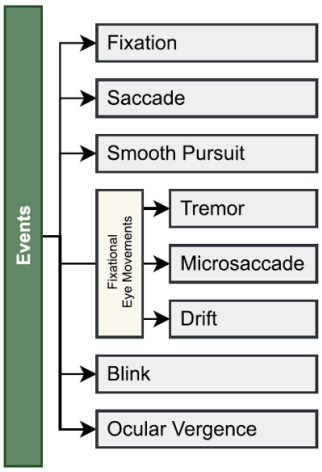
Chez l'humain, la vision constitue le système sensoriel le plus développé (Leigh & Zee, 2015). L'étude du comportement oculomoteur offre un précieux aperçu du fonctionnement des processus

cognitifs. Elle permet de mettre en lumière les liens entre certaines pathologies neurologiques et des anomalies des mouvements oculaires, telles que le nystagmus ou les *Rapid Eye Movements*, qui seraient directement associées à des propriétés biophysiques (Leigh & Zee, 2015). Certains types de mouvements oculaires sont spécifiques à des pathologies neurodégénératives, et leur analyse aide au diagnostic de celles-ci (Sekar et al., 2024). D'un point de vue cognitif, l'analyse oculomotrice renseigne sur divers processus tels que la prise de décision (Orquin & Mueller Loose, 2013) ou encore le multitâche (Huestegge, 2011). En psycholinguistique, l'oculométrie s'avère particulièrement utile pour déchiffrer la cascade d'opérations cognitives activées dans la production verbale (Griffin & Spieler, 2006).

4.1 Principes fondamentaux du comportement oculomoteur

La recherche visuelle, qui sollicite principalement les lobes frontaux et pariétaux, vise à minimiser la demande attentionnelle tout en optimisant l'acquisition de nouvelles informations fovéales. Lors de cette recherche, la fovéa, soit la partie de la rétine où la vision est la plus précise, doit en effet se déplacer de manière séquentielle pour explorer les points d'intérêt apparus en périphérie. Différents types d'événements oculaires peuvent être mesurés par l'oculométrie, dont deux sont les plus couramment étudiés : les saccades et les fixations. La figure 11, issue de Mahanama et al. (2022), présente une classification des événements oculomoteurs mesurables par une technique d'oculométrie.

Figure 11. Classification des événements oculomoteurs mesurables par une technique d'oculométrie.

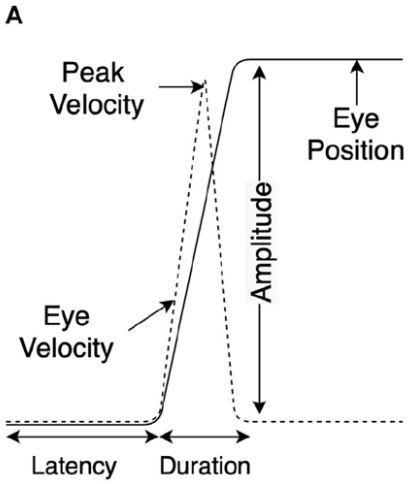


Note : Reproduit, avec permission, de Mahanama et al. (2022). 2022 Frontiers.

Les saccades sont des mouvements rapides du globe oculaire. Les saccades « volontaires », motivées par des facteurs internes, ont pour but d'explorer l'environnement visuel. Elles dirigent la fovéa vers une cible sans mouvement de la tête, en réponse à un stimulus visuel ou auditif en périphérie. Ces mouvements sont rapides et précis afin de ne pas perturber le traitement visuel. En

général, une latence d'environ 200 ms sépare l'apparition du stimulus de la réalisation de la saccade. Ce délai correspond au temps de traitement neuronal impliquant le cortex cérébral rétinal, le colliculus supérieur, les ganglions de la base, le thalamus et le cervelet. Les saccades sont accompagnées de phases correctives qui ajustent la position de l'œil lorsqu'un mouvement oculaire sous-estime ou dépasse l'amplitude idéale pour atteindre la cible (Pélisson et al., 2010). La figure 12, tirée de Mahanama et al. (2022), illustre les mesures des saccades : l'amplitude du mouvement, la durée de la saccade en millisecondes et le pic de vélocité du mouvement, mesuré en degrés par seconde.

Figure 12. Schéma des mesures oculomotrices : amplitude du mouvement, durée de la saccade en millisecondes et pic de vélocité du mouvement en degrés par seconde.



Note : Reproduit, avec permission, de Mahanama et al. (2022). 2022 Frontiers.

Ces trois paramètres montrent une corrélation régulière appelée « *main sequence* ». En d'autres termes, la vitesse et l'amplitude d'une saccade dépendent de sa durée. Des déviations de cette séquence seraient spécifiques à certaines pathologies (Gibaldi & Sabatini, 2021). En effet, ces paramètres ne sont pas sciemment contrôlables, bien que la vitesse puisse être influencée par divers facteurs, tels que la luminosité ambiante, la présence de paupières fermées, ou encore le fait de viser une cible mémorisée (Leigh & Zee, 2015). La rapidité des saccades peut également varier selon la direction du mouvement et les positions initiales et finales de l'orbite. Enfin, la fatigue, en particulier la fatigue mentale, peut ralentir la vitesse des saccades (Di Stasi et al., 2014).

Les fixations, quant à elles, sont composées de micro-saccades dans une région d'intérêt délimitée durant une période suffisamment longue pour traiter 'information visuelle. L'œil n'est jamais complètement stable, surtout lorsqu'il fixe de petites cibles. Chaque fixation de la fovéa délivre donc une partie infime de l'information, de sorte que les détails sur le monde visuel sont acquis par une succession de fixations (Griffin & Spieler, 2006). Toutefois, bien que sa résolution soit moins précise, le traitement parafovéal permet également d'extraire les informations nécessaires à l'orientation de l'attention visuelle vers une nouvelle cible. La fixation d'un objet n'est donc pas toujours indispensable à son traitement visuel. En revanche, la probabilité de fixer une cible augmenterait lorsque la qualité de celle-ci est dégradée ou lorsque son niveau de détails est élevé

(Griffin & Spieler, 2006). Cet effet n'est pour autant pas retrouvé lors de la production de parole. Dans une étude sur l'impact de la dégradation de la qualité visuelle des objets sur la production de disfluences, Pistono et Hartsuiker (2023) ont non seulement constaté que la présentation d'images floutées n'entraînait pas une augmentation des disfluences, mais que le temps passé à explorer l'image décroissait. Les auteurs suggèrent que gêner l'accès à la représentation visuelle d'un objet ne ralentit pas son identification ni le processus de dénomination. L'idée selon laquelle une mauvaise qualité visuelle ralentirait la production orale est donc remise en question. Du reste, même dans des contextes simples, les individus ont tendance à fixer une cible lors de tâches cognitives coûteuses, telles que la production d'un message verbal. Le temps passé à fixer la cible refléterait alors la durée de traitement de cette tâche (Griffin & Bock, 2000). Il faut noter cependant que plusieurs parties d'un même objet peuvent être fixées de manière successive. Dans ce cas, l'analyse individuelle du temps de fixation sur chacune de ces parties ne donne pas d'information pertinente. Ces multiples fixations sur une même cible, appelées « *gaze* », doivent alors être mesurées comme un événement unique en définissant des régions d'intérêt.

4.2 Interaction entre parole et recherche visuelle

L'une des premières études à explorer la relation en temps réel entre le langage et les fonctions cognitives à l'aide de l'oculométrie est celle menée par R. M. Cooper en 1974 (Cooper, 1974). Ce travail a donné naissance au « paradigme du monde visuel », qui suggère que lorsqu'un auditeur entend un message verbal, il tend à diriger son regard vers des objets partageant une similarité visuelle, de forme ou de couleur par exemple, avec les items imageables contenus dans ce message (Eberhard et al., 1995). Huettig et Altmann (2005) ont par exemple observé qu'à l'écoute du mot « serpent », l'attention visuelle se tourne vers l'image d'un câble (Huettig & Altmann, 2005).

Vision incrémentale de la parole

En ce qui concerne le volet de la production de la parole, le paradigme des mondes visuels est couramment employé pour analyser les mécanismes de sélection lexicale. La haute résolution temporelle et spatiale du suivi oculaire permet d'étudier la parole à mesure qu'elle se déroule. Ainsi, le temps passé à regarder un élément informatif d'une image témoignerait des étapes de récupération lexicale et d'encodage phonologique des mots (Griffin & Bock, 2000). Parler serait donc un processus incrémental, c'est-à-dire que les mots seraient sélectionnés et phonologiquement encodés les uns à la suite des autres.

Dans un discours non préparé, un locuteur commence à articuler les premiers mots d'un message avant même d'avoir terminé l'encodage phonologique des mots suivants (Ferreira & Swets, 2002). La durée de fixation sur un item se trouve néanmoins raccourcie par un effet de récence observé lorsque le locuteur a déjà été confronté à l'item dénommé (Griffin & Spieler, 2006). La récupération

de cet item en lexique est donc facilitée par sa précédente activation. De même, la fréquence d'un mot, mais pas sa longueur, affecte la durée de sa sélection. Autrement dit, moins un mot est fréquent, plus sa recherche en mémoire sera longue. Meyer et al. (1998) ont comparé l'effet de l'intégrité d'une image (des dessins complets vs. des dessins auxquels il manque des traits) et l'effet de la fréquence lexicale sur le délai de dénomination d'item lexicaux imageables. La qualité dégradée des images n'induisait pas un temps de dénomination plus long que celui des images complètes. En outre, le temps de latence était plus élevé (entre 750 et 800 ms) lorsque les participants devaient nommer l'objet, que lorsqu'il devait simplement l'identifier sans le nommer (environ 650 ms). La durée est accrue par un effet lexical lors de la tâche de dénomination, alors que cet effet ne se produit pas lors de la tâche d'identification. Les auteurs concluent que l'effet de la fréquence lexicale se manifeste principalement au moment de l'accès aux noms des objets (Meyer et al., 1998).

Outre la recherche lexicale, la planification syntaxique peut également être éclaircie par l'ordre dans lequel les items sont fixés durant la production orale (Griffin & Spieler, 2006). En effet, les éléments sont généralement regardés dans l'ordre de leur apparition dans le discours (Griffin & Bock, 2000). Par exemple, sur présentation d'une image qui dépeint un agent effectuant une action sur un patient, un locuteur fixera d'abord l'agent puis le patient. Des allers-retours entre les éléments pourraient suggérer que les locuteurs n'ont pas encore pleinement décidé de la structure de leur message (Griffin, 2004).

Au regard de ces éléments, la vision incrémentale de la parole explique la présence d'interruptions du flux de parole même dans un discours fluent. L'étude des mouvements oculaires pourrait donc se révéler particulièrement informative pour comprendre le manque du mot dans une population n'ayant pas de troubles du langage à proprement parler. En effet, le comportement oculaire concomitant à la recherche d'un mot chez des patients en oncologie pourrait éclairer sur les mécanismes sous-jacents de cette dernière.

Perception visuelle en quête de sens

La vision se trouve largement impliquée dans l'orientation de l'attention (Leigh & Zee, 2015). En effet, la fixation d'un élément visuel permet d'extraire et de traiter les informations qui lui sont associées. Deux théories concurrentes tentent d'expliquer l'orientation de l'attention lors de la perception d'une scène visuelle. D'une part, la théorie de la saillance visuelle (*image guidance*) stipule que la perception d'une scène est principalement guidée par la proéminence des propriétés d'une image ou d'une scène, telles que les couleurs ou les contrastes. D'autre part, la théorie de l'orientation cognitive (*cognitive guidance*) suggère que la perception d'une scène est plutôt dirigée par les régions visuelles porteuses de sens et par leur niveau d'information (p. ex. un fruit posé sur une table comparé au pied de la table)

Afin de tester ces théories, J. M. Henderson et Hayes (2017) et J. M. Henderson et al. (2018) ont développé une méthode permettant de créer des cartes sémantiques à partir de scores attribués

à différents extraits de l'image. L'équipe a ensuite comparé une carte de chaleur d'« attention visuelle » représentant les points de fixation de participants, aux cartes de sémantique et de saillance. La figure 13, issue de J. M. Henderson et al. (2018), représente (de gauche à droite) un exemple de scène présentée aux participants avec la carte de fixation visuelle, la carte de sens ainsi que la carte de saillance lui correspondant. Une analyse de correspondance d'histogrammes indique que la carte sémantique explique mieux la variance de la distribution des points de fixation que celle de la saillance. Cette découverte renforce la théorie de la perception visuelle guidée par le sens. Notons cependant que les éléments porteurs de sens peuvent tout à fait chevaucher les éléments proéminents (Peacock et al., 2023). La création de cartes sémantiques et saillantes s'avère également prometteuse pour étudier les processus attentionnels de personnes présentant un CRCI.

Figure 13. Exemple de scène présentée aux participants (gauche) et les différentes cartes de chaleur lui correspondant : carte de fixation visuelle (centre gauche), carte sémantique (centre droit) et carte de saillance (droite).

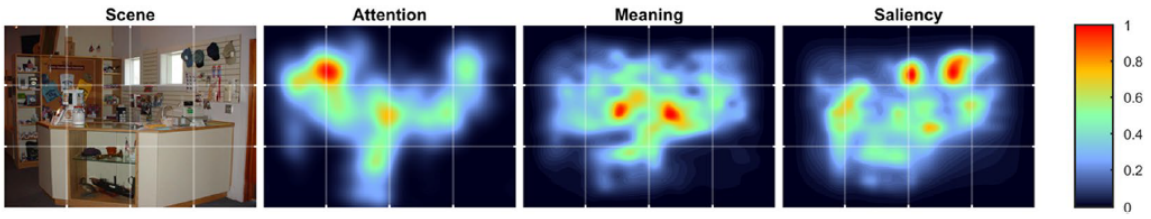


Figure 1. An example scene with its attention map empirically determined from viewer fixations, along with its meaning and saliency maps. (Grid lines are included for ease of comparison and were not present in the scenes).

Note : Reproduit, avec permission, de J. M. Henderson et al. (2018). 2018 Springer Nature.

MESSAGES CLÉS — INVESTIGUER LES PROCESSUS LINGUISTIQUES PAR LE REGARD

La parole guide le regard

- La vision est un système sensoriel largement impliqué dans l'orientation de l'attention.
- Le comportement oculaire se décompose en deux types d'événements principaux : les saccades, mouvements rapides qui dirigent la fovéa vers une cible, et les fixations, succession de microsaccades sur une même cible dans un laps de temps commun.
- L'étude du comportement oculaire au moyen de techniques d'oculométrie permet de mieux comprendre les processus cognitifs sous-tendant la recherche lexicale.
- Lors de la perception d'une scène visuelle, le regard serait plutôt guidé par la recherche d'éléments de sens (p. ex., des objets, des animés, etc.) que des éléments saillants (p. ex., une couleur vive).

5 Manifestations du CRCI à travers la parole et le regard

Un nombre croissant d'études mettent en évidence l'intérêt de l'analyse de la parole en tant qu'outil précis de détection précoce de troubles cognitifs. En effet, les indices linguistiques contenus dans la parole permettraient d'identifier des changements subtils de la cognition, à l'instar de ceux observés dans le cadre du CRCI. Dès lors, comment le CRCI s'exprime-t-il à travers la parole et le comportement oculomoteur ? La réponse à cette question implique de sélectionner les paramètres de parole appropriés. Or, la littérature sur les marqueurs linguistiques de troubles cognitifs est vaste, et les paramètres étudiés demeurent encore mal définis. De fait, l'interprétation de leur efficacité varie sensiblement d'une étude à l'autre, ce qui rend leur fiabilité hasardeuse. Cette limite pourra être contournée par l'aboutissement d'une revue des paramètres actuellement utilisés en recherche. Les résultats de cette revue faciliteront le choix de variables linguistiques les plus susceptibles d'être associée à un CRCI.

En outre, seules deux études se sont penchées sur les aspects linguistiques du CRCI (Myers et al., 2022 ; K. Williams et al., 2021). Ces travaux se sont limités à l'analyse de la complexité syntaxique, un facteur trop étroit pour cerner toute la complexité des altérations cognitives induites par un cancer et ses traitements. La fréquence élevée des plaintes concernant le manque du mot, chez les survivants d'un cancer, souligne le besoin d'appréhender les mécanismes cognitifs sous-jacents à ce phénomène. Cette thèse souhaite contribuer à une compréhension plus fine de ces troubles en apportant les fondements théoriques des caractéristiques prosodiques du CRCI. Ces connaissances serviront au développement d'un outil de diagnostic automatisé basé sur la parole.

Le **premier article** de cette thèse, Investiguer les troubles subtils de la cognition, offre un panorama des paramètres linguistiques utilisés à la détection de troubles cognitifs subtils dans diverses étiologies et syndromes. Les résultats de cette synthèse permettront de présélectionner les potentiels marqueurs de CRCI. Le **deuxième article**, Dévoiler les marqueurs prosodiques du CRCI, vise à déterminer le pouvoir prédictif de ces marqueurs au moyen d'un modèle d'apprentissage automatique (*machine learning*). Cet article expérimental compare des échantillons de parole de survivantes d'un cancer du sein ayant une plainte cognitive, de survivantes sans plainte cognitive et de sujets contrôles. Enfin, le **dernier article**, Comprendre l'anomie en contexte de CRCI grâce à l'oculométrie, examine les processus cognitifs impliqués dans la recherche lexicale de survivantes d'un cancer du sein grâce à une technique d'oculométrie.

Méthodologie

Ce chapitre présente la méthode élaborée spécifiquement pour cette thèse. Il détaille les étapes d'analyse de la parole, décrit l'analyse du comportement oculomoteur, et justifie le choix des outils statistiques et d'apprentissage machine.

1 Démarches éthiques

Les données analysées dans ce projet de thèse proviennent de trois études plus larges. D'une part l'étude ACTICOG, ayant pour objectif d'évaluer l'efficacité d'un programme de réhabilitation physique proposé à des patientes ayant terminé la phase curative de leurs traitements oncologiques, et l'étude REMCOG, qui évalue l'efficacité d'un programme de ré-entraînement cognitif à destination de patients ayant des séquelles cognitives des suites d'un AVC ou d'un cancer. Ces deux programmes sont proposés à l'hôpital Henry Gabrielle (Saint-Genis-Laval, France) et ont été validés par le comité d'éthique des Hospices Civils de Lyon aux numéros d'enregistrement n°22-5009 et n°22-121. D'autre part, l'étude Discours, Cancer et Cognition (DisCCo) qui a été conçue dans le cadre de cette thèse, et porte sur les troubles cognitifs induit par un cancer du sein et ses traitements. Cette étude a reçu la validation du Comité de Protection des Personnes Est-III en août 2023. Cette étude est enregistrée au numéro national 2023-A00267-38.

2 Population

Cette section présente les conditions d'inclusion des participantes à nos analyses. Soixante-neuf volontaires ont accepté de prendre part à cette recherche, soit cinquante-cinq survivantes d'un cancer et quatorze sujets témoins. Les volontaires de l'étude DisCCo ont reçu une indemnité pour leur participation.

2.1 Conditions d'éligibilité

Critères d'inclusion

Pour être éligibles à cette recherche, toutes les volontaires devaient (i) être une femme entre 30 et 65 ans, (ii) avoir le français pour langue maternelle ou avoir appris le français avant l'âge de 6 ans, (iii) avoir une vue normale ou corrigée, (iv) avoir une audition normale ou corrigée, (v) être affiliée à un régime de sécurité sociale ou assimilée. En plus des critères sus-cités, les survivantes d'un cancer devaient (vi) avoir terminé la phase de traitements curatifs (chimiothérapie, immunothérapie, thérapie ciblée, radiothérapie, chirurgie à visée curative [c'est-à-dire mastectomie partielle ou complète ou curage ganglionnaire]) depuis plus de quatre semaines et moins d'un an. Les contrôles de routine et l'hormonothérapie pouvaient être en cours.

Critères de non-inclusion

N'ont pas été analysées les données des participantes qui présentaient (i) des antécédents de troubles neuro-développementaux du langage oral (dysphasie, bégaiement), (ii) des antécédents psychiatriques (dépression, schizophrénie, éthylisme/cannabis/drogue), (iii) un traitement médicamenteux susceptible d'entraîner des modifications cognitives (anxiolytique, antidépresseurs), (iv) des antécédents neurologiques (AVC, traumatisme crânien, épilepsie), (v) des antécédents familiaux de pathologies neurodégénératives (maladie d'Alzheimer, maladie de Parkinson), (vi) des antécédents de cancers des voies aéro-digestives supérieures susceptibles d'entraîner une modification de la parole, et (vii) la présence d'une dyslexie. En plus des critères précédents, les survivantes d'un cancer devaient ne pas présenter de métastases.

2.2 Cohorte

Au total, cinquante-sept participantes ont été éligibles à cette recherche ($N_{ACTICOG}$ = 29, N_{REMCOG} = 6, N_{DisCCo} = 9 survivantes d'un cancer et 13 contrôles). Douze participantes n'ont pas été incluses dans les analyses pour antécédents neurologiques (épilepsie, Accident Ischémique Transitoire, aphasie post-traumatique, autres cancers, Sclérose En Plaques), psychiatriques (dépression), présence de métastases, présence de dyslexie et langue française non-fluente.

3 Méthode d'analyse de la parole

Cette section présente la méthode d'analyse de la parole développée pour cette thèse. Cette méthodologie a fait l'objet de deux articles, dont une version française simplifiée à destination

d'orthophonistes et de chercheurs en Sciences du Langage publiée dans le numéro 142 de la revue *Glossa* parue le 28 décembre 2024 (Cf. Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment).

Référence : Richard, A. B., Hirsch, F., & Jacquin-Courtois, S. (2024). Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment. *Glossa, UNADREO,* 142, 49–64.

Une version plus détaillée à destination d'un public international a été soumise à une revue à comité de lecture (Cf. Assessing Cancer-Related Cognitive Impairment for breast cancer survivors with speech analysis).

3.1 Tâches de narration

L'ensemble des participantes avait pour consigne de raconter une histoire à partir d'une séquence de cinq vignettes issue de la batterie GréMots (voir figure 1) (Bézy et al., 2016) sans limite de temps, puis d'en inventer la fin. Les vignettes retracent l'histoire d'une famille qui arrive à la plage. Le père fait de la planche à voile pendant que son fils et sa femme restent sur le rivage. Soudain, le vent se lève et le père tombe à l'eau. La dernière vignette dépeint la mère et le fils alarmés, la mère appelant les secours au téléphone. Les participantes ont eu pour consigne de raconter l'histoire qu'elles voient, puis d'en inventer une fin. Pour les participantes des études ACTICOG et REMCOG, la tâche était présentée sur format papier. Pour les participantes de l'étude DisCCo, la tâche était présentée sur écran d'ordinateur.

Figure 1. Images présentées lors de la tâche narrative.



En plus de cette tâche, les participantes des études ACTICOG et REMCOG ont passé deux autres

tâches discursives, à savoir le rappel immédiat d'un court texte de cinq paragraphes (« Michel et son puits ») issu du Protocole Montréal d'Evaluation de la Communication (Joanette et al., 2004) lu par l'examinatrice et une conversation thématique. Pour la première tâche, les instructions étaient « Je vais vous lire un texte de cinq courts paragraphes. Vous allez le mémoriser et quand j'aurai terminé, ce sera à vous de me raconter l'histoire de mémoire ». Les instructions de la tâche conversationnelle étaient « J'aimerais que vous me parliez de vos prochaines vacances, envisagées ou fictives ». Les seules interventions de l'examinatrice visaient à encourager les participantes à continuer leur discours.

Les participantes de l'étude DisCCo ont, quant à elles, lu et mémorisé deux très courts textes issus de la batterie ("Wechsler Memory Scale–Fourth Edition", 2012) d'évaluation de la mémoire. Les textes étaient affichés sur écran pendant quarante secondes pour permettre plusieurs lectures. À l'issue du décompte, il a été demandé aux participantes de rappeler l'histoire immédiatement, puis de la raconter une seconde fois après une tâche d'interférence de quarante secondes ne faisant pas intervenir de langage.

Un échantillon de ces productions discursives a été analysé lors d'études pilotes (voir Annexes Assessing Cancer-Related Cognitive Impairment for breast cancer survivors with speech analysis). Les résultats ont montré le pouvoir discriminant de la tâche narrative comportant les vignettes par rapport aux deux autres tâches discursives. Pour cette raison, seules les données de cette tâche ont été analysées dans cette thèse.

3.2 Matériel et conditions d'enregistrement

Les productions discursives des participantes des études ACTICOG et REMCOG ont été enregistrées dans une pièce calme de l'hôpital de rééducation Henry Gabrielle (Hospices Civils de Lyon, Saint-Genis-Laval, France) avec un micro-cravate omnidirectionnel Rode Lavalier Go® branché à un ordinateur portable. Quant aux participantes de l'étude DisCCo, les enregistrements ont eu lieu dans une cabine anéchoïque du Centre de Recherche en Neurosciences de Lyon (Bron, France) au moyen d'un micro-cravate omnidirectionnel Shure® relié à une interface audio M-Audio®, elle-même connectée à une unité centrale. Dans les deux cas, les microphones étaient placés à 15 cm en dessous de la bouche des participantes.

3.3 Paramètres linguistiques

Le tableau 1 décrit les différents paramètres prosodiques que nous avons analysés ainsi que leur mesure. L'analyse de la parole offre un large éventail d'options concernant les paramètres qu'il est possible d'étudier. Afin de restreindre ce choix, des études pilotes ont été réalisées pour présélectionner les paramètres les plus susceptibles de renseigner sur le CRCI. Cette méthode a permis de sélectionner sept paramètres prosodiques (c.-à-d. pauses silencieuses, pauses remplies,

allongements vocaliques, fréquence fondamentale, débit de parole, débit articulatoire et ratio de parole sur silence) ainsi que différentes mesures pour les pauses et allongements vocaliques (p. ex., durée moyenne, durée médiane).

Table 1. Paramètres prosodiques.

Nom du paramètre	**Caractéristiques**	**Mesure**
Pauses silencieuses	Segments caractérisés par l'absence de parole, de toux ou de bâillement d'une durée supérieure ou égale à 200 ms	Durée (ms) moyenne et durée médiane de la durée (ms) totale des pauses Ratio : durée (ms) totale des pauses / durée totale du discours (ms)
Pauses remplies	Segments voisés de type « euh », « hum » d'une durée supérieure ou égale à 200 ms ayant une courbe intonative (F0) plate	Durée (ms) moyenne et durée médiane de la durée (ms) totale des pauses Ratio : durée (ms) totale des pauses / durée totale du discours (ms)
Allongements vocaliques	Tenue d'un phonème vocalique d'une durée supérieure ou égale à 200 ms dont la courbe intonative (F0) est plate	Durée (ms) moyenne et durée médiane de la durée (ms) totale des allongements vocaliques Ratio : durée (ms) totale des allongements vocaliques / durée totale du discours (ms)
F0 SD	Ecart-type de la fréquence fondamentale (Hz)	Mesuré sur l'ensemble de l'extrait audio en conservant les paramètres d'origine de Praat
Débit de parole	Vitesse de parole	Nombre de syllabes/durée totale du discours
Débit d'articulation	Vitesse d'articulation	(Nombre de syllabes/durée totale du discours) - durée totale des pauses
Ratio parole / silence	Proportion de silence au sein du discours	Durée totale du discours / durée totale des pauses silencieuses

Note. Durée totale (ms) du discours : mesurée à parole de l'émission du premier phonème à la fin du dernier phonème. Durée totale des pauses : somme de la durée (ms) des pauses silencieuses, remplies et des allongements vocaliques.

3.4 Analyse des pauses

La section suivante décrit les différentes étapes de l'analyse prosodique réalisées à partir du signal acoustique.

Pré-traitement du signal acoustique

Avant de commencer les analyses prosodiques, les productions narratives de chaque participante ont été extraites des enregistrements audio au moyen du logiciel libre Audacity (Team, 1999). Les parties qui n'ont pas été étudiées (p. ex., intervention de l'examinatrice) ont été enlevées.

Transcription

La transcription des productions a été réalisée automatiquement avec le modèle large V3 Whisper (Radford et al., 2022). Whisper est un grand modèle de langage (*Large Language Model*) code source ouvert (*open source*) dédié à la reconnaissance automatique de la parole. Une de ses fonctionnalités consiste à transcrire de la parole en texte rédigé à partir du signal acoustique.

Chaque transcription a été vérifiée afin de corriger les éventuelles erreurs commises par le modèle. Les erreurs, peu nombreuses, ont surtout concerné les noms propres et les troncatures de mots des enregistrements dont la qualité était altérée (p. ex., bruits parasites). Les pauses silencieuses (#) et les pauses remplies (euh) ont ensuite été ajoutées manuellement. Il est à noter que Whisper élimine par défaut toutes les disfluences (p. ex., pauses, répétitions, faux-départs, etc.). Enfin, les textes ont été nettoyés en supprimant les marques de ponctuation ou en remplaçant les chiffres par leur équivalent en lettres. Le tableau 2 montre des exemples d'erreurs de transcription selon la qualité audio des enregistrements.

Table 2. Exemples d'erreurs de transcription automatique.

Erreurs de transcription	**Correction**	**Nature de l'erreur**
Enregistrements de bonne qualité		
[les vagues] qui **onduent**	[les vagues] qui **ondulent**	*Phonologie* : le dernier phonème est manquant /ɔ̃dy/ et /ɔ̃dyl/
son scooter des **mères**	son scooter des **mers**	*Sémantique* : homophones « mère » et « mer »

(A suivre)

(Suite)

Erreurs de transcription	**Correction**	**Nature de l'erreur**
il prend quelque chose dans la **glaciaire**	il prend quelque chose dans la **glacière**	*Sémantique* : homophones « glaciaire » et « glacière »
un monsieur **où** papa	un monsieur **ou** papa	*Syntaxe* : homophones « où » et « ou »
la planche à voile du papa qui du coup est tomb**ée** dans l'eau	la planche à voile du papa qui du coup est tomb**é** dans l'eau	*Morphosyntaxe* : la vignette dépeint un homme tombant de sa planche à voile, or Whisper accorde « tomber » à la « planche » au lieu de l'accorder avec « papa ».
Enregistrements de qualité moyenne		
une planche de **sœur**	une planche de sur**f**	*Phonologie* : le dernier phonème est manquant, /sœr/ et /sœrf/
pas loin des coco**tis**	pas loin des coco**tiers**	*Phonologie* : le dernier phonème est manquant /kokoti/ et /kokotje/
sans que ni la maman ni le petit garçon ne remarqu**e**	sans que ni la maman ni le petit garçon ne remarqu**ent**	*Morphosyntaxe* : absence de la marque du pluriel au verbe « remarquer » car Whisper ne rend pas compte du lien entre « maman » et « garçon » introduit par « ni ... ni »

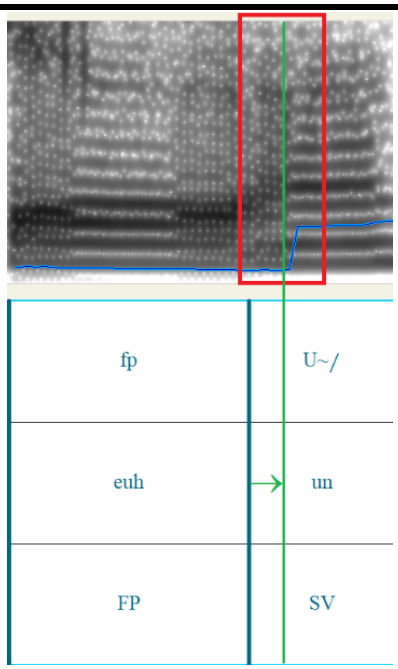
Note. Enregistrements de bonne qualité : échantillons de parole enregistrés dans une cabine insonorisée. Enregistrements de qualité moyenne : échantillons de parole enregistrés dans une pièce calme d'un hôpital. Présence possible d'écho, de bruits de fond, de voix étouffées provenant du couloir, rares bruits extérieurs (p. ex., voiture, hélicoptère).

Annotation des pauses

Les pauses ont été automatiquement annotées en utilisant le logiciel SPPAS (Bigi, 2015). SPPAS est un logiciel *open source* spécifiquement développé pour l'analyse de la parole. Il fournit de nombreuses analyses automatiques telles que la détection de pauses silencieuses, la conversion graphème-phonème ou encore la syllabification de la transcription alignée sur le signal acoustique. Son interface graphique et son caractère interopérable en font un outil autant applicable en recherche qu'en clinique. En complément, le logiciel libre Praat (Boersma & Weenink, 2024) a été utilisé pour affiner les analyse. Celui-ci a été conçu pour l'analyse du signal acoustique et permet de visualiser avec précision le signal pour en extraire de nombreuses variables prosodiques (p. ex., fréquence fondamentale).

Les enregistrements (format `.wav`) et les transcriptions (format `.txt`) ont d'abord été chargés dans SPPAS. Ensuite, les unités inter-pausales (*Inter Pausal Units*, IPUs) ont été recherchées depuis l'onglet « Annotation », puis la complétion automatique de la transcription a été lancé. La première fonction, « Search for IPUs », détecte automatiquement les pauses silencieuses, et la seconde, « Fill in IPUs », complète les espaces vides par la transcription des segments de parole. Ces fonctions sont accessibles via le bouton « STANDALONE ». Le résultat de ces requêtes a été généré au format de sortie « `.TextGrid` », format traité par Praat. Les fichiers `.TextGrid` ont été manuellement vérifiés dans Praat et les frontières des pauses silencieuses ont été réajustées quand celles-ci ne correspondaient pas au signal acoustique. La figure 2 montre une erreur d'alignement en le signal acoustique et sa transcription phonémique. Les erreurs d'alignement étaient nombreuses pour les enregistrements de qualité moyenne et rares pour les enregistrements de bonne qualité.

Figure 2. Exemple d'erreur d'alignement. La pause « fp » est partiellement tronquée et la frontière (barre bleue) doit être tirée vers la droite jusqu'à la fin de l'émission du phonème (barre verte).



Les fonctions « Text Normalization », « Phonetization » et « Alignment » simultanément depuis l'onglet « Annotation » de SPPAS. Cette étape a permis de segmenter l'audio en phonèmes à partir du texte, et d'aligner temporellement les phonèmes sur le signal acoustique. De cette manière, les allongements vocaliques et les pauses remplies sont facilement repérables par leur durée égale ou supérieure à un seuil temporel défini (ici 200 ms). Les fichiers `.TextGrid` nommés `-palign` ont été inspecté dans Praat pour vérifier l'alignement des phonèmes sur le signal acoustique. Les fichiers `.TextGrid` intitulés `-palign` ont été importés dans l'onglet « Analyze » de SPPAS pour sélectionner la ligne d'annotation (c.-à-d. *tier*) « TokenAlign ». Un filtre (c.-à-d. *single filter*) a ensuite été appliqué pour créer un nouveau *tier* nommé « pauses ». Ce *tier* contient les phonèmes vocaliques et les pauses (c.-à-d. selon les conventions de SPPAS U /, A/, O/, @, i, E, e, j, H, a , O , y, 2, 9, u, fp, #) d'une durée égale ou supérieure à 200 ms. Seules les annotations des allongements vocaliques dont la courbe

intonative plate ont été conservées. Enfin, l'ensemble des annotations a été harmonisé en remplaçant celles déjà existantes du *tier* « pauses » par les étiquettes suivantes : SP pour les pauses silencieuses (*Silent Pauses*), FP pour les pauses remplies (*Filled Pauses*) et SV pour les allongements vocaliques (*Sustained Vowels*) grâce au script de remplacement de texte développé par DiCanio (2011).

Accord inter-annotateur

L'accord inter-annotateur consiste à donner un certain pourcentage des données (environ 20%) ainsi que son protocole d'annotation (c.-à-d. un descriptif des éléments à annoter) à une ou plusieurs personnes n'ayant pas de connaissances du corpus afin de vérifier la validité et la fiabilité du processus d'annotation. Il se calcule à l'aide du coefficient $\varkappa$ de Cohen, dont le score s'échelonne de 0 à 1. Un score égal ou supérieur à .81 atteste la validité et la fiabilité des annotations. En clinique, cette étape permet d'assurer une annotation comportant le moins de biais de subjectivité possible. En effet, la connaissance du patient et de ses troubles peut conduire à surestimer ou sous-estimer le nombre de pauses. Il est vrai cependant que l'accord inter-annotateur peut être difficile à réaliser en clinique (p. ex., manque de disponibilité de ses collègues), aussi cette étape peut être optionnelle. 20% des fichiers `.TextGrid` ont été confiés à deux annotatrices cliniciennes non-orthophonistes naïves de l'analyse de la parole pour comparer les annotations des pauses. Un document détaillé du système d'annotation leur a été fourni en amont. L'accord inter-annotateur a atteint 98,71% d'accord pour l'annotation des pauses silencieuses, soit un agrément presque parfait ($\varkappa$ = .97), 90,98% pour l'annotation des pauses remplies, soit un agrément presque parfait ($\varkappa$ = .81), et 86,88% pour l'annotation des allongements vocaliques, soit un agrément substantiel ($\varkappa$ = .74).

Extraction des données prosodiques

Le nombre, la durée totale, moyenne, médiane et l'écart-type des trois types de pauses ont été extraits de chaque échantillon audio à partir de SPPAS. Après sélection de tous les tiers « pauses » dans l'onglet « Analyze », l'option « Statistics » a permis d'accéder aux données statistiques des pauses sous forme de tableur.

4 Étude du comportement oculomoteur

Outre leur production narrative, le comportement oculomoteur des participantes de l'étude DisCCo (N = 22) a été enregistré au moyen d'une technique de suivi oculaire (*eye-tracking*).

4.1 Matériel d'oculométrie

L'expérience a été réalisée avec une barre d'*eye-tracking* Tobii Pro Fusion. Cette technologie dotée de caméras infrarouge offre plusieurs avantages quant à l'étude de la parole. En effet, ce système ne requiert aucune contrainte physique, ce qui permet une liberté de mouvement de la tête et rend la production de la parole plus proche des conditions naturelles. L'absence de mentonnière constitue également un critère essentiel pour une population oncologique pouvant facilement être distraite par des stimulations sensorielles et pouvant avoir été exposée à des situations traumatisantes, telles que la répétition d'examens d'imagerie médicale lourde.

Configuration

La figure 3 illustre la configuration de l'expérience. La barre d'*eye-tracking* de résolution 120 Hz était montée à 7 cm en dessous d'un écran haute définition de dimension 487 x 289 mm de résolution 1920 x 1080 pixels à 60 Hz. Les participantes étaient assises à environ 70 cm de l'écran. L'enregistrement des données oculomotrices a entièrement été effectué par le logiciel Tobii Pro Lab (Tobii AB, 2023) installé sur un ordinateur fixe de marque Dell. L'ordinateur était équipé d'un processeur Intel i7 et doté du système d'exploitation Windows 10.

Figure 3. Configuration de l'expérience de l'étude DisCCo. De gauche à droite : Encadré bleu : enregistreur annexe. Cercle vert : caméra de surveillance. Cercle rouge : micro-cravate. Encadré orange : *eye-tracker*. Encadré violet : écran. Encadré cyan : interface audio.



4.2 Tâche narrative sur écran

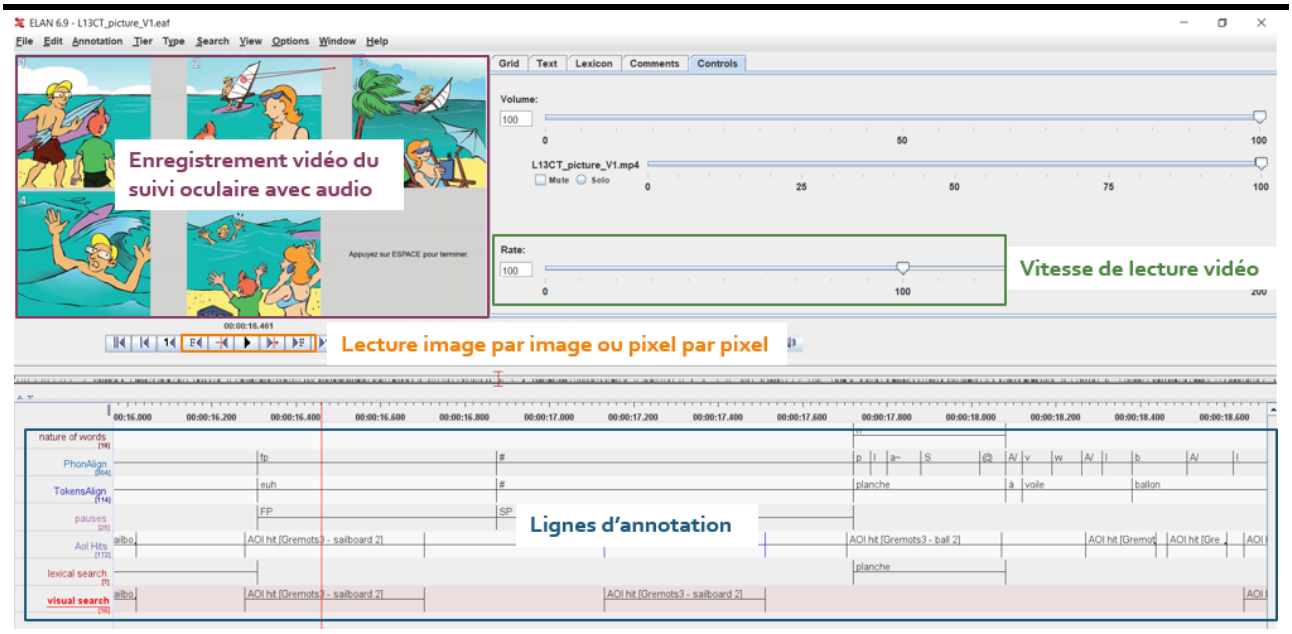
L'expérience s'est déroulée dans une cabine insonorisée du Centre de Recherche en Neurosciences de Lyon. Elles ont eu pour consigne de raconter l'histoire imagée s'affichant à l'écran. La tâche était précédée d'une croix de fixation de 250 ms au centre de l'écran et n'était pas soumise à une limite de temps.

4.3 Extraction et analyse des données oculomotrices

Eye-Voice Span

Des régions d'intérêt (*Areas of Interest*, AoI) ont été définies pour délimiter les éléments visuels porteurs de sens tels que « casquette », « visage d'enfant » ou « parasol » (par contraste avec un trait ou une couleur vive). Les données de suivi oculaire ont ensuite été converties au format `.TextGrid` et temporellement alignées sur les données vocales à l'aide d'un script Python. De cette manière, il a été possible de repérer les AoI correspondant aux items lexicaux de classe ouverte précédés d'une pause. La figure 4 présente l'alignement temporel entre les mouvements oculaires et la parole sur le logiciel ELAN (Max Planck Institute for Psycholinguistics, 2024).

Figure 4. Alignement temporel de la parole et des mouvements oculaires dans le logiciel ELAN.



Le délai entre l'apparition de la première fixation sur un élément visuel porteur de sens et le début de la dénomination de son concept, c'est-à-dire l'*Eye-Voice Span* (EVS) a été mesuré avec le logiciel ELAN (Max Planck Institute for Psycholinguistics, 2024). La figure 5 illustre le calcul de la durée de

l'EVS. Les concepts pouvaient référer à différentes dénominations, p. ex., le dessin d'une planche à voile peut être appelé « planche à voile », mais également « activité nautique » ou « surf ». En plus de l'EVS, le nombre de visites de ces éléments a été également été quantifié.

Figure 5. Calcul du délai (vert) entre la première visualisation d'un élément sémantique (orange) et sa dénomination (bleu).

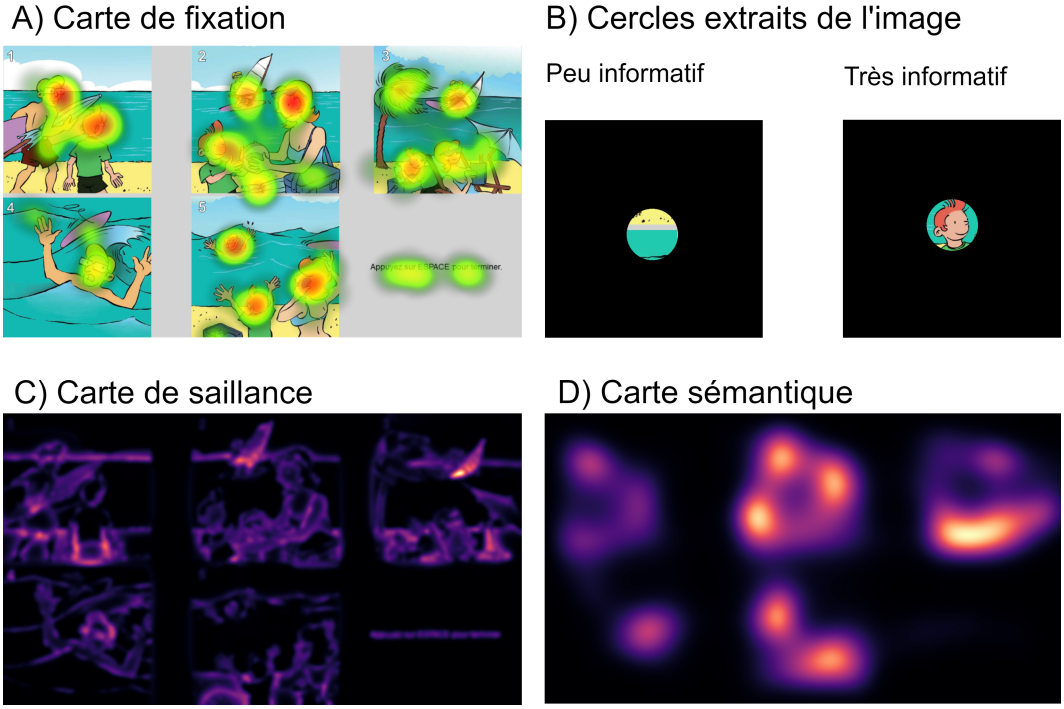


Cartes de fixation

Les données de suivi oculaire ont été extraites et analysées à l'aide de bibliothèques Python pour la visualisation. Tous les scripts ont été développés pour les besoins de cette thèse en utilisant les bibliothèques OpenCV (Bradski, 2000), SciPy (Virtanen et al., 2020), PIL (Umesh, 2012) et scikit-image (Van Der Walt et al., 2014).

Les fixations ont été définies comme une séquence de points sur une zone où la fovéa s'est stabilisée pendant au moins 100 ms. Ce seuil a été considéré comme le temps minimum requis pour extraire et traiter l'information visuelle (Blignaut, 2009). Les cartes de fixation de chaque participante ont été réalisées à partir des coordonnées des points de fixation à l'aide d'un script Python (Cf. figure 6A).

Figure 6. A) Carte de fixation superposée à l'image globale. B) Cercles extraits de l'image globale, à gauche : cercle peu informatif, à droite : cercle très informatif. C) Carte de saillance de l'image globale. D) Carte sémantique de l'image globale.



Une carte de saillance (*saliency map*) a été générée à partir de l'image globale (soit les cinq vignettes affichées en plein écran) utilisée dans la tâche narrative à l'aide d'un script Python (Cf. figure 6C). La détection de saillance d'OpenCV a permis d'extraire les régions proéminentes de l'image globale en s'appuyant sur le contraste entre les pixels adjacents.

Une carte sémantique (*meaning map*) a été créée en suivant la méthode développée par J. M. Henderson et Hayes (2017). Deux cent vingt cercles d'un rayon de 75 px, dont un chevauchement de 50 px entre chaque cercle, ont été extraits de l'image originale (Cf. figure 6B). Les cercles ont été présentés de manière aléatoire à 86 volontaires (62% de femmes, âge médian = 37,5 ans) qui ont été invités à évaluer le niveau d'information de chaque cercle à l'aide d'une échelle de Likert à 6 points (« Très faible », « Faible », « Plutôt faible », « Plutôt élevé », « Élevé », « Très élevé »). Seuls dix volontaires ont entièrement répondu à l'enquête. Cependant, le taux de participation pour chaque cercle était d'environ 20%, de sorte que chacun d'entre eux a été évalué par 15 évaluateurs en moyenne. Ensuite, les scores moyens ont été calculés et pondérés pour chaque cercle selon une logique en base 2 (c'est-à-dire de Très faible = 0, à Très élevé = 32) afin d'augmenter les contrastes entre les cercles peu informatifs et les cercles très informatifs. Enfin, une carte sémantique a été générée à partir de ces notations à l'aide d'un script Python (Cf. figure 6D).

Les cartes de fixation des participants et la carte sémantique, ainsi que les cartes de fixation et la carte de saillance, ont été superposées pour calculer le nombre de pixels correspondant dans chaque paire de cartes. En d'autres termes, la valeur d'intensité d'un pixel dans une carte de fixation a été

soustraite de la valeur du pixel correspondant, en termes de coordonnées spatiales, dans les cartes de saillance et de sémantique. Deux analyses ont été réalisées à l'aide de cette méthode : l'une axée sur le chevauchement spatial et l'autre prenant en compte le temps passé sur chaque zone des cartes.

5 Évaluation cognitive et physiologique, psychologique

L'ensemble des participantes a passé une évaluation cognitive, physiologique et psychologique composée de questionnaires et de tests neuropsychologiques. Cette évaluation permet de contrôler les variables confondantes susceptibles d'influencer la production de parole.

5.1 Évaluation cognitive

Les participantes ont complété la version française du questionnaire *Functional Assessment of Cancer Therapy – Cognitive Function* (FACT-Cog v3 [Joly et al., 2012]). Ce questionnaire autoadministré a spécialement été conçu pour déceler l'éventuelle présence d'une plainte cognitive subjective chez des patients en oncologie. Seule la sous-échelle *Perceived Cognitive Impairment* (PCI) bénéficie d'un seuil pathologique. Un score inférieur ou égal à 55/72 à cette sous-échelle indique la présence d'une plainte cognitive significative (Van Dyk et al., 2020). Ce score a permis de différencier les survivantes d'un cancer ayant une plainte cognitive (BC_{yes}) de celles n'ayant pas de plainte cognitive (BC_{no}).

Les participantes ont passé le test de repérage cognitif *Montreal Cognitive Assessement* (MoCA [Nasreddine et al., 2005]). Ce test rapide à administrer est conçu pour détecter des troubles légers à modérés de la cognition en évaluant plusieurs domaines cognitifs (p. ex., mémoire, capacités visuo-spatiales, langage). Trois items sont consacrés au langage :

- une dénomination de trois animaux,
- une fluence verbale phonémique d'une durée d'une minute,
- une répétition de deux phrases.

Un score inférieur à 26/30 au MoCA indique la présence d'un trouble cognitif.

Les participantes ont passé le test de vitesse de traitement *Symbol Digit Modalities Test* (SMDT [A. Smith, 1973]) en modalité orale. Le test consiste à associer le plus possible de chiffres correspondant à des symboles en 90 secondes en s'aidant d'une clé d'interprétation. Ce test standardisé donne un score brut convertible en écart-type en fonction de l'âge et du niveau d'étude des participants (Strober et al., 2020).

5.2 Évaluation psychologique

Les participantes ont complété l'échelle *Hospital Anxiety Depression Scale* (HADS [Zigmond & Snaith, 1983]). Ce questionnaire comprend deux sous-échelles destinées à évaluer l'anxiété (HAD-A) et la dépression (HAD-D). Un score supérieur à 11/21 à l'une de ces sous-échelles atteste de la présence d'anxiété ou de dépression.

5.3 Évaluation de la fatigue

Les participantes ont complété le questionnaire *Functional Assessment of Chronic Illness Therapy – Fatigue*, (FACIT-F [Bonomi et al., 1996 ; Cella et al., 1993]). Ce questionnaire est spécialement conçu pour évaluer la fatigue globale liée au cancer. Plus le score est élevé, moins la fatigue est importante. À ce jour, il n'existe pas de seuil standardisé pour détecter une fatigue accrue chez des patientes ayant un cancer du sein.

6 Modèle d'apprentissage machine

Le meilleur sous-ensemble de variables de parole permettant d'identifier les participantes avec plainte cognitive de celles sans plainte cognitive a été recherché au moyen de modèles de *machine learning*. Autrement dit, ces analyses ont permis de sélectionner la combinaison de variables de parole discriminant le mieux les groupes.

L'analyse de *machine learning* a consisté en un processus de deux étapes comprenant le classement et la sélection des variables de parole, ainsi que la prédiction des résultats. Un évaluateur de caractéristiques 1R (Holte, 1993) a été utilisé pour le classement de caractéristiques. Celui-ci évalue individuellement le pouvoir prédictif de chaque variable en définissant un ensemble de règles pour les valeurs discrétisées des variables (c'est-à-dire les variables continues transformées en variables discrètes), afin de classer au mieux les échantillons en fonction du résultat. Une liste de variables classées est obtenue à l'issue de cette première sous-étape. Un processus itératif a ensuite été utilisé pour identifier le sous-ensemble optimal de variables de la liste classée. Ce processus consistait à former un premier classificateur en utilisant toutes les variables, puis à éliminer progressivement les variables les moins performantes des données d'entraînement, jusqu'à ce que seules deux variables restent. Le sous-ensemble de caractéristiques optimal est celui conduisant au score f1 le plus élevé, ce qui représente la moyenne harmonique entre la précision et le rappel. Enfin, un classificateur *naïve Bayes* a été utilisé (John & Langley, 2013) pour procéder à la prédiction des résultats. Ces classificateurs postulent que les variables sont indépendantes les unes des autres. Il utilise les probabilités de classe et les probabilités conditionnelles de la valeur des variables en fonction de la classe calculée à partir des données entraînées, afin de déterminer la classe des échantillons pour la

phase de test. Cet algorithme s'est avéré très efficace pour divers ensembles de données, et est connu pour obtenir de bons résultats sur de petits ensembles de données, même dans les cas où l'hypothèse d'indépendance des variables est rejetée (Zhang, 2004). Les modèles ont été évalués grâce à une étape validation croisée (*cross-validation*) répétée dix fois, qui consiste à utiliser itérativement 90% des données pour l'apprentissage du modèle et à conserver 10% des données pour les tests, ce jusqu'à ce que tous les échantillons soient utilisés. Le classement des variables et l'entraînement du modèle ont été effectués de manière imbriquée afin d'éviter qu'une même variable ne soit employée deux fois (Lo Vercio et al., 2020). Les analyses de *machine learning* ont été réalisées à l'aide de la version 3.8 de Weka (Witten et al., 2011).

7 Analyses statistiques

7.1 Variables démographiques et cofacteurs

Les analyses statistiques de cette thèse ont été effectuées à l'aide du langage R (R Core Team, 2021), au sein de l'environnement logiciel RStudio (RStudio, BPC, 2020). Ce logiciel est gratuit, hors-ligne et son code source est accessible. Compte-tenu du nombre de participantes, des tests non-paramétriques ont été sélectionnés pour analyser l'ensemble des données. Des U de Mann-Whitney (Mann & Whitney, 1947) ont été employés pour comparer le groupe de survivantes et le groupe de contrôles. Lorsque la cohorte était divisée en trois groupes, des tests de Kruskal-Wallis (Kruskal & Wallis, 1952) ont comparé les variables des survivantes d'un cancer ayant une plainte cognitive (BC_{yes}) à celles des survivantes d'un cancer sans plainte cognitive BC_{no} et des sujets témoins (HC). Ces tests ont été complétés par des analyses post-hoc de Dunn (Dunn, 1964) pour comparaisons multiples, afin d'identifier les groupes se différenciant significativement les uns des autres. Enfin, une corrélation de Spearman (Spearman, 1904) a permis de tester la possibilité d'un effet des variables démographiques et psychologiques (c.-à-d. âge, anxiété, dépression, fatigue) sur les pauses.

7.2 Variables de parole

Les variables de parole ont fait l'objet d'une double analyse statistique. Des modèles linéaires mixtes ont d'abord été élaborés afin de mieux comprendre la relation entre les variables de parole et le CRCI, tout en tenant compte de l'influence de cofacteurs tels que la dépression ou la fatigue. Des analyses bayésiennes ont ensuite été réalisées pour confirmer ou infirmer ces résultats.

Modèles linéaires mixtes

L'hypothèse nulle (H0) postule qu'il n'y a pas de différence dans la variabilité des variables de parole en fonction des groupes et des conditions (c.-à-d. des tâches). Pour tester H0 par rapport à H1 (c.-à-d. la présence d'une différence significative), des modèles mixtes linéaires ont été conçus et comparés pour chacune des variables de parole dépendantes. Un intercept aléatoire a été attribué à la variable participante, ainsi que des effets fixes ont été attribués aux groupes (niveaux BC_{yes}, BC_{no} et HC), à la condition (niveaux narration sur images et narration sur mémoire), aux co-facteurs (c.-à-d. âge, anxiété, dépression). Les différents termes d'interaction ont été combinés avec tous ces effets fixes. Les modèles ont été réalisés sous R (R Core Team, 2021) dans RStudio (RStudio, BPC, 2020) à l'aide du package lme4 (Bates et al., 2015). La formule ci-dessous montre un exemple de modèle linéaire mixte que nous avons utilisé.

```
lmer(speech_rate   +group +cond +age +anxiety +depression +(1|id), data=data)
```

Statistiques Bayésiennes

Une analyse Bayésienne a été effectuée pour confirmer les résultats des modèles mixtes. Étant donné qu'aucun *a priori* robuste quant aux valeurs des paramètres des modèles n'a été déterminer, des *a priori* non informatifs ont été définis pour tous les paramètres de tous les modèles avec une distribution *a priori* normale N(0,3). Le facteur de Bayes a été calculé entre chaque comparaison de modèle pour les tests d'hypothèse et, pour le modèle le plus probable, le résumé de l'estimation des paramètres avec un intervalle de crédibilité de 95% a été pris en compte. L'analyse bayésienne a été réalisée sous R à l'aide du package brms (Bürkner, 2021).

Investiguer les troubles subtils de la cognition

1 Contribution de l'article à la thèse

Le premier article de cette thèse passe en revue les études évaluant les troubles subtils de la cognition au moyen de variables de parole continue. L'objectif est non seulement de recenser les paramètres linguistiques en vigueur, mais également d'appréhender les tâches d'élicitation du discours et les méthodes d'analyse de la parole. À l'issue de cette étude, sera effectuée une présélection des marqueurs de parole potentiels du CRCI, laquelle servira à la recherche suivante.

2 Résumé de l'article

Introduction

L'appellation Troubles NeuroCognitifs (TNC) couvre un large éventail d'étiologies caractérisées par un déclin des performances cognitives, lesquelles sont constatées par l'individu ou ses proches (American Psychiatric Association, 2022). Les TNC sont considérés comme mineurs (TNCm) lorsque l'autonomie de la personne est préservée. Ces troubles comprennent autant les pathologies dégénératives et vasculaires que les syndromes cognitifs induits par un contexte infectieux, toxique ou oncologique et ceux dont l'étiologie sous-jacente n'est pas diagnostiquée.

Les TNCm, desquels le *Cancer-Related Cognitive Impairment* fait partie, altèrent la qualité de vie d'un individu. Malgré la variété des étiologies et syndromes, les personnes atteintes de TNCm décrivent des symptômes d'apparence similaire tels que des problèmes d'ordre mnésique, exécutif, attentionnel et linguistique. La particularité des TNCm réside en ce que l'évaluation neuropsychologique ne correspond pas à l'intensité des difficultés rapportées par les patients. Une des raisons couramment avancée est que les tests sont conçus pour repérer des troubles plus sévères dans un environnement contrôlé, or les TNCm s'expriment lors de situations quotidiennes par des manifestations discrètes. En l'absence de résultats pathologiques clairs, les patients atteints de TNCm restent souvent sous-diagnostiqués et vivent avec un handicap invisible alors même que des options thérapeutiques pourraient leur être proposées. De nouvelles approches sont donc nécessaires pour détecter ces changements subtils de la cognition.

À ce titre, l'analyse de la parole fait preuve d'un intérêt exponentiel en tant qu'outil de surveillance des TNCm. Le caractère non invasif de la parole, sa fine sensibilité et sa validité écologique sont autant d'atouts qui expliquent sa popularité croissante. En raison de son interaction avec d'autres fonctions cognitives, le langage verbal est susceptible de contenir des marqueurs fiables de changements cognitifs. Toutefois, il n'existe pas de consensus sur les paramètres linguistiques permettant d'identifier le mieux les personnes atteintes d'un TNCm. De là découle une importante quantité de variables diverses, dont la terminologie souvent imprécise entrave la reproductibilité de ces études. Une classification des paramètres bien décrite et facile à mettre en œuvre semble donc requise.

2.1 Objectifs

L'objectif de cette revue est d'offrir une vue d'ensemble des marqueurs actuellement utilisés pour détecter ou étudier les troubles cognitifs subtils indépendamment de l'étiologie sous-jacente ou du syndrome cognitivo-comportemental. Dans un second temps, cet article fournit une synthèse actuelle des méthodes d'analyse de la parole à des fins cliniques.

2.2 Méthode

Après consultation de cinq bases de données bibliographiques électroniques, cinquante-et-unes études publiées entre 2012 et juillet 2023 sur 815 articles identifiés ont été incluses à cette revue. Pour être éligibles, les études devaient comporter un groupe de patients ayant un TNCm et un groupe contrôle d'au moins dix participants chacun et des tâches de parole continue (hors lecture). N'ont pas été retenus les articles comprenant des personnes souffrant de troubles moteurs de la parole, des personnes âgées de moins de 18 ans ou de plus de 75 ans et des participants présentant des troubles cognitifs modérés à sévères (c'est-à-dire des scores inférieurs à 21/30 pour le *Mini-Mental State Examination* [MMSE] ou 22/30 pour le *Montreal Cognitive Assessment* [MoCA]).

2.3 Principaux résultats

Les résultats de cette étude ont montré que la plus grande partie des articles examinés portaient sur des patients ayant une maladie d'Alzheimer à un stade précoce et que les locuteurs étaient pour la majorité anglophones.

Au total, 384 paramètres linguistiques ont été identifiés, lesquels étaient étiquetés par 335 noms différents. La plupart des variables et leur mesure dépendaient des études, bien qu'elles étaient analogues d'une étude à l'autre. Autrement dit, deux articles pouvaient s'intéresser au débit de parole, l'un en calculant le taux de mots par minute (appelé *word per minute rate*) et l'autre en mesurant le

nombre de syllabes par seconde (appelé *speech rate*). Les domaines linguistiques les plus explorés ont concerné la phonétique-prosodie et la sémantique-lexicale, dont les paramètres récurrents étaient, respectivement, les pauses et le débit de parole, et les items lexicaux.

En ce qui concerne les tâches de parole continue, 12 études ont inclus plus d'une tâche, et 27 études ont étiqueté le type de discours par une appellation erronée. Par exemple, plusieurs d'entre elles ont déclaré inclure du discours spontané tandis que la tâche administrée a consisté en une description d'image. Seulement quatre études ont utilisé du discours semi-spontané de type entretien semi-dirigé et aucune n'a inclus du discours spontané. L'évaluation de la qualité des études analysées a révélé qu'un peu moins de la moitié d'entre elles a respecté des critères suffisants pour être répliquée par d'autres équipes.

Pour finir, près d'un paramètre sur deux ont été capables de discriminer significativement les groupes, avec cependant une large variabilité entre les études. La variable « débit de parole » a montré un pouvoir discriminant dans environ 50% des études dans lesquelles elle a été utilisée.

2.4 Discussion

Malgré l'intérêt qu'elle représente, l'absence de consensus concernant les marqueurs linguistiques de TNCm limite l'applicabilité de l'analyse de la parole à des fins cliniques. En effet, la littérature recense un nombre élevé de paramètres, lesquels répondent à des définitions souvent peu claires et dont les mesures divergent. Une variable qu'une étude donnée nomme « pause » ne correspond pas toujours aux pauses d'une autre étude, ce qui pose problème lorsqu'il s'agit de reproduire une expérience ou de répliquer un effet. Étant donné que la taille des cohortes de ces études est souvent restreinte, l'accumulation de preuves à travers les études pourrait rendre un tel résultat généralisable. La disparité des mesures et des variables rend cette comparaison aujourd'hui impossible.

De même, le manque d'informations à propos des populations observées, généralement induite par l'exploitation de corpus en libre accès, restreint les interprétations des résultats. Une dizaine d'études seulement se sont servies d'outils d'intelligence artificielle (c.-à-d., des modèles d'apprentissage machine simples et des modèles d'apprentissage profond) pour extraire ou examiner les variables de parole. Ces méthodes, lorsqu'elles sont suffisamment détaillées, présenteraient l'avantage d'être plus objectives qu'un procédé manuel. L'usage de programmes informatisés pour l'analyse de la parole pourrait donc un choix judicieux.

Enfin, cette revue systématique révèle qu'une poignée d'études ont choisi des paramètres linguistiques et un genre discursif en fonction des aspects de la cognition qu'elles ont souhaité explorer. Ces études ont montré une association entre les paramètres et l'évaluation objective des TNCm, soit par des tests neuropsychologiques soit par de l'imagerie médicale, renforçant ainsi les hypothèses sur les corrélats cognitifs ou neuronaux des marqueurs.

3 Apport de l'article à la recherche

Cette revue systématique est la première à mettre en exergue le besoin d'une standardisation des variables et des méthodes employées pour investiguer la cognition. En effet, elle souligne l'importance d'une définition précise des paramètres linguistiques, afin de pouvoir comparer les résultats entre études et d'assurer la reproductibilité des recherches dans ce domaine. À cet égard, elle offre une classification unifiée de paramètres linguistiques, qui pourra être enrichie ultérieurement.

Cette revue systématique a été publiée dans le volume 67 du *Journal of Speech, Language and Hearing Research* paru le 09 décembre 2024. Référence : Richard, A. B., Lelandais, M., Reilly, K. T., & Jacquin-Courtois, S. (2024). Linguistic markers of subtle cognitive impairment in connected speech : A Systematic Review. *Journal of Speech, Language, and Hearing Research*, 67(12), 4714–4733. https://doi.org/10.1044/2024_JSLHR-24-00274

Linguistic markers of subtle cognitive impairment in connected speech: A Systematic Review

Amélie B. Richard[a,b], Manon Lelandais[c], Karen T. Reilly[b], Sophie Jacquin-Courtois[b,d]

[a] Université Paul-Valéry Montpellier 3, CNRS UMR5267; Praxiling, 34090 Montpellier, France [b] Université de Lyon, Université Lyon 1, INSERM U1028; CNRS UMR5292; Lyon Neuroscience Research Centre, Trajectoires Team, F-69676 Lyon, France [c] Université Paris Cité, CLILLAC-ARP UR3967, 75006 Paris, France [d] Physical Medicine and Rehabilitation Department, Henry-Gabrielle Hospital, Hospices Civils de Lyon, 69610 Pierre-Bénite, France

Correspondence to Amélie B. Richard: amelie.richard@inserm.fr. Amélie B. Richard and Manon Lelandais contributed equally to this work. Sophie Jacquin-Courtois and Karen T. Reilly contributed equally to this work. *Disclosure*: The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.

ABSTRACT

Purpose: This systematic review covers the current stage of research on subtle cognitive impairment with connected speech. It aims at surveying the linguistic features in use to single out those that can best identify patients with mild NeuroCognitive Disorders (mNCDs), whose cognitive changes remain underdiagnosed. **Method:** We followed the PRISMA guidelines and proposed a full definition of features for the analysis of speech features. Fifty-one studies met the inclusion criteria. Most of them focused on age-related progressive diseases, and included fewer than thirty subjects. **Results:** A total of 384 features labelled with 335 different names was retrieved, yielding various results in discriminating individuals with mNCDs from controls. **Conclusion:** This finding highlights the need for harmonized labels to further investigate mNCDs with linguistic markers. We suggest two different ways of assessing a feature's reliability. We also point out potential methodological issues that remain to be resolved, along with recommendations for reproducible research in the field.

Keywords: linguistic markers, connected speech, mild neurocognitive disorders

Neurocognitive Disorders cover a large range of etiologies characterized by a decline in cognitive performance, either noticed by the individual or their relatives (American Psychiatric Association, 2022). Neurocognitive Disorders are considered mild (Mild Neurocognitive Disorders [mNCDs]) when the individual's independency is preserved. These include degenerative or acquired

diseases, such as neurodegenerative pathologies (e.g., early-stage Alzheimer's disease [AD], Lewy body dementia), vascular pathologies (e.g., stroke), brain injuries (e.g., Traumatic Brain Injury), and infectious or oncological contexts (e.g., Covid-19 sequalae, HIV, Cancer-Related Cognitive Impairment). Mild NCDs also refer to cognitive-behavioral syndromes such as Mild Cognitive Impairment (Petersen, 2016) and Subjective Cognitive Decline (Röhr et al., 2020), which are objective or subjective cognitive disorders for which no underlying neuropathology have been diagnosed. In some cases, these syndromes are the manifestation of the onset of a neurodegenerative disease.

mNCDs substantially affect an individual's quality of life (see for instance F. M. Henderson et al. [2019]about Cancer-Related Cognitive Impairment). Individuals with mNCDs report suffering from forgetfulness, high distractibility, trouble with multitasking, and difficulties with learning and language (e.g., Petersen [2016] for Mild Cognitive Impairment). Patients report substantial difficulties in their everyday lives, but current neuropsychological tests often partially reveal or fail to detect such subtle cognitive impairment. For instance, questionnaires designed to assess self-reported Cancer-Related Cognitive Impairment are weakly correlated with objective measures. This results in a discrepancy between the intensity of the cognitive difficulties reported by the patients and their scores, deemed in the range of norms (Areklett et al., 2024). One explanation is that these screening or diagnosis tests are typically conceived for the detection of more severe or advanced disorders, such as dementia for the Mini Mental State Examination (Tangalos et al., 1996). Another explanation might be the lack of ecological validity of such tests, which test a single cognitive function in optimal conditions (Costa et al., 2018), whereas mNCDs' symptoms express themselves in the patient's everyday life. Furthermore, the presence of confounding factors, such as fatigue, anxiety or depression, make the assessment complex to interpret (Vos et al., 2020). Finally, an individual's expectations about their own cognitive performances may differ depending on their daily life functioning as well as their professional activity.

The absence of clear pathological results leads to underdiagnose individuals with mNCDs, leaving them with an invisible disability. New approaches to detect subtle cognitive changes are thus needed.

Investigating mNCDs with markers in connected speech

The literature in health research shows a growing number of analyses focusing on speech and discourse. Studies have shown the interest of such linguistic analyses to track cognitive changes in an individual. These analyses can be conducted along with other objective (i.e., neuropsychological tests) and subjective (i.e., questionnaires) evaluations. Their non-invasive character, high sensitivity, ecological validity and feasibility are some of the assets accounting for their increasing popularity (Bryant et al., 2017; Lanzi et al., 2023). Language is deemed susceptible to contain markers of subtle cognitive change because of its interaction with other cognitive functions such as memory, attentional processes and executive functions. Such markers are likely to be present in all types of language production (e.g., reading, sentence repetition), but might be more easily detected in

language in context, i.e., connected speech. Connected speech is defined as stretches of speech whose production is organized according to phonological, syntactic and pragmatic rules. There is, however, little consensus on the best linguistic markers for identifying individuals with mNCDs in the general population. This results in an extensive number of features with unclear labels and various measures, as highlighted by Boschi et al. (2017), hindering the replicability of such studies. There is thus a need for a well-described easy-to-use feature classification.

One scoping and two systematic reviews have been carried out on methods involving speech. These studies investigated Alzheimer's Disease in association with more severe cognitive impairment than mNCDs and MCI. They have found that both manual and automatic methods for speech recognition or analysis can distinguish individuals with Alzheimer's Disease from healthy controls or from individuals with Mild Cognitive Impairment (de La Fuente Garcia et al., 2020; Filiou et al., 2020; Ivanova et al., 2023; Martínez-Nicolás et al., 2021). As pointed out by Filiou et al. (2020), the numerous methods and linguistic markers highlighted by these reviews raise the question of their clinical applicability. However, these reviews focused on two conditions, namely Alzheimer's Disease and MCI. Regarding the use of features, Filiou et al. (2020) highlighted that semantic and fluency measures are the most likely to find a significant effect of group difference in early-stage Alzheimer's Disease or MCI. Since mNCDs cover a large range of underlying neuropathologies, we expect different patterns of linguistic markers depending on the etiology, as instigated in Boschi et al. (2017), even though a direct mapping between linguistic features and pathologies / syndromes will not be the focus of this review. Furthermore, while linguistic markers for Alzheimer's Disease have been identified, little is known about those for subtle cognitive impairment.

Objectives of this review

The aim of this review is to give an overview of the markers currently in use to detect or investigate mNCDs independently from the underlying pathology or cognitive-behavioral syndrome. This systematic review covers the current stage of research on subtle cognitive impairment with markers in connected speech by examining a large number of recent publications. It aims at surveying the long list of speech features in use to identify those that can best detect subtle cognitive impairment in patients with mNCDs, a population whose cognitive changes remain underdiagnosed, particularly in the case of acquired lesions. At the same time, it provides a current synthesis of the methods in use for the detection of subtle cognitive impairment with the help of linguistic markers.

Terminology used in this review

The terms "linguistic variable", "linguistic feature" and "linguistic marker" can be found in the literature either with different or interchangeable meanings. In this review, we use "feature" to refer to a specific speech unit or phenomenon, "variable" when the speech unit/phenomenon is an entry

of a statistical or machine learning method, and "marker" when the speech unit/phenomenon can successfully identify the case population from controls. We also made a distinction between "feature" and "measure". "Measure" refers to the way the feature is analyzed (e.g., count, ratio). However, some linguistic features are inseparable from their measure such as "speech rate" and "articulation rate".

Organization of the review

The Introduction gave a brief description of mild Neurocognitive Disorders and spells out the main benefits and challenges of using connected speech for their detection. In the Method section, we describe the article selection process, the data extraction and synthesis, and the study quality scale we used to assess the data. The Results section gives an overview of the characteristics of the studies selected, then focuses on trends in the extracted linguistic features and their power to discriminate individuals with mNCDs from a control population. The Discussion addresses the critical topics and challenges in using features of connected speech for detecting mNCDs. We first focus on the high variability in the reported results in the studies selected, and suggest two different ways of assessing a feature's reliability. We then point out potential methodological issues that remain to be resolved, along with recommendations for reproducible research in the field. In the Conclusion, we raise questions that could prompt new methodological developments and offer a number of starting points for future studies in health and speech sciences. Finally, we make our data collection tables available in the supplementary material, along with our search equations and study assessment scale.

Method

We searched the following electronic bibliographic databases: PubMed, ScienceDirect, Embase, Web of Science, Google Scholar. The protocol of this study was registered on 02/04/2023 in the PROSPERO database (ID number: CRD42023394729) and its preprint was uploaded to a non-peer review repository for protocols: https://doi.org/10.21203/rs.3.pex-2276/v1

Eligibility criteria

Only Randomized Controlled Trials and case-control studies written and published in English or French as primary studies in journals with a peer-review process were included, even if the focus of the paper was on another language. Productions referred as "grey literature" were not searched.

Population

We selected studies that included a healthy control group and individuals with mNCDs. This incorporates individuals with a cognitive-behavioral syndrome (e.g., MCI, Subjective Cognitive Decline). We also selected studies that included a pathological control group other than subjects with mNCDs since speech features may help discriminate one pathology from another one. Studies including individuals with major motor speech impairment (e.g., Apraxia of Speech) and/or individuals younger than 18 or older than 75 were excluded. Studies with fewer than 10 participants in the patient group, studies that included participants with moderate-to-severe or severe cognitive impairment were also excluded (i.e. scores below 21/30 for the Mini-Mental State Examination [MMSE] or 22/30 for the Montreal Cognitive Assessment [MoCA]). As for language, bilingualism was either an exclusion criterion stated by the reviewed studies, or not mentioned.

Connected speech tasks and neuropsychological tests

Only studies using recorded connected speech tasks were included. Reading or writing were not considered as connected speech tasks. All studies had to include at least one standardized screening test or a full standardized battery for assessing cognition.

Search strategy

We did our last search in July 2023 and included studies published between 2012 and July 2023. With the help of literature, we selected main keywords for the case population (W. M. McDonald, 2017) and connected speech tasks (Boschi et al., 2017). We added exclusion criteria such as "therapy", "care" to avoid studies unrelated to diagnosis and "articulation" to avoid studies focusing on motor speech only. Using Boolean operators, our prototypical syntax was "case population" AND "connected speech task" AND "methods" NOT "exclusion criteria" (e.g., "mild cognitive impairment" AND "picture-based description" NOT "therapy"). Exact search equations are provided in Appendix A.

Selection process

We followed the PRISMA statement (Moher et al., 2009) and reported our process in a PRISMA flowchart using a modified version of the screening and selection procedure found in Patti et al. (2015) and (Mateo, 2020). Patti et al. (2015) propose a step-by-step method to conduct a PRISMA systematic review with extensive details and examples. (Mateo, 2020) describes helpful, easy-to-use methodological tools implemented in Excel and Zotero.

We selected our articles for inclusion following a three-step procedure:

1. Pre-screening: focused on publication language, publication type, duplicates, incorrect metadata — conducted by one author (A.B.R).

2. Title and abstract screening: excluded studies that did not focus on connected speech, including studies that used naming tasks or verbal fluency tasks. The following exclusion criteria were applied: (a) studies involving individuals with major motor disorders (e.g., apraxia of speech), (b) studies focusing exclusively on reading and writing tasks, (c) studies treating unrelated topics (e.g., qualitative studies in sociology focusing on discourse analysis). This step was conducted independently by two authors (A.B.R & M.L.). We ran an inter-rater agreement test, which reached a substantial 84% agreement (Cohen's $\kappa = 0.68$). Agreement for the remaining abstracts was achieved by both authors reviewing them again and discussing their eligibility.

3. Full-text selection: articles were excluded using the following criteria: (a) *population*: No control group; mean age below 18 or above 75; fewer than 10 participants per case group; scores below normal cognition or mild disorders cutoffs at screening tests, (b) *tasks*: Studies that did not include a connected speech task, (c) *analysis*: Studies that did not explicitly analyze speech features, (d) *methods, statistics and study design*: Descriptive statistics only, meta-analyses; systematic reviews, case studies; qualitative studies; theoretical or methodological articles.

This step was conducted independently by two authors (A.B.R & M.L.). An inter-rater agreement test was run and reached near perfect agreement (96% and Cohen's $\kappa = .90$).

Data extraction process

Data were classed into four categories: (a) population, (b) tasks, (c) tests, and (d) methods. The data in each category are described in Table 1. When studies included more than 10 linguistic variables, we chose to only report those markers discriminating patients with mNCDs from controls. This specifically applies to studies using Machine Learning methods where more than a hundred features were analyzed.

Due to the nature of the studies included in our review we used a modified version of the National Institute of Health Quality Assessment of Case-Control Studies scale to evaluate the quality of each study and its risk of bias. We removed questions on concurrent controls and exposure/risks which were not applicable, and we added five items on the connected speech tasks, the linguistic variables, and the statistical or machine learning methods used to compare case and control populations. Appendix B describes the guidance and questions for assessing studies with our modified version of the NIH grid. Items 1 to 12 have been left unchanged, and items 13 to 17 were added for the purpose of this systematic review. Two authors (A.B.R & M.L.) independently conducted the risk of bias and quality assessment and then discussed to reach consensus on a paper's quality (Low, Fair, Good). A study reached a Good score when its method was deemed fully replicable, a Fair score when it lacked descriptive information on the linguistic variables or when it mislabeled discourse

Table 1. Types of data extracted from the reviewed studies.

Category	**Data extracted**
Population	• Number, age, language, for both case and control populations • Pathology for case population
Connected speech task	• Name of the connected speech task • Type of elicited discourse according to the authors • Type of elicited discourse according to our categorization
Tests	• Name of the neuropsychological tests used • Cognitive domains assessed according to the tests' authors
Methods	• Name of the linguistic features extracted • Feature extraction mode (manual vs automatic) • Statistical tests used for discriminating case population from control and their *p*-value • Machine Learning models used

type, and a Low score when information was missing in two or more items of the scale, hindering its replicability.

Synthesis methods

Feature synthesis

The assessed cognitive functions were grouped based on the functions reported by the authors of the included papers and on the neuropsychological tests they reported using.

To facilitate the analysis of the linguistic features, they were clustered into specific, hierarchized categories by two language scientists (A.B.R, M.L.) on the basis of similar characteristics (e.g., sound, meaning, word order). We for instance grouped all speech pause variables and then split them into silent pauses and filled pauses. When no detail was given on the characteristics of a pause (whether silent or filled), we left it into the larger, superior category labelled pauses (not specified). We finally grouped the categories into four large main linguistic domains: phonetics-prosody, lexicon-semantics, syntax-morphosyntax, and discourse. These four large domains are visible in Table 2 (first column). We chose a loose classification system because boundaries between smaller-sized domains can be fuzzy, as some features might belong to more than one. Features which did not fit any specific categories either by being not readable by humans (which is sometimes the case with machine learning feature extraction) or by lacking information were ascribed to a linguistic domain. Our classification system is available on OSF: https://osf.io/qhtzp/?view_only=249a568932d24cf7986159df2a0eb30d

Boschi et al. (2017) suggested classifying markers based on the descriptions reported by the study authors. In total, they reported one hundred and twenty markers, and provided a definition for each of them. In addition to surveying the descriptions provided by the study authors we included, we searched the literature in language sciences for comprehensive feature definitions. We particularly focused on studies in speech production with experimental data if available, in an effort to build an interface between clinical studies and linguistic studies. This process resulted in twenty-three definitions with clear labels for markers. Table 2 gives full definitions as well as examples of the twenty-three labels we identified. We kept measurement types (e.g., count, duration, rate) apart from observed variables (i.e., linguistic feature) whenever possible to enhance precision and replicability for specialists in other fields.

Table 2. Proposal for a harmonized terminology of linguistic markers.

Linguistic domain ***Category***	**Linguistic features**	**Definition**	**Examples**
Phonetics-Prosody *Disfluencies*	Silent pauses	Absence of intensity and F0, any interval where the amplitude is undistinguishable from that of the background noise above 200 ms. (see for instance, Duez [1982])	A: and this woman drives her kids to # school
	Filled pauses	Any type of filler particle that delays speech production. "A phonetic exponent which is segmentally structured, semantically empty, syntactically unconstrained, and does not show an interjectional function". (Betz et al., 2023)	A: and this woman drives her kids to uh school
	Breaks	Self-interruption from the speaker involving an incomplete delivery of a phonetic, syntactic, or morphological unit as a result of its abandonment. Can lead to the resumption of a new unit as part of self-repair. At the lexical level, includes deletions, substitutions, insertions, and articulation errors linked to missed phonological targets. (Pallaud et al., 2019)	With self-repair: A: and this woman drives her daugh- her kids to school Without self-repair: A: and this woman drives her kids to #
	Repetitions	Duplication(s) of a phoneme, syllable, word, or phrase. (Fox Tree, 1995)	A: and this woman drives her k- kids school
Speed of speech	Speech rate	Measure for speed and density per time unit, as the number of syllables produced per minute of speech. (De Jong & Wempe, 2009) For studies relying on a fully automated data extraction, a robust definition can be found in (He et al., 2023): "number of nuclei / (Total nuclei duration + Total internuclei duration)"	
	Articulation rate	Measure for speed and density per time unit, as the number of syllables produced per minute of speech, without silent pause or laughter time. (Miller et al., 1984)	
Length of speech units	IPU duration	Inter-pausal Units are speech segments separated by a 250 ms silent pause, with a minimum duration of 300 ms. (Bigi & Priego-Valverde, 2022)	A: and this woman drives her kids to school (2390 ms) # (250 ms) I didn't know she had kids (1770 ms)
	Turn duration	Conversational turn-taking system. A speech turn (turn constructional unit) is a stretch of speech by one speaker during which other participants assume the role of listeners, until a point of "projected completion" (Sacks et al., 1974), called a transition relevance place (TRP), which can include a silent pause. A turn shift can occur (but does not have to) at a TRP. If no turn shift occurs at the TRP, the turn continues.	A: and this woman drives her kids to school (2390 ms) # B: Did you know she had kids?
	Speech duration	Total speech time duration. Whether this includes silent pauses and other vocal phenomena such as laughter has to be indicated.	A: and this woman drives her kids to school # I didn't know she had kids (4410 ms)

(Table continues)

(Continued)

Linguistic domain ***Category***	**Linguistic features**	**Definition**	**Examples**
Lexicon-semantics *Words*	Lexical items	Refers to content (i.e., concepts, actions, beings or objects with clear mental representation) and open-class words (i.e., nouns, [non-auxiliary] verbs, adjectives, and [some] adverbs).	A: and this woman drives her kids to school
	Grammatical items	Refers to function and closed-class words. No independent meaning on its own or no full-fledged semantic features. Can be prepositions, pronouns, conjunctions, auxiliary verbs, or [some] adverbs.	A: and this woman drives her kids to school
	Lexical richness	Also referred to as lexical diversity and density, is used to assess the number produced by a speaker, showing the proficiency of a discourse. Often measured with a type-token ratio (number of different words divided by the total number of words) or Shannon's Entropy.	
Syntax-morphosyntax	Clause	A subject and a predicate (**verb** + any complement). Sometimes labeled as sentence, utterance, phrase, or proposition.	A: and this woman **drives** her kids to school
	Clause completion	Refers to the ability for a participant to finish a clause in accordance with grammatical rules.	
	Clause length	Number of lexical and grammatical items in a clause. Whether this excludes or includes filler particles has to be specified.	
	Mean Length of Utterance (MLU)	In the studies we reviewed, the definition was the following: MLU is calculated by dividing the total number of words by the number of utterances. (Galetto et al., 2013; Lowit et al., 2022)	
	Syntactic errors	Clauses with inconsistent syntax, grammar, or morphology (for instance using a noun in a verb slot, or using an incorrect inflection for tense). Dependent on context of use and language variety (e.g., African American Vernacular English, General American, Southern American English).	A: and this woman have driven her kids to school.
	Item class	Also called word classes or lexical categories. POS tagset: syntactic class of the item. This list includes: nouns, verbs, adjectives, adverbs, pronouns, prepositions, conjunctions, articles, interjections.	

(Table continues)

(Continued)

Linguistic domain ***Category***	**Linguistic features**	**Definition**	**Examples**
	Syntactic complexity	No consensus yet. This is what we propose based on the studies we reviewed. Number of items which are additional to the obligatory constituents in a clause. Includes subordinators and coordinators. Can also be Number of complex sentences (made up of several clauses) as opposed to simple sentences (made up of one clause). This has to be made explicit.	
Discourse	Informativeness	Notion of relevance (Grice's maxim Be relevant). Proportion of discourse segments on topic vs. off topic in the case of a thematic task (i.e., description or narrative). Main discourse structure vs. side discourse structure. (Van Kuppevelt, 1995)	A: and this woman drives her kids to school # I didn't know she had kids # then she goes to the office
	Coherence	How different segments of discourse are made to relate to one another. (Kehler, 2006; Labov, 2001; Wright et al., 2014)	Coherence: A: and this woman drives her kids to school # I didn't know she had kids # then she goes to the office Incoherence: A: and this woman drives her kids to school # I didn't know she had kids # we were only two at the office
	Cohesion	Consistent use of reference in speech. Consistent use of cohesive devices such as anaphora (e.g. through pronouns and noun phrases), connectives and simultaneity markers. Can also include temporal cohesion. (McNamara et al., 2010)	A: and this woman drives her kids to school # but she [another woman] doesn't know that
	Epistemic and Pragmatic events	Epistemic events: Use of any hedging expression (Schiffrin, 1987). A speaker's lexical reference to their own uncertainty or inability; use of modal auxiliaries; use of hedging discourse markers (see [Cuenca & Crible, 2019]). Pragmatic events: speech overlapping with another participant's turn; turn change (can be forced turn-taking as in interruptions); turn retention; use of conversational feedbacks (also called backchannels).	Epistemic events: A: and this woman drives her kids to school # I can't remember her name # then she goes to the office Pragmatic events: A: and this woman drives her kids to school # B: mhm # A: then she goes to the office
	Theory of Mind (ToM)	Items showing the speaker is aware of others' different representations and mental states. Includes items that refer to the acknowledgment/consciousness of shared knowledge between participants.	A: and this woman drives her kids to school (i.e., no need to name the woman because the speaker knows that their co-speaker knows the woman in question)

Note. IPU = interpausal unit; POS = part-of-speech.

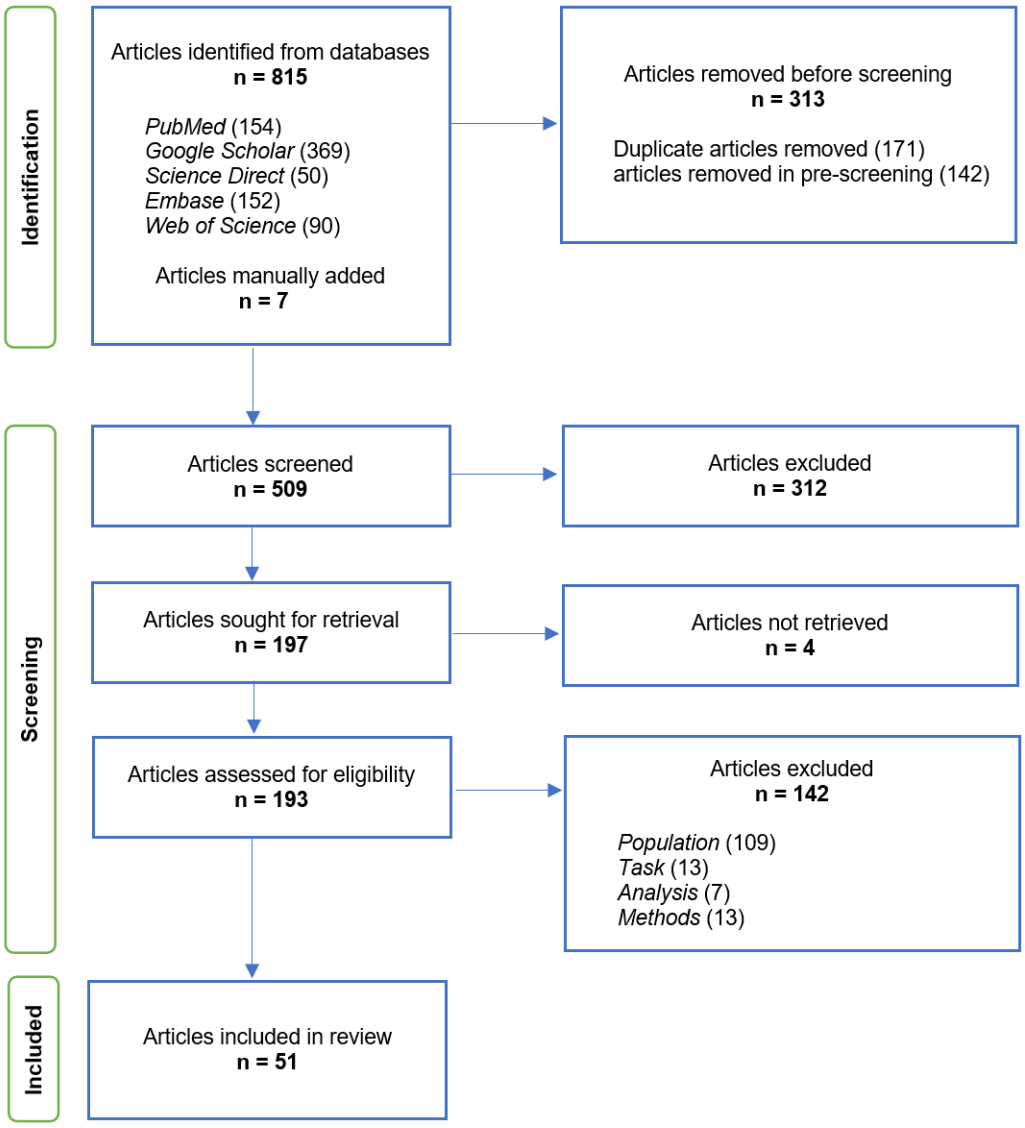
Discourse types

To compare the features depending on the speech task, we classified the task into four discourse types. Descriptive discourse gives details about a referent (i.e., a picture, an event, etc.) without any interpretation while narrative discourse refers to story-telling. Procedural discourse follows a logical structure to explain the steps of a process. Finally, semispontaneous discourse refers to speech that is elicited by context rather than a task, as is the case for everyday speech.

Results

The database search retrieved 815 records. A further 7 records were manually identified for a total of 821 records. Duplicates and unrelated topic papers were discarded in a pre-screening phase leaving 313 records. Of the 509 screened records 312 were removed based on the title and abstract, leaving 197 records for full-text retrieval. Despite direct requests to the authors four full texts were irretrievable. Out of the 193 full-texts that were screened, we excluded 142 for various reasons, including inappropriate population, tasks, methods and/or analyses. Figure 1 shows the full flowchart for the literature screening and selection process.

Figure 1. Literature screening and selection flowchart following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) guidelines.

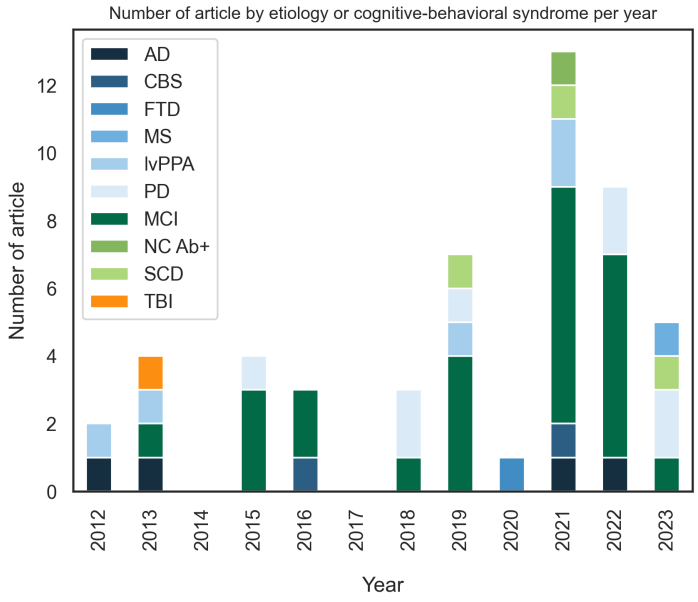


Population overview

In this section, we summarize the main results. A comprehensive summary of the extracted information is available on OSF (https://osf.io/qhtzp/?view_only=249a568932d24cf7986159df2a0eb30d). The 51 studies included in this review came from 36 different research teams, and the majority were published within the last three years. Figure 2 shows the evolution in the number of studies per pathology from 2012 to July 2023. The number of studies was relatively stable between 2012 and 2018, and more than doubled in the last 5 years, with the exception of 2020 — due to the Covid-19 period, showing the growing interest for assessing subtle cognition impairment through connected speech analysis. The most represented etiologies and cognitive-behavioral syndromes across studies were neurodegenerative pathologies. Within

these, studies mostly focused on Mild Cognitive Impairment and early onset Alzheimer's Disease, followed by Parkinson's Disease and the logopenic variant of Primary Progressive Aphasia. Only five studies focused on non-neurodegenerative diseases (i.e., Traumatic Brain Injury, Subjective Cognitive Decline, Normal Cognition with β-amyloid markers).

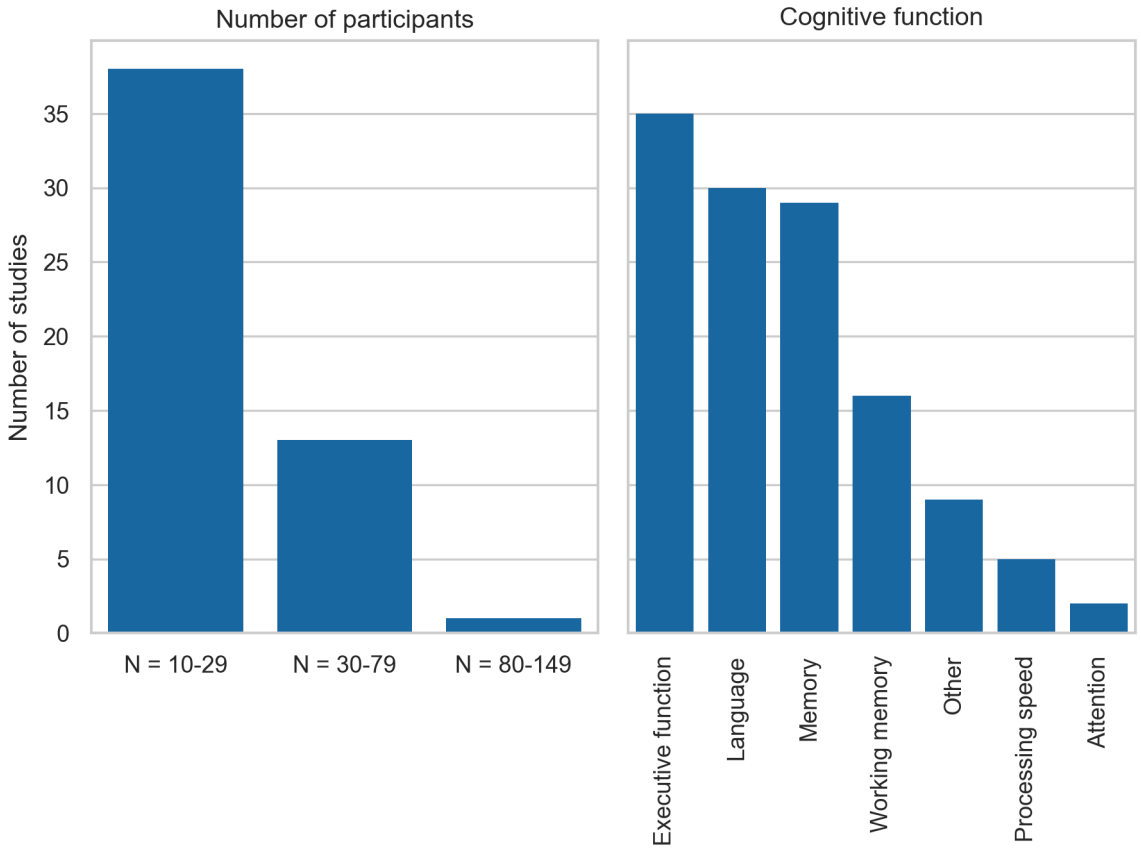
Figure 2. Number of studies per pathology per year. AD: Alzheimer's Disease, CBS: Cortico-Basal Syndrome, FTD: Fronto Temporal Dementia, lvPPA: logopenic variant of Primary Progressive Aphasia, MCI, Mild Cognitive Impairment, MS: Multiple Sclerosis, NC Ab+: Normal Cognitive with β-amyloid markers, PD: Parkinson's Disease, SCI: Subjective Cognitive Decline, TBI: Traumatic Brain Injury. Blue shades: neurodegenerative diseases, green shades: cognitive-behavioral syndromes, orange: non-neurodegenerative diseases.



Almost half of the studies included English speaking participants (24) but a total of 14 languages were identified. Other frequent languages included Hungarian (six), French (three), Korean (three), and Swedish (three). One study did not report the participants' language.

Figure 3 shows the sample size (left) across studies as well as the assessed cognitive functions using neuropsychological tests (right). A majority of studies featured between 10 and 29 participants. Very few studies included more than 80 participants. As for cognitive evaluation, executive functions were very frequently assessed, as well as language and memory. However, attention was very rarely focused on.

Figure 3. Number of participants (left) and cognitive function assessed (right) across studies.



Linguistic features

We now turn to linguistic features. We retrieved 384 features labelled with 335 different names, indicating that most of the features as reported verbatim were study-specific, although they often referred to analogous features across studies.

Most of the studies did not use specifically developed software for the analysis of linguistic data. For studies that did use such software, Praat (Boersma & Weenink, 2024) was the most frequent tool. Features were equally extracted with manual methods and with semi- or fully automatized methods. Most of the studies did not use software for extracting the linguistic features (30/51 studies). When it was the case, python scripts or machine learning (ML) models were used. The 16 studies using ML models for analyzing the features mostly used Support Vector Machine models. An interesting point is that the phonetics-prosody domain was the most investigated domain with ML models (11/16 studies).

Figures 4 and 5 show the relation between linguistic features and case populations as well as types of assessed discourse. Figure 4 shows the different linguistic features mapped out across the case pathologies or syndromes. Studies investigating MCI were by far the most frequent (25/51).

Figure 4. Linguistic feature frequency (count) according to case population. Lighter colors represent higher numbers of studies. n = number of studies. Colors of the y-axis labels refer to the linguistic domains: blue = Phonetics-prosody, red = Lexical-semantics, green = Morphosyntax-syntax, orange = Discourse.

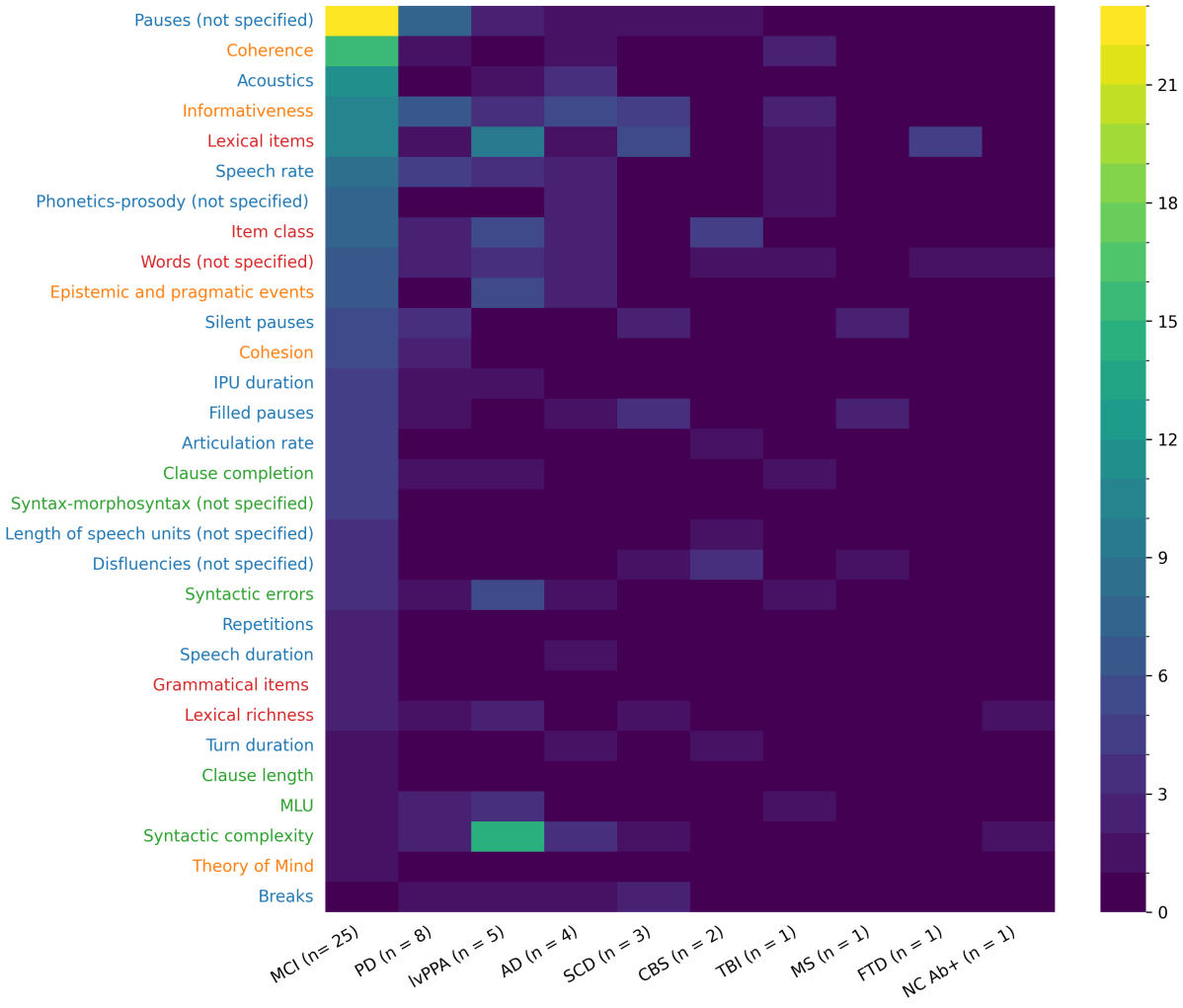
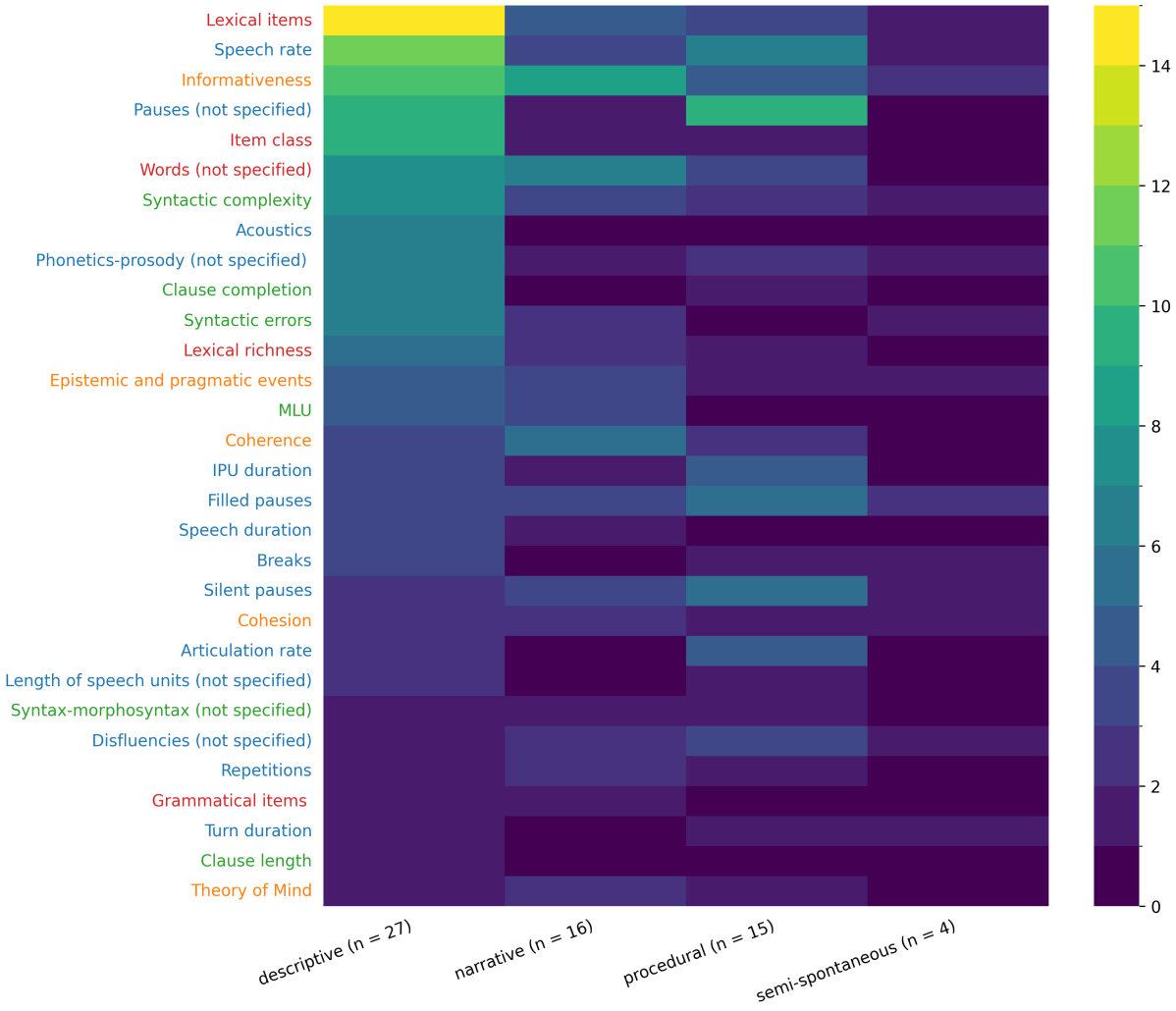


Figure 5. Linguistic feature frequency (count) according to type of discourse. Lighter colors represent higher numbers of studies. n = number of studies. Colors of the y-axis labels refer to the linguistic domains: blue = Phonetics-prosody, red = Lexical-semantics, green = Morphosyntax-syntax, orange = Discourse



These studies mostly used pauses to investigate patients'cognition, but did not specify or provide a description of the pauses. In the same domain (i.e., phonetics-prosody), acoustic features and speech rate were also frequently used. Other commonly analyzed features belong to the domain of semantics, at the word and discourse levels (such as coherence, informativeness and lexical items). Studies investigating lvPPA (5/51) stood out in that they primarily used features related to syntax (syntactic complexity, syntactic errors) and to the lexicon (i.e., a language's inventory of lexemes with features such as lexical items, item class). However, considering the variety of features and the different ways to measure them, the range of pathologies and the small sample size, we were unable to find consistent mapping between specific features and particular phenotypes.

Figure 5 shows the frequency of use of the linguistic features depending on the type of elicited discourse. Picture description was the most frequent type of assessed discourse (27 studies), often elicited by the Cookie Theft Picture (in 18 studies). Picture description was also investigated with

more features than the other discourse types. The most used linguistic features were lexical items and speech rate. The narrative and procedural discourse types were also frequently assessed (17 and 14 studies respectively), with informativeness and pauses respectively. Narrative discourse is different in the use of linguistic features in that it was mainly investigated with markers belonging to word and discourse semantics but not with markers related to rhythm. Procedural discourse, on the contrary, was investigated almost exclusively with markers related to rhythm (unspecified pauses, speech rate, silent and filled pauses).

Regarding connected speech tasks, 12 studies included more than one task, and 27 studies mislabeled at least one of the connected speech tasks (i.e., featured an incorrect classification of the assessed discourse type). The descriptive and procedural types of discourse were mislabeled in 17 and 13 studies respectively, described verbatim as either narrative or as spontaneous discourse. No study contained spontaneous speech strictly speaking (i.e., unconstrained, task-free speech in an ecological setting such as participants' homes), but four studies analyzed samples of semispontaneous discourse (i.e., speech elicited with open questions as part of semistructured interviews, which provided substantial scope for answer style or addressed topics). These four studies used filled pauses and informativeness.

Study quality assessment

Figure 6 shows the quality assessment ratings for each included study sorted by year of publication, while Figure 7 provides a summary of the quality assessment in function of each topic of interest.

The quality assessment showed that the objectives and populations were clearly described for most of the articles. The linguistic markers and methods lacked descriptive information in more than half of the studies. The tasks involving connected speech were fully described in 37 studies out of 51. Confounding variables were not assessed in most of the studies and 17 studies were deemed replicable despite missing information. Nine studies were rated Good, 21 Fair, and 22 Low. All studies remain equally important contributions to this rapidly growing field. These ratings solely apply in the context of the scale we have designed. In no way can they be considered as an indicator of the general quality of papers.

Figure 6. Quality assessment ratings.

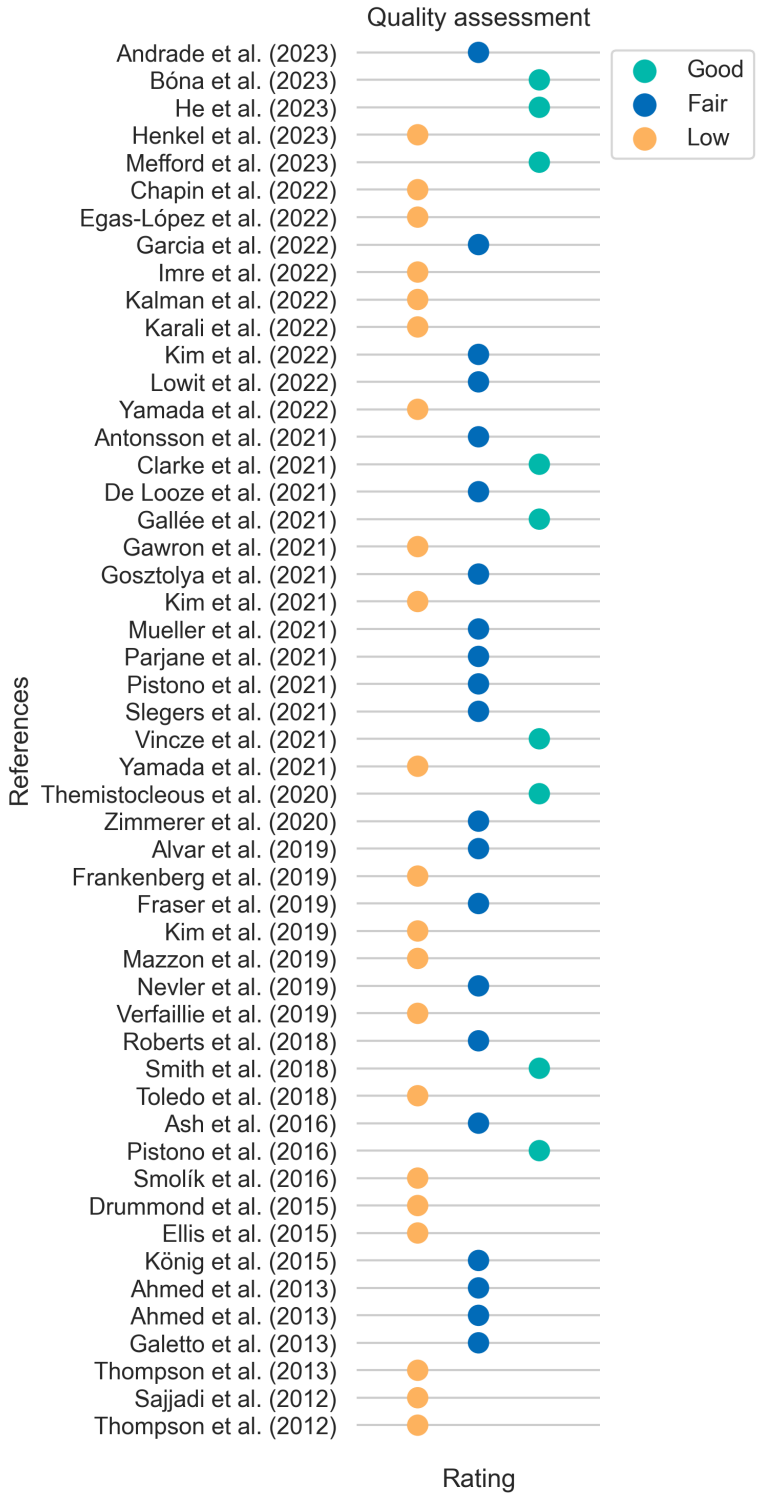
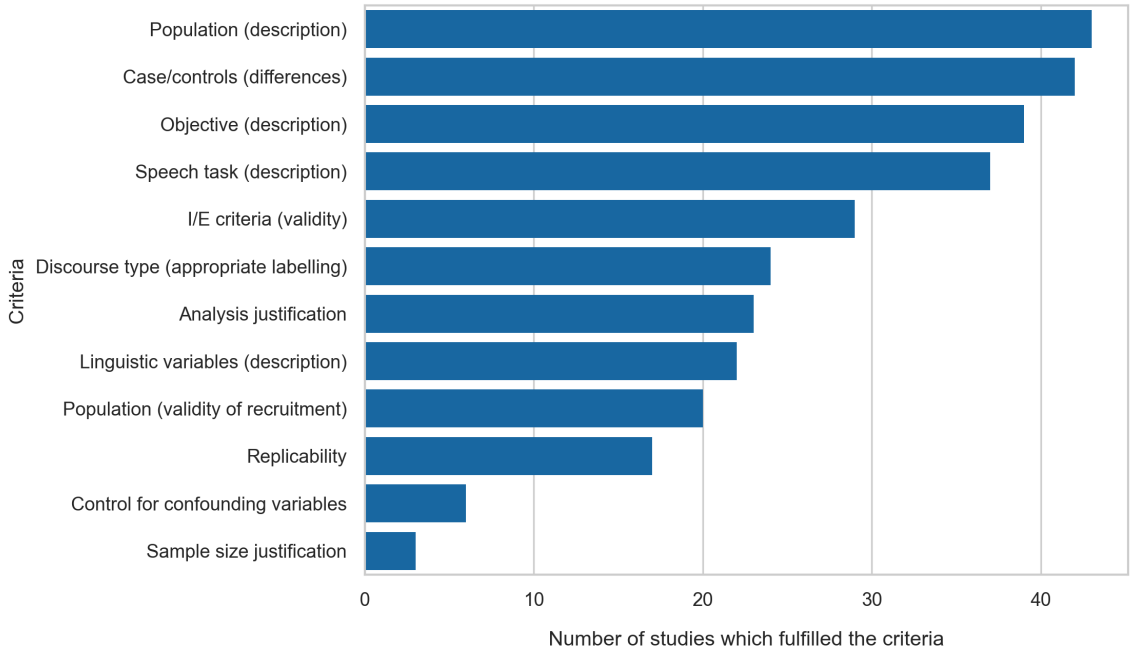


Figure 7. Summary of the quality assessment.



As shown in Figure 7, the population and objectives were often clearly stated, while less than half of the studies fulfilled the criteria related to replicability. While speech tasks were detailed in three quarters of the studies, linguistic variables were fully described (i.e., the feature characteristics) in only half of them, which might impact study replicability.

Statistical significance of group comparisons via linguistic features

We looked at the feature p-values that were reported in the studies to see whether some features showed consistent findings independently from the underlying etiologies or behavioral-cognitive syndromes. However, the reader should keep in mind that this study gathers various etiologies together, and that different results might arise from one pathology to another. Thirty-five studies used frequentist statistical tests to compare linguistic features between groups, mainly with ANOVA and Mann-Whitney U tests (Mann & Whitney, 1947). Across these studies, a significant p-value or equivalent threshold regarding the statistical tests performed to compare groups was reached for 144 linguistic features out of 258 features in total. Significance probability values were not reported for 6 of them.

Looking at linguistic features which were used in more than ten studies, lexical items, informativeness and item class showed a higher number of uses with significant p-value results than uses with non-significant p-values. Speech rate was used in 20 studies, with a similar number of significant and non-significant p-values. Equivalent numbers of uses with significant and non-significant p-values were also found for syntactic complexity and pauses (type unspecified). However,

the picture is different for specified pauses, which showed very contrasting results. Silent pauses showed a lot more uses with significant p-values, while filled pauses showed a lot fewer uses with significant p-values. Eventually, words, i.e. unspecified lexical or grammatical units, showed fewer uses with significant p-values than uses with non-significant p-values.

Summary

In terms of population, a majority of studies featured a small sample size, i.e. between 10 and 29 participants. Neuropsychological assessments of the population frequently targeted executive functions as well as language and memory, but very rarely focused on attention. Picture description was the most frequent type of elicited discourse. The linguistic features we retrieved were given study-specific labels, meaning that they were very diverse in terms of names and measures (which were often unspecified). Our quality assessment also showed that more than half of the studies lacked descriptive information about the way linguistic markers were annotated and measured. The most frequently used features across studies belonged to two different domains, namely phonetics and prosody (speech rate, pauses) and semantics (informativeness, lexical items). Pauses were also frequently analyzed, but studies did not often specify or provide a description of pause type. Results showed variability as well, with numerous features yielding equivalent numbers of significant and non-significant results. Lexical items, informativeness and item class showed a higher number of reported uses with significant p-value results than uses with non-significant p-values. Speech rate was used in 20 studies, with an equivalent number of significant and non-significant p-values. Equivalent numbers of uses with significant and non-significant p-values were also found for syntactic complexity and pauses (type unspecified). When pause type was specified, a sharp distinction appeared between silent pauses (more uses with significant p-values) and filled pauses (less uses with significant p-values).

Discussion

The studies we reviewed covered a large range of populations in terms of age, etiologies and cognitive-behavioral syndromes. The main finding of this review is the high variability in results across studies. In the first section, we relate this variability to the use of different labels for similar features, and to the use of many different frequentist statistical tests for similar variable groupings and datasets. We also address the diversity in the machine learning models in use. In the second section, building up on the contribution of the reviewed studies to the field, we break down two different ways of assessing a feature's reliability. The third section discusses the benefits of selecting a discourse type to elicit in function of its linguistic characteristics. Finally, the last section addresses the general implications of investigating underlying cognitive mechanisms in speech.

The overall objective of this section is to discuss potential keys to reach methodological consensus

while sustaining plural outlooks on subtle cognitive impairment in connected speech.

High variability in results

Different labels for similar features

Our findings show a high diversity across studies in the linguistic features reported as reliable markers for detecting subtle cognitive impairment. This mainly stems from disparities in the way studies measured the features. Specifically, markers that were measured in different ways were reported with a same label across studies. For instance, one study may measure the feature "word number" excluding fillers and word fragments, while another study may refer to the same feature while including fillers and word fragments. This is particularly true for features that were often conflated with their measures. For instance, speech rate was one of the most frequently used markers; it is a feature which also is a measure per se. While most of the studies described speech rate as the number of words per minute, other studies reported measuring speech rate as the number of syllables per second. In addition, the use of umbrella terms yielding multiple interpretations such as "word", "utterance", or "sentence" might contribute to disparities in measurement as well. For instance, "uhm" as a marker may be classified as a word or as a pause depending on studies. Similarly, we rarely found a statement on whether word fragments (e.g., "her daugh- her kids") were counted as words, repetitions, or were removed from the count.

Our findings regarding the diversity in marker measurements echo those of Boschi et al. (2017) and back their observation that a large diversity in markers hinders cross-study comparisons.

Multiplicity of methods

Another potential cause for the variability in results lies in the multiplicity of quantitative methods used to discriminate cases from healthy controls or non mNCD controls. In this section, we specifically address the variation found in the frequentist statistical methods and in the machine learning models that were used.

A large majority of studies used frequentist statistical methods. Such methods are however hard to implement because of the non-parametric nature of the data as well as the small sample sizes. The characteristics of language data, especially when acquired for clinical purposes, will usually entail many confounding variables such as participant age, participant treatments, or time gap between disease onset and speech recording. The suitability of classical statistical tests (i.e., ANOVAs, t-tests) is increasingly questioned for such data (Gries, 2015). Research in statistics for health sciences suggests that Mixed Effects Models might provide more accurate results by controlling the confounding variables as random effects (e.g., participants) (Winter & Grice, 2021). Although Mixed

Effects Models are not as straightforward to run and interpret as other statistical methods, open-source software such as JASP now provide user-friendly Linear Mixed Models tools (JASP Team, 2023).

Another way to discriminate mNCDs from controls is the use of machine learning models, which rely on the automatic classification of populations. The main advantage of using such methods relies on the number of features (sometimes more than a hundred) that can be implemented and analyzed at once, increasing the likelihood of detecting subtle cognitive change. However, the sixteen studies using machine learning showed inconsistent model accuracies, sometimes even within a single study. While model type (e.g., Support Vector Machine, Random Forest) may differ across studies, the heterogeneity of classifiers might not fully explain the variability in accuracy.

One other reason for the high variability in results might be related to data quality. This covers issues related to the quality of the speech output as conditioned by a specific type of elicited discourse, and issues concerning missing information or metadata. The first point has for instance been highlighted in Clarke et al. (2021), who noticed that discourse type impacts model accuracy. The SVM models they ran were indeed found to perform better with features from an overlearned narrative than with features from a picture book narrative. Likewise, He et al. (2023) showed that Random Forest models provide better accuracies using acoustic features than transcribed features from audio to letters (sometimes called speech-to-text features). The two previous systematic reviews on classifying AD patients' speech compared to healthy controls support the idea of a better accuracy for acoustics features (de La Fuente Garcia et al., 2020; Martínez-Nicolás et al., 2021). Despite the diversity of ML models, it is important to highlight their promising use on speech assessment in clinical context.

Another significant concern relates to the lack of information on the population from which the speech sample is drawn. Clinical data are hard to access and timely to acquire. To overcome this issue, some studies used linguistic data from readily-available datasets. While these datasets have the benefit of being open access (like the Pitt Corpus Becker et al. [1994]) their main drawback is that the researcher may lack information about the cognitive status of participants, or be obliged to work with poor-quality recordings. These data flaws might lead to biases, confounding variables, and classification errors.

The impact of data quality and model types on the results assessed in this review suggests that these elements deserve more careful consideration in future studies. Researchers aiming to use machine learning models for clinical purposes may refer to the guidelines outlined in Lo Vercio et al. (2020), where more information on the use of machine learning models can be found. Another useful resource is the Weka software (Frank et al., 2016) to easily run machine learning models, as a collection of open-source material developed by the University of Waikato, New-Zealand.

Assessing a feature's reliability: how can a feature become a marker?

There is no consensus regarding the metrics necessary for estimating a marker's reliability. In this section, we build up on the studies taken into account in this review and suggest considering a combination of two factors, namely discrimination power and replicability.

Discrimination power

The first essential factor to assess a marker's reliability is discrimination power. We define discrimination power as the number of times a feature significantly discriminates cases from controls in comparison with the number of times the feature yields non-significant results. In other words, discrimination power corresponds to the number of times the feature shows consistent (i.e. similar) results across studies.

This systematic review reveals that few features yielded significant results more times than they provided non-significant results. The high discrimination power for coherence relates to the fact that most of the studies using coherence as a marker followed consistent instructions on the evaluation of discourse coherence (see for instance Wright et al. [2014]). Similarly, equivalent numbers of uses with significant and non-significant p-values were found for pauses whose type was unspecified. Those pauses whose type was specified showed very contrasting results, with silent pauses showing high discrimination power, and no power for filled pauses. These findings highlight the need for clear marker definitions in order to enhance the replicability of results.

One limit of using discrimination power to assess a marker's reliability is linked to the general characteristics of language, where markers are used in contiguity with others, meaning that they interrelated. We found that subtle cognitive change in language is often interpreted through the prism of several markers analyzed in isolation. The odds for single markers to be sufficient to discriminate cases from controls are however not high. Compared to single markers in isolation, clusters of markers are much more likely to be informative in the detection of subtle cognitive impairment. For this reason, machine learning methods seem particularly relevant to explore cognition through linguistic markers and their patterns.

In addition to discrimination power, a marker's frequency of use is a useful indicator of its reliability. Frequency of use corresponds to the number of times the feature was used across all studies, independently from its outcome.

The results of this review showed that a majority of markers were used once, or in a small number of studies. For those features used in more than ten studies, most were able to distinguish cases from controls approximately as often as they were unable to make this distinction. For instance, lexical items and informativeness are recurrent features that showed statistically significant outcomes in more than half of the studies using them, while speech rate is a recurrent feature that showed

statistically significant outcomes in about half of the reviewed studies. Coherence, on the contrary, showed high discrimination power but was used in only eight studies out of fifty-one. This suggests that the significance of a given marker is best assessed when weighted by its frequency of use. A feature that shows either significant or non-significant consistent p-values across several studies is a good indicator of this feature's reliability. We suggest that features with inconsistent results should be further investigated to confirm or reject their reliability.

Replicability

One of the aims of this review was to improve the feasibility of speech analysis in clinical contexts. We found that very few studies reported having carried out an inter-rater agreement test as part of the cross-validation of annotations. This directly relates to the identifiability of markers, and such tests are crucial first steps for replicability (Hallgren, 2012; Holle & Rein, 2013). These findings highlight the need for harmonized methods, specifically annotation processes, when investigating speech for identifying subtle cognitive impairment.

While language may seem an accessible and relevant environment to study mNCDs, its analysis requires strong knowledge and expertise. Two major steps in speech and discourse analysis are the transcription and annotation of audio samples. These two steps are frequently done manually which leads to many human-related biases such as auditive fatigability or false interpretation of segments. To reduce these factors, the use of semi- or fully automatized tools might be worthwhile. These systems are however not perfect and need to be checked by humans. In both cases, an inter-rater agreement should be run to ensure the replicability of speech annotation. Replicability is ensured when the annotations performed by a researcher are consistent with those of another researcher using the same methods, without having access to the former annotations.

Linguistic characteristics of discourse types

This section addresses associations between patterns of features and discourse types. Discourse types differ in their structures and contents, as well as in how their ecological validity and in the cognitive load they put on the participant. When selecting discourse type, it is therefore important to keep in mind which language skills are being assessed, and by extension, which features are more appropriate. However, the results of this systematic review must be interpreted carefully as most of the included studies focus on English speakers. Indeed, each language has its own system and rules (e.g., accentuation), and linguistic markers may vary from a language family to another one. Much research is thus needed on languages other than English to consider the difference between linguistic systems.

Marker layout in function of discourse type

The results of our clustering of the linguistic markers show that markers related to the acoustic-temporal aspects of speech (i.e., speech rate) and the lexical-semantic aspects of language (i.e., lexical items) were the most frequent across all discourse types. They also show that each discourse type was explored via a different set of markers. For instance, procedural discourse was investigated with markers focusing on the temporal aspects of discourse (e.g., pauses, length of speech units), while narratives were studied with markers related to discourse structure and semantics (e.g., informativeness, words, coherence). This finding suggests that authors attempted to select features as function of discourse type. This approach seems particularly relevant since in procedural discourse participants recall elements or learned tasks in a sequential order, while in narrative tasks participants build a story incorporating elements of their own interpretation, using markers of coherence and cohesion.

This contrast does not hold for descriptive discourse, for which markers were more varied and belonged to all linguistic domains. The wider range of markers may be explained by the overrepresentation of descriptive speech. Description tasks are indeed easier to control and standardize. However, they are far from ecological, as describing entities on pictures does not reflect a participant's everyday language use in context. As highlighted Bryant et al. (2016), this raises the question of whether descriptive discourse is a good candidate for revealing/investigating subtle cognitive impairment as encountered by patients in their daily lives. Contrastively, spontaneous, or specifically semispontaneous discourse feature in very few investigations. While this discourse type is the closest to everyday language, it might also be the hardest to investigate, with many confounding variables.

A related point in question is that of the lack of harmonization regarding nomenclature on discourse types. Several studies sought to assess spontaneous speech, but used a picture-based description task, meaning that their results were uninformative with respect to their aim. Such mismatches might lead to strong biases and might create a focus on markers that are irrelevant to spontaneous speech.

In addition to previous research suggesting to mix discourse types to get a good overview of language skills (Brookshire & Nicholas, 1994), we advocate selecting different markers for different discourse types. We hypothesize that targeting specific markers sheds better light on the abilities and frailties of the population with mNCDs.

Exploring cognition through speech analysis

The final point of interest in this review relates to the selection of linguistic markers to investigate the underlying cognitive mechanisms of mNCDs. We specifically discuss the fact that few of the

reviewed studies explored patients' cognition in function of discourse type, and in function of the markers they focused on.

Hypotheses on underlying mechanisms of discourse type

Language use is cognitively complex in that it involves specific processing mechanisms such as semantics and phonology, but also relies on other cognitive functions such as memory, executive functions (including working memory [Baddeley, 2003] and processing speed) as well as attentional processes. The discourse type elicited by the speech production task will draw upon these cognitive functions in specific ways and intensities. For instance, narratives may require memory more than descriptive discourse. The output speech variables might also vary depending on discourse type. As suggested by He et al. (2023), morphological features are likely to be associated with procedural memory, and therefore may not be relevant for people with short-term or semantic memory deficits such as individuals with AD or MCI. Research in neuroscience on language learning has shown the implication of procedural memory in the implicit learning of linguistic sequences such as morphology and syntax rules (Nemeth et al., 2011; Ullman, 2016).

This can guide the selection of a particular type of discourse over another, depending on the targeted cognitive functions and the clinical context. Yet, we found very little justification about the choice of speech task relative to underlying cognitive functions, which suggests that such relations between discourse type and cognitive functions represent a promising avenue for future studies.

Neuropsychological and neural correlates of linguistic markers

When asked about their cognition, patients with mNCDs mostly report patterns of difficulties that cover a variety of cognitive domains (Verfaillie et al., 2019). It is therefore not surprising that executive functions, memory and/or language are often assessed altogether by authors. Some of the reviewed studies tested the presence of a correlation between speech markers and neuropsychological scores, and found significant results (see e.g. Parjane et al. [2021], K. M. Smith et al. [2018], and Zimmerer et al. [2020]). Such findings may help generate hypotheses on underlying cognitive impairment. For instance, De Looze et al. (2021) performed linear mixed effects models to examine whether the test scores of an isolated cognitive function (e.g., working memory) could predict temporal features in speech (e.g., number of pauses). In the procedural discourse of MCI patients, they observed associations between speech rate and memory, turn duration and working memory/attention, and Inter Pausal Units and memory. They suggest that a slower speech rate and a longer duration for pauses might signal deficits in episodic memory as well as lexical, semantic, and executive functioning processes.

Yet, these studies selected different markers along with different tests, and ran mixed models or linear regressions. Replicating such promising results that map validated test scores onto linguistic

markers is thus a long journey given the variability in methods. Furthermore, the small number of studies that have investigated associations between speech and neuropsychological variables means that the interpretation of such findings will benefit from further research.

In addition to neuropsychological assessments, some studies used neuro-imagery to investigate potential associations between linguistic features and brain regions. Pistono and Hartsuiker (2021) used both structural and functional MRI in MCI patients vs healthy controls during a picture-based narrative task to explore regions related to language and to executive functions. Although they found significant differences between groups for speech markers, they found no correlation between language network structure or functionality and language performance despite reporting a significant decrease in language gray matter compared to controls. De Looze et al. (2021) investigated the structural correlates of procedural features in MCI, targeting nine Regions Of Interest (ROIs) involved in speech production. Combining both neuropsychological assessment and fMRI techniques, they found interactions between ROIs and speech rate, turn duration, and IPU duration. The authors suggest that in addition to revealing lexical and semantic deficits, temporal features may reflect planning difficulties in speech production due to damaged working memory and attention skills.

Limitations of this review and future research

This systematic review targeted a wide array of pathologies and cognitive-behavioral syndromes to better reflect an ongoing clinical reality. The results of this review are therefore relevant for an overview of subtle cognitive impairment. The literature on linguistic markers of mNCDs is likely to increase the next few years. We suggest that future research should focus on subsets of pathologies or cognitive-behavioral syndromes. In particular there is a need to focus on underrepresented pathologies, such as Multiple Sclerosis, to better grasp the relationship between the linguistic features and the underlying cognitive mechanisms.

One bias of this review is linked to the fact that we analyzed linguistic markers based on our labelling system. While this method offers a wide linguistic perspective, it might prevent a perfect match between the study results and ours. For instance, "informativeness" as a label encompasses features such as "idea density" and "information content", while a distinction between them was made by some authors. However, we estimate that our method does not impact the readability of our results since we provided a description of the labels we used.

This study has included both manual and automatic analyses of speech. Another limitation concerns automatic analyses, in that we did not rely on Machine Learning-specific criteria to further compare these methods and models. Further studies covering specific aspects in Machine Learning methods for detecting subtle cognitive impairment are thus necessary. Since the majority of the reviewed studies focused on English, we were unable to draw any conclusion regarding possible differences in linguistic markers between language. Another study with a special focus on cross-languages comparison is thus needed.

This review was designed to offer a number of starting points for further research in clinical and language science to gain better insight on linguistic markers of subtle cognitive impairment. However, the detection of mNCDs through speech is a rapidly-evolving field, as new articles have been published since our last database search. We provide our working definitions for features, data tables, and assessment grids as their value and relevance can persist through their re-use.

Contributions of This Review This review covers different fields of interest. In this section, we state what this review adds to previous literature depending on the reader's background and interests.

1. Interest in analyzing connected speech:
    - A list of currently examined linguistic markers is available in the Results, as well as our clustering approach for categorizing linguistic features in the Method.
    - A detailed overview of the features and their definitions are proposed in Table 2.
    - The Exploring cognition through speech analysis section in the Discussion presents an opinion on current challenges in the domain of speech analysis.
2. Interest in assessing cognition using a discourse task:
    - The Method draws up a list of the various pathologies and cognitive-behavioral syndromes we included.
    - The variability in feature outcomes is revealed the Results section.
    - The Discussion provides an overview of the association between linguistic markers and cognitive impairment, as this will provide guidance when choosing features and discourse tasks.

Conclusion

Our systematic review targeted mNCDs, which covers a large range of underlying pathologies characterized by a subtle decline in cognitive performance This review aimed at surveying the speech features in use to identify those that can best detect subtle cognitive impairment in individuals with mNCDs, whose prevalence is increasing. The underlying physiopathology of such subtle cognitive changes often remain undiagnosed. Patients complain of cognitive impairment which can hardly be detected by current neuropsychological tools. In order to improve the reliability of markers, future studies should specify the markers they used, as well as the type of discourse they elicited along with the methods to annotate and analyze the data. This systematic review highlights the need for a harmonized terminology. A first step towards this is a detailed description of linguistic markers to improve the interpretation and the generalization of studies. A further development of this review is

therefore to provide a detailed paper on harmonized labels and definitions of the linguistic markers in connected speech.

Author contributions

Amélie B. Richard: Conceptualization, Methodology, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration, Funding acquisition. **Manon Lelandais**: Conceptualization, Methodology, Data Curation, Writing - Original Draft, Writing - Review & Editing, Supervision. **Sophie Jacquin-Courtois**: Conceptualization, Writing - Review & Editing, Supervision. **Karen T. Reilly**: Conceptualization, Writing - Review & Editing, Supervision.

Data Availability Statement

The appendices and our classification system are available on OSF: https://osf.io/qhtzp/?view_only=249a568932d24cf7986159df2a0eb30d

Acknowledgement

Amélie B. Richard is financed by a PhD grant awarded by Paul-Valery Montpellier 3 University and funded by the French Ministry of Higher Education (Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation — MESRI).

Dévoiler les marqueurs prosodiques du CRCI

1 Contribution de l'article à la thèse

Cette étude répond au premier objectif de cette thèse, à savoir d'identifier les caractéristiques de parole chez une population décrivant un CRCI. Une présélection de paramètres prosodiques a été inspectée grâce à une technique d'apprentissage automatique (*machine learning*) explicable pour trouver l'ensemble des marqueurs permettant de classer un échantillon de survivantes d'un cancer du sein ayant une plainte cognitive de celles n'en ayant pas et de sujets contrôles sains.

2 Résumé de l'article

Introduction

Environ 53 millions de personnes dans le monde vivent aujourd'hui après un cancer. L'avancée des connaissances sur les cancers accroît les chances de survie. L'apparition de séquelles de cancer, parmi lesquelles le *Cancer-Related Cognitive Impairment*, représente désormais un enjeu majeur de santé publique. Alors qu'environ trois quarts des survivants déclarent souffrir de difficultés cognitives, à ce jour les tests neuropsychologiques peinent à établir un diagnostic net.

Des questionnaires standardisés, tels que le FACT-Cog, sont souvent administrés pour contourner ce problème. Ils permettent autant d'évaluer la présence d'un CRCI que d'estimer son retentissement sur le fonctionnement quotidien des patients. Ces questionnaires font preuve d'une sensibilité et spécificité suffisantes pour différencier des survivants d'un cancer de la population générale. Néanmoins, l'utilisation de ces questionnaires est limitée par son risque de biais lié à la réponse. En effet, les patients peuvent être influencés par leurs connaissances de la méthode d'évaluation.

L'analyse de parole continue pourrait permettre de détecter un CRCI chez des survivants d'un cancer. Pour autant, aucune recherche sur les potentiels marqueurs de ce trouble n'a été menée jusqu'à présent. De plus, la palette de variables prosodiques complique le choix de paramètres susceptibles d'être informatifs.

2.1 Objectifs

L'objectif principal de cette étude est de déterminer le sous-ensemble optimal de marqueurs de parole capable d'identifier des survivantes d'un cancer présentant un CRCI, parmi celles n'en ayant pas et parmi des contrôles. La classification de ces groupes est réalisée par des modèles d'apprentissage automatiques simples, compréhensibles par un humain. De cette manière, ces modèles peuvent facilement être reproduits en clinique.

2.2 Méthode

Cinquante-sept femmes atteintes d'un cancer du sein et 14 témoins sains se sont portées volontaires pour cette étude. Les volontaires du groupe « cancer » avaient toutes terminé la phase curative de leur traitement (c.-à-d. intervention chirurgicale, chimiothérapie et/ou radiothérapie) depuis quatre à 52 semaines. L'ensemble des participantes éligibles a complété le questionnaire FACT-Cog, puis a été réparti en trois groupes selon leur score : les participantes atteintes d'un cancer du sein présentant une plainte cognitive (N_{BCyes} = 33), les participantes atteintes d'un cancer du sein ne présentant pas de plainte cognitive (N_{BCno} = 11), et les témoins sains (N_{HC} = 13). En plus du FACT-Cog, les participantes ont passé une évaluation cognitive et psychologique. Cette évaluation a compris le report de la présence ou l'absence de difficultés linguistiques (p. ex., « je cherche mes mots »), un test de dépistage cognitif (MoCA), un questionnaire concernant l'anxiété et la dépression (HADS) et la fatigue (FACIT-F).

La procédure de cette étude suit la méthode décrite dans le chapitre Méthodologie. Lors de l'expérience, les participantes ont eu pour consigne de raconter une histoire à partir d'une séquence de cinq vignettes. Leur production orale a été enregistrée, puis transcrite de manière semi-automatique et annotée. À partir d'études pilotes menées par notre équipe, quatorze variables identifiées comme potentiels marqueurs de CRCI ont été présélectionnées : durée moyenne et médiane et ratio (durée des pauses / durée totale du discours) des pauses silencieuses, pauses remplies et prolongations vocaliques, le ratio de l'ensemble des pauses (durée) sur la durée du discours, le ratio parole/silence, le débit de parole (nombre de syllabes / durée totale du discours), débit articulatoire ([nombre de syllabes / durée totale du discours] - durée des silences) et l'écart type de F0 (Hz) mesurée sur l'ensemble d'un échantillon donné.

Les paramètres de parole ont été intégrés à des classifieurs *Naive Bayes*. Ce type de classifieurs de *machine learning* se montrent particulièrement performants sur des jeux de données restreints. Deux modèles ont été testés : un premier à deux classes ayant pour but de séparer les participantes avec plainte cognitives des participantes sans plainte cognitive, un deuxième dont l'objectif était de classer les participantes selon le groupe auquel elles appartenaient.

Enfin, une analyse statistique Bayésienne utilisant des modèles linéaires généralisés multivariés

a été conduite pour estimer dans quelle mesure les paramètres de parole interagissaient avec les symptômes déclarés de fatigue, de dépression et d'état anxieux. Cependant, bien qu'il soit important de contrôler leur corrélation, il semble peu réaliste de déterminer un lien de cause à effet entre ces facteurs.

2.3 Principaux résultats

La sélection de quatre variables de parole (c.-à-d. durée moyenne des pauses silencieuses et des prolongations, ratio parole/silence et débit de parole) ainsi que de l'âge des participantes a permis de différencier les participantes qui avaient une plainte cognitive de celles qui n'en avaient pas avec une précision de 77,2% (ROC = 0,78). La combinaison du rapport parole/silence, de la durée moyenne des pauses silencieuses et des pauses remplies a permis de répartir les participantes post-cancer ayant une plainte cognitive, celles sans plainte et les sujets contrôles (ROC = 0,74, précision = 73,7%). Les prédictions erronées du modèle ont concerné principalement les participantes post-cancer sans plainte cognitive classées à tort comme ayant une plainte cognitive (6 participantes sur 11).

La comparaison des modèles bayésiens a montré que le rapport parole/silence et la durée moyenne des pauses silencieuses étaient mieux expliqués par un modèle incluant le groupe et le score de dépression comme facteurs ($SpeechSilence_{ratio}$ Groupe + Depressiontextsubscriptscore et $SilentPause_{moyenne}$ Groupe + Dépression). Les facteurs de Bayes supérieurs à 3 sont généralement considérés comme un niveau de preuves modéré en faveur du modèle au numérateur, les valeurs supérieures à 10 ont un niveau de preuves fort et les valeurs supérieures à 30 ont un niveau de preuve très fort (Jeffreys, 1998 ; Schmalz et al., 2023). L'ajout de scores de fatigue à ces modèles a considérablement dégradé leur qualité ($BF_{SilenceSpeech[Group + Depression][Group + Depression + Fatigue]}$ = 37,60, $BF_{SilentPauseMean[Group + Depression][Group + Depression + Fatigue]}$ = 509,51). De même, l'ajout de scores d'anxiété a considérablement dégradé le modèle ($BF_{SilenceSpeech[Group + Depression][Group + Depression + Anxiety]}$ = 8,72, $BF_{SilentPauseMean[Group + Depression][Group + Depression + Anxiety]}$ = 509,51). Cette comparaison a donc suggéré d'ôter la fatigue et l'anxiété de l'analyse. Ce résultat indique que le silence dans le discours des participants ne dépend pas de la fatigue et de l'anxiété, mais plutôt du fait qu'elles aient des troubles cognitifs liés au cancer et qu'elles déclarent ou non des symptômes dépressifs. Cependant, les Intervalles de Crédibilité du coefficient de régression n'ont montré qu'une tendance pour la dépression (95% CI [-0,36, -0,00]). En ce qui concerne la durée moyenne des pauses remplies, la comparaison des modèles bayésiens a montré que les données étaient mieux expliquées par le modèle ne comportant que l'ordonnée à l'origine. Autrement dit, l'ajout d'autres facteurs, tels que la fatigue, l'anxiété, la dépression et le groupe, a diminué sa qualité.

Ces résultats sont corroborés par ceux des tests de Kruskal-Wallis visant à comparer les cofacteurs démographiques et psychologiques entre les groupes. L'âge ne différait pas entre les trois groupes (X^2 (2, N = 57) = 4,85, p = 0,09). Parmi les participantes atteintes de cancer, 39 sur 44 ont signalé

un changement dans le langage et la communication depuis leur cancer (BC_{yes}, N = 31/33, BC_{no}, N = 8/11). La différence au test de dépistage cognitif (MoCA, médiane = 28, écart-type = 2,22) entre les trois groupes (X^2 (2, N = 57) = 2,69, p = 0,26) n'était pas significative. En revanche, l'anxiété et la dépression déclarées différaient significativement entre les trois groupes (X^2 (2, N = 56) = 7,71, p = .02, X^2 (2, N = 56) = 20,84, $p < 1e^{-3}$). Les scores médians HAD-A et HAD-D de BC_{yes} étaient significativement plus élevés que ceux de HC (p = .03 et $p < 1e^{-3}$). Toutefois, les scores ne différaient pas entre les groupes BC_{yes} et BC_{no}, p = 0,20 et p = 0,06, ni entre les groupes BC_{yes} et HC, p = 0,56 et p = 0,11. Il est important de noter que les scores d'anxiété et de dépression restaient inférieurs aux seuils d'alerte pour tous les groupes et ne sont donc pas pertinents d'un point de vue clinique. Une différence significative concernant la fatigue déclarée a été trouvée lors de la comparaison des trois groupes (X^2 (2, N = 56) = 26,42, $p < 1e^{-3}$). Le score médian FACIT-F pour BC_{yes} était significativement plus élevé que pour BC_{no} et HC ($p < 1e^{-3}$), mais ne différait pas entre les groupes BC_{no} et HC (p = 0,17).

2.4 Discussion

Ses sous-ensembles optimaux de caractéristiques de la parole ont été identifiés grâce à deux expériences utilisant des modèles d'apprentissage automatique. Ces caractéristiques permettent de distinguer les participantes présentant une plainte cognitive, selon le questionnaire FACT-Cog, de celles sans plainte (participantes post-cancer et sujets contrôles confondus). La première expérience a comparé les participantes avec et sans plainte cognitive. La deuxième expérience a étendu le modèle pour distinguer les participantes post-cancer ayant une plainte cognitive de celles n'en n'ayant pas et des contrôles grâce à la combinaison de la durée moyenne des pauses silencieuses et remplie, ainsi que le rapport parole/silence.

Les deux modèles ont montré une précision comparable, bien que le modèle à trois groupes ait eu plus de difficultés à identifier correctement les participantes post-cancer sans plainte. En effet, ce modèle a classé six de ces onze participantes comme ayant une plainte cognitive. Un examen plus approfondi de ces erreurs indique que cinq de ces participantes ont reçu une classification qui correspondait à leur déclaration de troubles linguistiques. En d'autres termes, les modèles ont mieux prédit les difficultés linguistiques déclarées que la présence d'un trouble cognitif général.

D'un point de vue clinique, le questionnaire FACT-Cog est plus pertinent pour évaluer la présence d'une plainte cognitive qu'une question ouverte posée en entretien. Ce questionnaire a été validé auprès d'une population oncologique et offre une bonne sensibilité et spécificité grâce à sa structure multidimensionnelle (Costa et al., 2018). Cependant, il a été conçu pour évaluer la présence d'une plainte cognitive générale et non par domaine cognitif. Dès lors, chacune de ses dimensions ne peut pas être interprétée indépendamment. En revanche, les questions posées aux participants de cette étude ciblaient le langage, ce qui peut expliquer pourquoi l'analyse de la parole correspond mieux aux difficultés linguistiques rapportées qu'à une plainte cognitive générale.

Le manque du mot, symptôme très fréquent parmi la population oncologique, se répercute sur la prosodie, notamment en entraînant des variations de débit. Il est intéressant de noter que deux paramètres de silence ont été partagés par les modèles d'apprentissage automatique. Le silence pourrait refléter davantage les déficits d'accès lexical que d'autres caractéristiques de la parole. En général, la présence de longues pauses silencieuses dans la production du discours pourrait être un signe de difficultés cognitives chez les survivantes d'un cancer du sein.

3 Apport de l'article à la recherche

Cette étude montre les avantages de l'analyse de la parole pour évaluer les troubles cognitifs liés au cancer. Tout d'abord, les données sont faciles et rapides à obtenir avec un équipement minimal. Cette méthode évite l'administration exhaustive de tests, qui rend l'évaluation chronophage pour des patients susceptibles d'être déjà fatigués. La narration basée sur l'image ne nécessite pas plus de quelques minutes. Deuxièmement, l'enregistrement de la parole n'est pas invasif. Les patients atteints de cancer subissent de longues évaluations tout au long de la phase de traitements, qui multiplient souvent les imageries médicales lourdes, telles que les scanners ou les IRM. La production de la parole évite la réminiscence de périodes stressantes, qui peuvent induire des changements dans la cognition. Enfin, le langage est un outil quotidien et écologique qui peut compléter le questionnaire FACT-Cog. En effet, l'analyse de la parole est une méthode d'évaluation implicite dans la mesure où les patients ne sont pas informés de ce sur quoi ils sont évalués.

Cette étude contribue à la recherche en psycholinguistique portant sur les troubles subtils de la cognition engendrés par une pathologie non-dégénérative. Elle ouvre ainsi de nouvelles perspectives pour le diagnostic de ces handicaps invisibles, mais néanmoins largement répandus. Enfin, cette recherche fait l'objet d'un article soumis à une revue internationale.

Using machine learning to identify speech markers of reported cognitive impairment in breast cancer survivors

Amélie B. Richard[1,2], Alexandre Foncelle[2], Sophie Jacquin-Courtois[2,3] Pauline Mouchès[2]

(1) Université Montpellier Paul-Valéry, CNRS UMR5267; Praxiling, 34090 Montpellier, France, (2) Université de Lyon, Université Lyon 1, INSERM U1028; CNRS UMR5292; Lyon Neuroscience Research Centre, F-69676 Lyon, France, (3) Physical Medicine and Rehabilitation Department, Henry-Gabrielle Hospital, Hospices Civils de Lyon, 69610 Pierre-Bénite, France

Corresponding author: Amélie B. Richard, amelie.richard@inserm.fr

Disclosure: The authors declare no competing financial or nonfinancial interests.

ABSTRACT

Purpose: Cancer-Related Cognitive Impairment (CRCI) refers to a cognitive decline induced by cancer and its treatments which may affect up to 75% of survivors and substantially decreases their quality of life. While CRCI is still hardly revealed by neuropsychological tests, and its detection remains a challenge for public health, speech analysis shows good potential to track subtle cognitive decline. **Method:** Forty-four women who completed breast cancer treatments and 13 controls were asked to narrate a picture-based story. Speech productions was recorded and analyzed using semi-automated methods. Fourteen speech features were extracted and used in machine learning models. In addition, all participants underwent psychological assessment to control for confounding factors. **Result:** The combination of speech-to-silence ratio, silent pause mean duration, and filled pause mean duration best predicted cancer survivors with CRCI (ROC = 0.74, accuracy = 73.7%). Bayesian Generalized Linear Models showed that silence in participants' speech depended on whether they had CRCI and reported depression. Credibility Intervals of regression coefficients showed a trend for depression (95% CI [-0.36, -0.00]), even though scores were below threshold alerts for all participants. **Conclusions:** Long silent pauses in discourse may be a sign of Cancer-Related Cognitive Impairment in survivors of breast cancer. This study demonstrates the benefits of using speech analysis as a fast, ecological, and non-invasive technique for assessing CRCI.

Keywords: machine learning, linguistic markers, cognitive impairment, breast cancer

Introduction

Cancer is a polyfactorial disease worldwide. According to the International Agency for Research on Cancer, the incidence of cancer in the world in 2022 is approximately 20 million new cases per year, and the number of 5-year or more cancer survivors is approximately 53 million (Ferlay et al., 2024). With advances in knowledge and technologies, some cancers, such as breast cancer, are better diagnosed and treated, increasing their chances of survival. Cancer sequelae in survivors represents nowadays a challenge to public health. One of these side effects is Cancer-Related Cognitive Impairment (CRCI) (Hurria & Lachs, 2007; Pembroke et al., 2024). CRCI refers to cognitive changes described by cancer patients. This cognitive sequelae can affect up to 75% of patients with non-Central Nervous System tumors (Janelsins et al., 2014) and can last several years after treatments ended (Dijkshoorn et al., 2021).

CRCI is characterized by reported cognitive difficulties that cover various aspects of cognition (e.g., memory, attention, language, and processing speed) (Yao et al., 2017). Despite the intensity of the complaint, scores on current neuropsychological tests struggle to fully match the difficulties reported by patients, as they are often within the norm (Ahles & Root, 2018; Costa & Fardell, 2019). One explanation is that these tests are generally designed to assess more severe disorders and might lack sensitivity for detecting subtle cognitive impairment, such as CRCI. Furthermore, they test isolated cognitive function in optimal conditions and might therefore not reflect the functioning of a patient in their everyday life.

To overcome this point, standardized questionnaires are often administered to assess subjective cognitive difficulties. The Functional Assessment of Cancer Therapy-Cognitive Function (FACT-Cog) (Costa et al., 2018; Wagner et al., 2009) was specifically designed for cancer patients. Although it demonstrated good sensitivity in distinguishing cancer participants from controls, the use of questionnaires is nevertheless limited by the potential for response bias, as participants' answers may be influenced by the awareness of the evaluation method. New tools with implicit measures are needed to fully assess subjective or objective CRCI.

The ecological nature of language along with its relationship with other cognitive mechanisms (i.e., memory, attention, executive functions, and motor functions) arouse a growing interest in the analysis of speech to monitor cognition (Lanzi et al., 2023). Speech analysis is sensitive enough to track subtle cognitive decline (Ivanova et al., 2023). While these findings make speech analysis an ideal tool for detecting cognitive impairment in cancer patients, no research on speech markers of CRCI has been conducted. Unfortunately, the wide range of speech variables found in the literature complicates feature selection (Richard, Lelandais, et al., 2024). The main goal of this study was to find the optimal subset of speech markers that can best detect CRCI using machine learning.

In this study, 14 speech features that might be good candidates for revealing CRCI in cancer survivors were preselected based on previous studies (Richard, Foncelle, et al., 2024; Richard et al., 2023). The preselected features were incorporated into a first machine learning model to discriminate participants with a cognitive complaint from those without. The group without any cognitive complaint included cancer survivors as well as healthy controls. A second model was run to classify cancer patients with a cognitive complaint, those without any cognitive complaint and healthy controls. Furthermore, for both experiments, the optimal subset of features for the task was identified within the 14 originally selected.

Finally, Bayesian statistical analysis was conducted using generalized linear multivariate multilevel models to assess the relationship between speech features and reported fatigue, anxiety, and depression. In other words, we investigated the extent to which these cofactors were correlated with speech. Moreover, while it may seem ambitious to determine causal relationships between cognitive impairment and these cofactors, it is crucial to control for their correlation.

Methods

Ethics approval statement

This research involving human participants will be performed in accordance with the ethical standards of the Declaration of Helsinki and its later amendments. The participants recruited in this study were part of three larger projects. The first project aimed to evaluate the effectiveness of a physical rehabilitation program offered to women with breast cancer who have completed the main phase of their treatment (ACTICOG), and the second project evaluated the effectiveness of a cognitive rehabilitation program for individuals with subtle cognitive impairment due to stroke or cancer (REMCOG). Both were approved by the ethics committee of the Hospices Civils de Lyon (n°22-5009 and n°22-121). The third project focuses on cognitive impairment induced by cancer and its treatments in women with breast cancer (DisCCo), and was validated by the French ethics committee Comité de Protection des Personnes Est-III in August 2023 (national number: 2023-A00267-38).

Population

Fifty-seven women with breast cancer and 14 healthy controls volunteered for this study. Cancer survivors completed surgery, chemotherapy, and/or radiation therapy within a period ranging from four weeks to one year. A threshold of 55/72 or lower on the Perceived Cognitive Impairment (PCI) FACT-Cog subscale indicates the presence of a cognitive complaint (Van Dyk et al., 2020). This score was used to divide the population into three groups: breast cancer participants with a cognitive complaint (BC_{yes}, N = 33), breast cancer participants without any cognitive complaint (BC_{no}, N = 11),

and healthy controls (HC = 13).

Speech analysis and feature extraction

Narrative discourse task

Participants performed a narrative discourse task from a five-picture sequence and were told to make up the sixtieth one (Bézy et al., 2016) with no time limit. The sequence tells the story of a family arriving at a beach. The father is windsurfing while his son and wife are on the shore. Suddenly, the wind picks up, and the father falls into the water. The last picture portrays the alarmed son and mother calling for help on the phone.

Speech recording material

The speech productions of participants in the ACTICOG and REMCOG studies were recorded in a quiet room at the Henry Gabrielle Rehabilitation Hospital (Saint-Genis-Laval, France) using a Rode Lavalier Go® omnidirectional lavalier microphone, while those of DisCCo participants were recorded in a soundproof booth at the Lyon Neuroscience Research Centre (Bron, France) using a Shure® omnidirectional lapel microphone.

Linguistic feature preselection

Table 1 describes the 14 preselected features, along with their characteristics and measures.

Table 1. Description of the speech variables characteristics and measures.

Speech features	Data Characteristics	Measures
Silent pauses	Segments characterized by an absence of speech, cough, or yawn during 200 ms or more, from the first to the last speech phonemes. Silent pauses might include lip or tongue noise or breathing	• Mean duration (ms) of pause • Median duration (ms) of pause • Ratio: duration of pauses / duration of speech

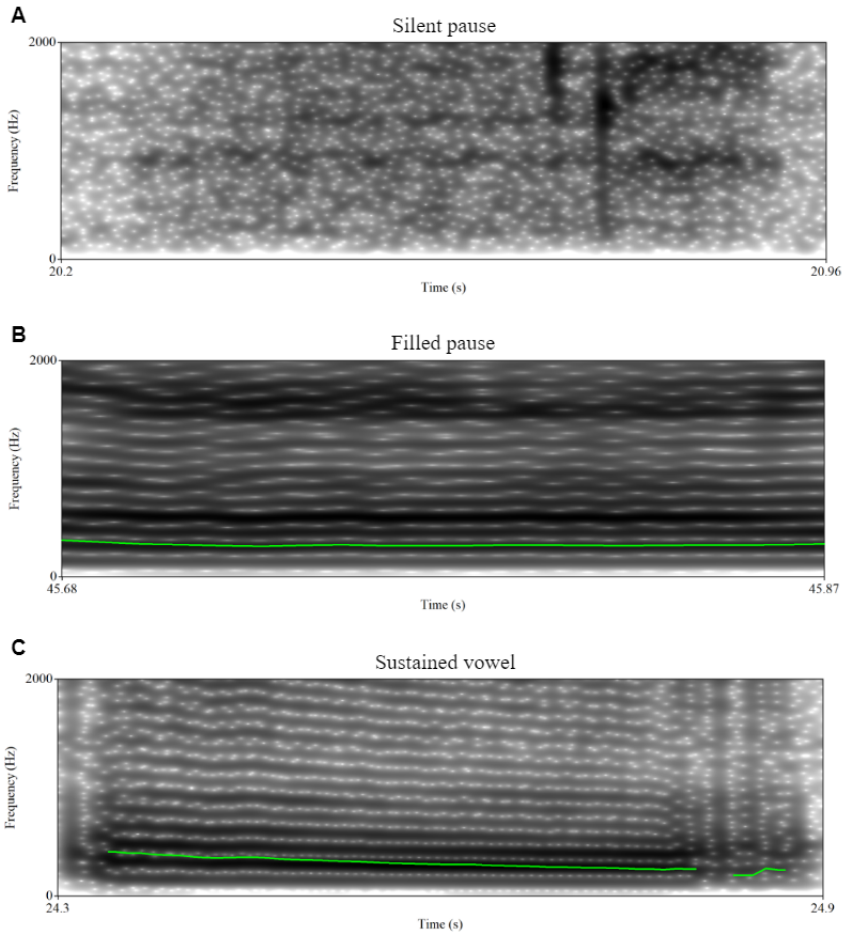
(Table continues)

(Continued)

Speech features	Data Characteristics	Measures
Filled pauses	Non-semantic voiced segments such as "uh" and "uhm" during 200 ms or more with a flat F0 contour	• Mean duration (ms) of pause • Median duration (ms) of pause • Ratio: duration of pauses / duration of speech
Sustained vowels	Prolongated vowels during 200 ms or more with a flat F0 contour	• Mean duration (ms) of vowels • Median duration (ms) of vowels • Ratio: duration of vowels / duration of speech
Overall pause ratio	Sum of pause duration (ms), i.e., silent pauses + filled pauses + sustained vowels	Ratio: duration of pauses / duration of speech
Speech-to-silence ratio	Sum of pause duration (ms), i.e., silent pauses + filled pauses + sustained vowels	Duration of speech / the duration of silence
Speech rate		Number of syllables / total duration of speech
Articulation rate		(Number of syllables / total duration of speech) - silence duration
F0 SD	Standard deviation of the fundamental frequency (Hz)	Standard deviation calculated on total speech sample

The spectrogram of silent pauses (Figure 1A) contains no formant, whereas those of filled pauses (Figure 1B) and sustained vowels (Figure 1C) display a flat, long fundamental frequency (F0) in Hz (green line).

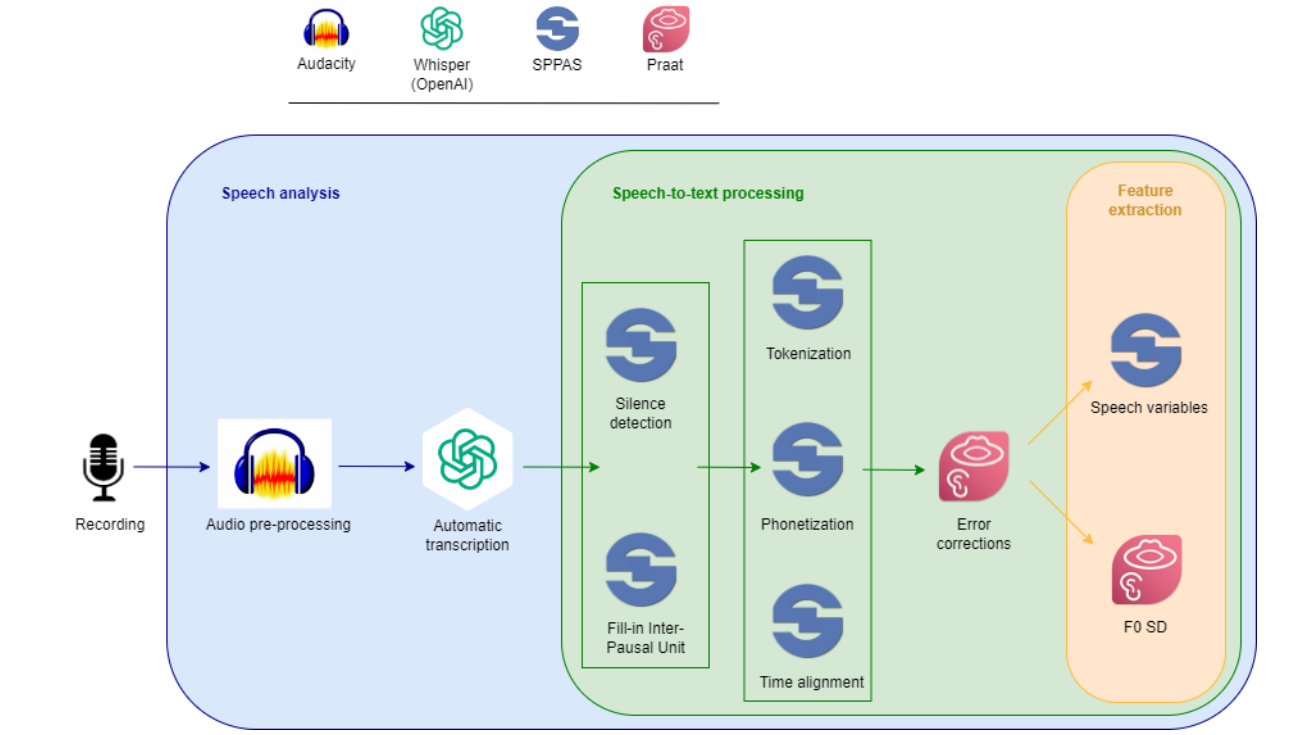
Figure 1. Canonical spectrograms of (A) silent pauses, (B) filled pauses, and (C) sustained vowels. The green line represents the fundamental frequency (F0) in Hz.



Speech analysis and feature extraction

We used open-source, offline AI tools to ensure the security of our data and replicability of this study. We first edited recordings on Audacity to ensure audio quality and to remove pre- and post-sample segments. We converted the speech samples into mono `.wav` format at 44.1kHz to automatically transcribe them using the Whisper large-v3 model (Radford et al., 2022). We then semi-automatically annotated silent pauses and extracted the features using SPPAS (Bigi, 2015). When necessary, we adjusted the boundaries to time-align audio signals with a transcription on Praat (Boersma & Weenink, 2024). Figure 2 shows the steps of speech analysis from recording to feature extraction.

Figure 2. Steps of the speech analysis.



All transcription and annotation steps were manually checked to avoid errors. Finally, we provided 20% of our dataset to two naïve annotators to verify that our definitions of pauses and sustained vowels were replicable. The inter-annotator agreement reached 98.71% agreement for silent pauses, near perfect agreement (κ = .97) for silent pauses, 90.98% agreement for filled pauses, near perfect agreement (κ = .81), and 86.88% agreement for sustained vowels, indicating substantial agreement (κ = .74).

Cognitive, psychological and physiological evaluation

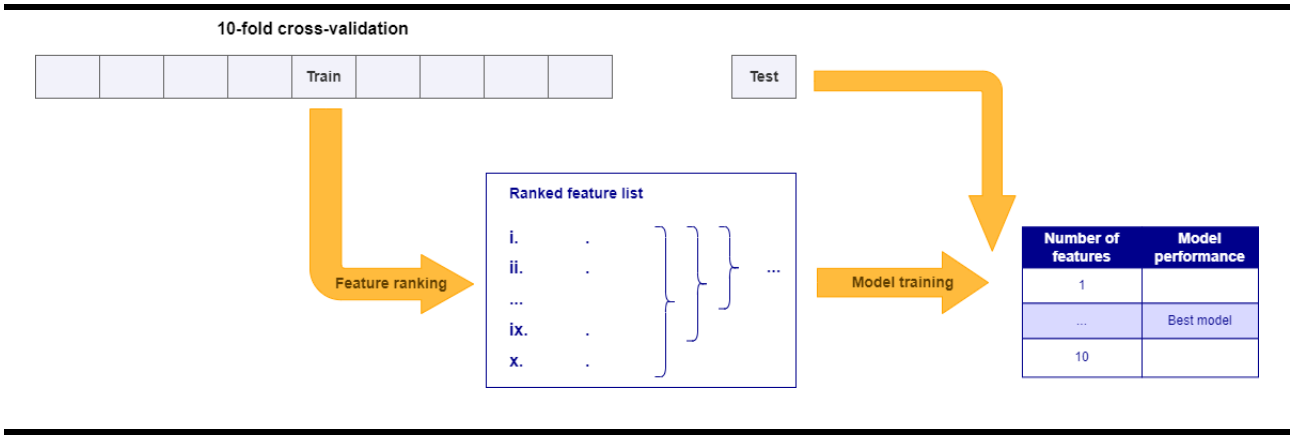
To complement the subjective evaluation of CRCI using the FACT-Cog questionnaire, participants underwent cognitive screening and psychological assessment to control for confounding variables that might influence speech production. Participants took the Montreal Cognitive Assessment (MoCA) test (Nasreddine et al., 2005), a rapid, self-administered screening tool designed to detect fine-grained cognitive impairment by evaluating several domains (e.g., visuospatial abilities, memory, and language). A score below 26/30 indicates a cognitive decline. Participants completed the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) questionnaire (Bonomi et al., 1996; Cella et al., 1993), designed to estimate cancer-related fatigue. The higher the score, the greater the level of fatigue. There was no standardized threshold for the questionnaire. Participants also completed the Hospital Anxiety Depression (HAD) scale (Zigmond & Snaith, 1983), which comprises two subscales designed to assess anxiety (HAD-A) and depression (HAD-D). A score above 11/21 on either subscale

indicates the presence of anxiety or depression. Finally, for informative purposes, cancer survivors were asked to report any change in their language or communication since their cancer diagnosis, and were allowed to describe their difficulties. In this study, their answers were analyzed as either "yes" or "no."

Machine learning method

The machine learning analysis consisted of a two-step process, including feature ranking and selection and outcome prediction, for both experiments, as illustrated in Figure 3. For feature ranking, we used the 1R feature evaluator (Holte, 1993). It evaluates the predictive power of each feature individually by defining a set of rules on the discretized feature values to best classify the samples based on the outcome. We thus obtained a list of the ranked features. To identify the optimal subset of features within the ranked list, we employed an iterative process consisting of training a first classifier using all features and incrementally removing the fewest ranked features from the training data until only two were left. The optimal subset of features was considered as that leading to the highest f1-score, which represents the harmonic mean between precision and recall. A naïve Bayes classifier was used (John & Langley, 2013). Naive Bayes classifiers assume that features are independent of each other. It uses the class probabilities and conditional probabilities of feature values given the class computed on the training data to determine the class of test samples. This algorithm has been shown to be highly efficient among various datasets and is known to perform well on small datasets, even in cases where the assumption of feature independence does not hold (Zhang, 2004). The models were evaluated using 10-fold cross-validation, consisting of iteratively using 90% of the data for model training and holding 10% of the data for testing, until all samples were used for testing once. Feature ranking and model training were performed in a nested fashion to avoid double-dipping (Lo Vercio et al., 2020). Machine learning analyses were performed using Weka version 3.8 (Witten et al., 2011).

Figure 3. Graphical illustration of the machine learning pipeline. Training data was used for feature ranking followed by model training iteratively discarding the least ranked feature. 10 folds cross-validation was employed for train/test data splitting.



Statistical analysis

All statistical analyses were performed using the last version of R language (R Core Team, 2021) in RStudio software (RStudio, BPC, 2020). We performed Bayesian analysis using generalized linear multivariate multilevel models with the brms package (Bürkner, 2021) to estimate the effects of fatigue, anxiety, depression, and group on speech variables. Non-informative priors were specified as normal distributions with a mean of 0 and a standard deviation of 3 for both the intercept and regression coefficients. The models were fitted using 43,000 iterations, with the first 2,000 iterations allocated for the warm-up. Kruskal-Wallis tests (Kruskal & Wallis, 1952) were run to compare scores between groups. Eventually, we ran a Spearman (Spearman, 1904) correlation to check the relationship between the cofactors using base R.

Results

Participants characteristics

Table 2 summarizes the results of the Kruskal-Wallis tests comparing the demographic and psychological cofactors between the groups. Age did not differ among the three groups (X^2 (2, N = 57) = 4.85, p = 0.09). Among participants with cancer, 39 out of 44 reported a change in language and communication since their cancer (BC_{yes}, N = 31/33, BC_{no}, N = 8/11). There was no significant difference in cognitive screening using MoCA (median = 28, SD = 2.22) across the three groups (X^2 (2, N = 57) = 2.69, p = 0.26). There was a significant difference in both reported anxiety and depression across the three groups (X^2 (2, N = 56) = 7.71, p = .02, X^2 (2, N = 56) = 20.84, $p < 1e^{-3}$). The median HAD-A and HAD-D scores for BC_{yes} were significantly higher than those for HC (p = .03 and $p < 1e^{-3}$). However, scores did not differ between the BC_{yes} and BC_{no}, p = 0.20 and p = 0.06, and between the BC_{yes} and HC groups, p = 0.56 and p = 0.11. Note that the anxiety and depression scores remained below the alert thresholds for all groups and are therefore not clinically relevant. There was a significant difference in reported fatigue across the three groups (X^2 (2, N = 56) = 26.42, $p < 1e^{-3}$). The median FACIT-F score for BC_{yes} was significantly higher than that for BC_{no} and HC ($p < 1e^{-3}$), but did not differ between the BC_{no} and HC groups (p = 0.17).

Machine learning results

We conducted two experiments using a naïve Bayes classifier with 14 speech features. Age was included as an additional variable in the models. We then chose four metrics (accuracy, ROC, precision, and F1-score) to evaluate the efficiency of the models.

Table 2. Characteristics of the population and scores at neuropsychological tests. Mean, SD.

Variables	**BC_{yes}**	**BC_{no}**	**HC**	***p*-value**
N =	33	11	13	
Age	50.2 (8.9)	48.8 (8.8)	43.7 (10.6)	0.09
Reported difficulties	N = 31	N = 8	N = 0	NA
FACT-Cog PCI (/72)	34.3 (12.4) ∗ †	63.3 (5.5) †	66.7 (4.8) ∗	$< 1e^{-3}$
MoCA (/30)	27.0 (2.2)	28.2 (1.9)	27.4 (2.4)	0.26
FACIT-F (/52)	25.8 (10.0) ∗ †	37.6 (6.5) †	44.3 (7.1) ∗	$< 1e^{-3}$
HAD-A (/21)	7.1 (4.5) ∗	9.9 (2.9)	6.1 (4.2) ∗	0.02
HAD-D (/21)	7.3 (3.8) ∗	4.3 (3.1)	1.8 (1.7) ∗	$< 1e^{-3}$

Note: BC_{yes} = breast cancer participants with a significant cognitive complaint, BC_{no} = breast cancer participants without significant cognitive complaint, HC = healthy controls. * = significant difference BC_{yes} vs. HC. † = significant difference BC_{yes} vs. BC_{no}.

Experiment 1: Classification of the cognitive complaint

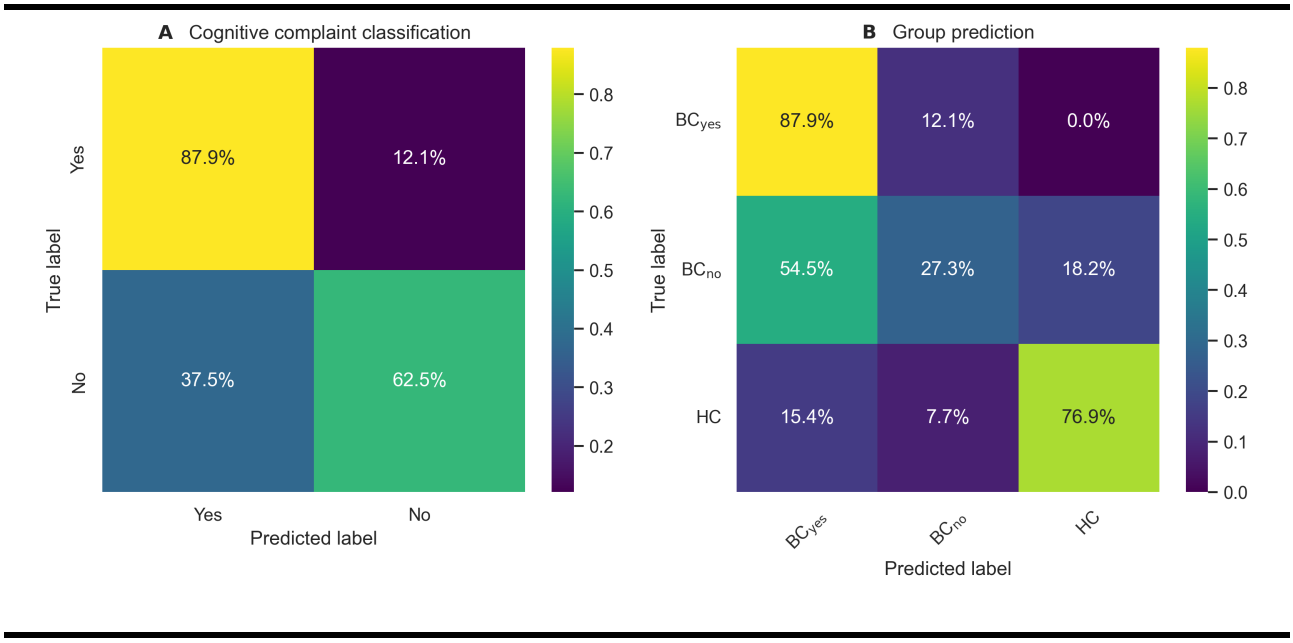
Table 3 summarizes the results and the evaluation of the two-class model, which classifies participants with a cognitive complaint from those without any cognitive complaint according to their scores on the FACT-Cog questionnaire. On average, the model was 77.2% accurate using 4 out of 14 speech features along with age. It discriminates between positive and negative classes (ROC = 0.78) and makes 78.9% of the correct positive predictions.

Figure 4 shows the confusion matrix of the two-class model with better accuracy for true positives and a high percentage of false positives, namely, participants without any cognitive complaint who have been classified as having a complaint. Specifically, the model best predicted participants who reported a cognitive complaint (N errors = 4/33) compared to those without complaint (N_{errors} = 9/24).

Table 3. Results of the two-class machine learning model for the classification of participants with a cognitive complaint and participants without any cognitive complaint. CogComplaint(positive class: Yes): participants with a cognitive complaint. CogComplaint(positive class: No): participants without any cognitive complaint. CogComplaint(average): average of the two classes.

Target	True positive	False positive	Precision	ROC	Accuracy	F1-score	Selected features
CogComplaint (positive class: Yes)	0.88	0.38	0.76	0.78	77.2	0.82	—
CogComplaint (positive class: No)	0.63	0.12	0.79	0.78	77.2	0.69	—
CogComplaint (average)	0.77	0.27	0.77	0.78	77.2	0.77	• $SilentPause_{mean}$ • SpeechSilence • SpeechRate • $SustainedVowel_{mean}$

Figure 4. (A) Heatmap of the two-class machine learning model results model for the classification of participants with a cognitive complaint and participants without any cognitive complaint. True label refers to actual groups, and predicted label refers to groups assigned by the model. Brighter colors represent a higher accuracy in prediction, while darker colors represent more classification errors. (B) Heatmap of the results of the three-group machine learning model for the classification of cancer survivors with a cognitive complaint (BC_{yes}), cancer survivors without any cognitive complaint (BC_{no}), and healthy controls (HC). True label refers to actual groups, and predicted label refers to groups assigned by the model. Brighter colors represent a higher accuracy in prediction, while darker colors represent more classification errors.



Experiment 2: Group prediction

Table 4 presents the results and evaluation of the three-group model that classified cancer survivors with a cognitive complaint, those without, and controls. On average, the model was 73.7% accurate using three of the 14 speech features. It discriminated between positive and negative classes (ROC = 0.74) and made 73.7% of correct positive predictions.

Table 4. Results of the two-class machine learning model for the classification of participants with a cognitive complaint and participants without any cognitive complaint. CogComplaint(positive class: Yes): participants with a cognitive complaint. CogComplaint(positive class: No): participants without any cognitive complaint. CogComplaint(average): average of the two classes.

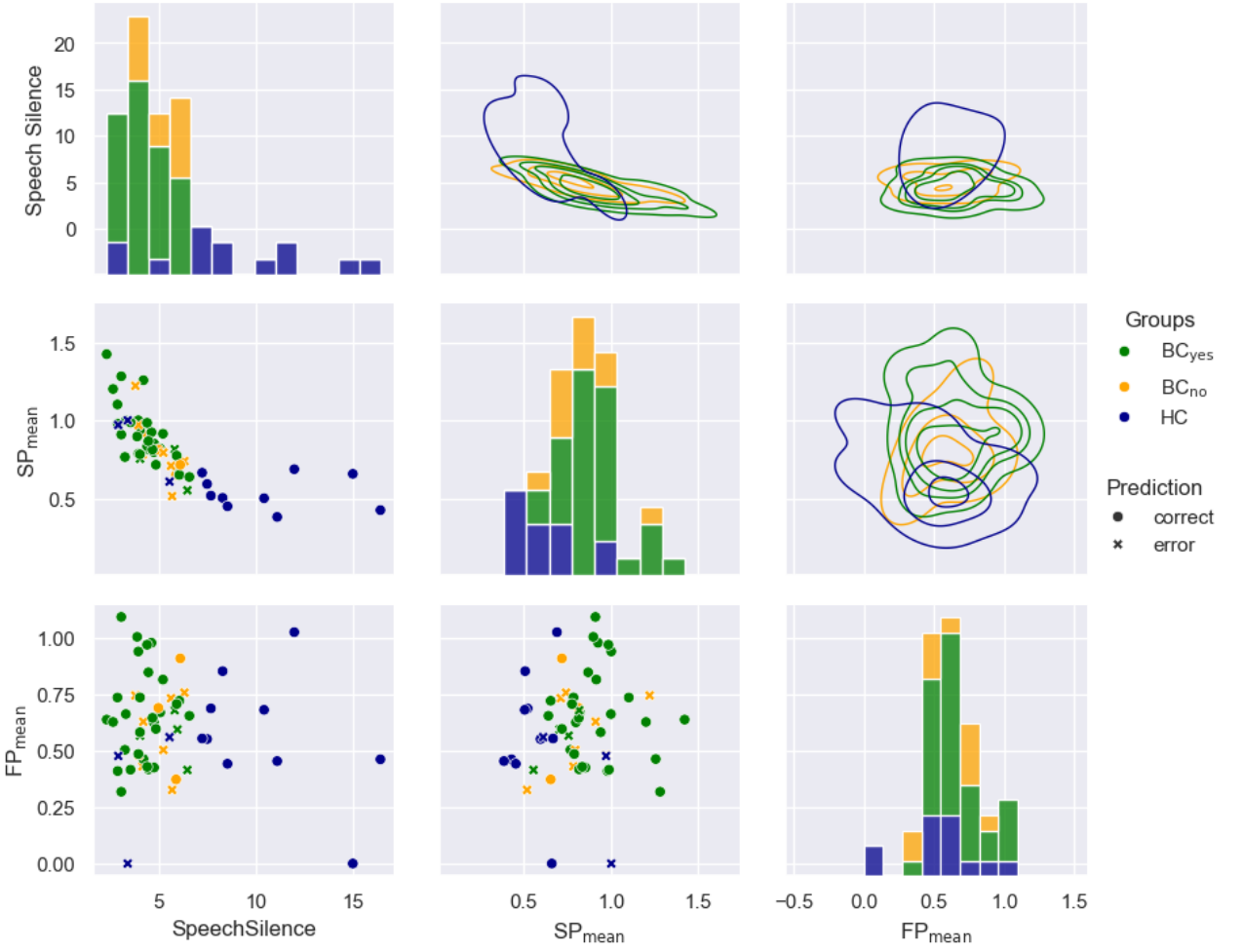
Target	True positive	False positive	Precision	ROC	Accuracy	F1-score	Selected features
Group (positive class: BC_{yes})	0.88	0.33	0.78	0.74	73.7	0.83	—
Group (positive class: BC_{no})	0.27	0.11	0.38	0.74	73.7	0.32	—
Group (positive class: HC)	0.77	0.05	0.83	0.74	73.7	0.80	—
Group (average)	0.74	0.23	0.72	0.74	73.7	0.72	• $SilentPause_{mean}$ • SpeechSilence • $FilledPause_{mean}$

Figure 4B shows the confusion matrix of the three-group model with better accuracy for true positives and a high percentage of false positives, that is, participants without any cognitive complaint who have been classified have a complaint. Specifically, the model best predicted cancer survivors who reported a cognitive complaint (N_{errors} = 4/33) and controls (N_{errors} = 3/10). In contrast, the model struggled to classify cancer survivors without any cognitive complaint (N_{errors} = 6/11).

Speech variables

We further investigated the three speech variables that were found to be optimal for discriminating between cancer survivors with a cognitive complaint, those without, and controls (see Section Experiment 3). Figure 5 shows the correlation and classification errors (lower corner), distribution (diagonal), and density (upper corner) of the three-group model results. While the group of cancer survivors with a cognitive complaint and the control group tended to stand out from each other, the group of cancer survivors without any cognitive complaint overlapped the group with a cognitive complaint. Not surprisingly, the scatter plots showed a trend between the mean duration of silent pauses and the speech-to-silence ratio, but not between these variables and the mean duration of filled pauses.

Figure 5. Correlation and classification errors (lower corner), distribution (diagonal), and density (upper corner) of the three-group model's results. $SpeechSilence_{ratio}$: speech-to-silence ratio. A higher ratio indicates a shorter duration of silence. $SilentPauses_{mean}$: mean duration of silent pauses. $FilledPauses_{mean}$: mean duration of filled pauses. BC_{yes}: cancer survivors with a cognitive complaint. BC_{no}: cancer survivors without any cognitive complaint. HC: healthy controls.



Influence of cofactors on speech variables

The Bayesian model comparison showed that the speech-to-silence ratio and the mean duration of silent pauses were better explained by a model including Group and Depressionscore as factors ($SpeechSilence_{ratio}$ Group + $Depression_{score}$ and $SilentPause_{mean}$ Group + $Depression_{score}$). Bayes factors greater than 3 are generally considered moderate evidence in favor of the model in the numerator, values greater than 10 are strong evidence, and values greater than 30 are very strong evidence (Jeffreys, 1998; Schmalz et al., 2023). Hence, adding fatigue scores to these models significantly degraded their quality ($BF_{SilenceSpeech[Group + Depression][Group + Depression + Fatigue]}$ = 37.60, $BF_{SilentPauseMean[Group + Depression][Group + Depression + Fatigue]}$ = 509.51). Similarly, adding anxiety scores significantly degraded the model ($BF_{SilenceSpeech[Group + Depression][Group + Depression + Anxiety]}$ = 8.72, $BF_{SilentPauseMean[Group + Depression][Group + Depression + Anxiety]}$ = 509.51). This comparison recommended

omitting Fatigue and Anxiety from the analysis. This suggests that silence in participants' speech did not depend on fatigue and anxiety but on whether they had Cancer-Related Cognitive Impairment and reported depression. However, the Credibility Intervals of the regression coefficient showed only a trend for depression (95% CI [-0.36, -0.00]) (see Appendices A and B).

Regarding the mean duration of filled pauses, the Bayesian model comparison showed that the data were better explained by the model with only intercept. Adding other factors, such as Fatigue, Anxiety, Depression and Group decreased its quality (see Appendix C).

Cofactors relationship

We ran non-parametric rank correlations to test the relationships between cofactors. Fatigue was assessed by using the FACIT-F questionnaire. A lower score indicated a higher report of fatigue. Anxiety and depression were assessed using the HAD scale. Higher scores represent higher levels of anxiety and depression. The Spearman correlation showed strong negative correlations between reported fatigue and reported depression ($r(54) = -0.50$, $p > 1e^{-3}$), and between reported fatigue and reported anxiety ($r(54)= -0.45$, $p > 1e^{-3}$), meaning that the more the patients reported fatigue, the more they were likely to report anxiety and depression. There was a positive relationship between reported fatigue and anxiety ($r(54) = 0.63$, $p > 1e^{-3}$).

Discussion

Incorporating 14 preselected speech features in two experiments using machine learning models, we found an optimal subset of markers that can identify post-cancer patients with a cognitive complaint from the general population. In the first experiment, we ran a two-class model that compared the speech production of participants with a cognitive complaint to those without any cognitive complaint. The optimal subset of features for this experiment included the mean duration of silent pauses, speech-to-silence ratio, speech rate, and mean duration of sustained vowels along with age. In the second experiment, we classified cancer survivors as those with a cognitive complaint, those without any cognitive complaint, and controls. The three-group model selected an optimal subset of features comprising the mean duration of silent pauses, speech-to-silence ratio, and the mean duration of filled pauses.

The two models obtained similar accuracy, even though the two-class model was slightly more efficient than the three-group model. The latter failed more often to provide a correct prediction for the cancer group without complaint in the second experiment. Eight out of eleven participants were incorrectly predicted compared to six out of eleven participants in the first experiment. Two of them were classified as controls, despite reporting language difficulties such as word-finding problems. A further glance at the dataset indicated that one of them had the maximum score on

the PCI subscale of the FACT-Cog, indicating a total absence of cognitive complaint. This participant did not fill out the other questionnaires; therefore, her FACT-Cog score would have resulted from noncompliance rather than the absence of cognitive complaint. Both the FACT-Cog and models failed to match the complaint reported by the participants. The second participant, mistaken for a control by the three-group model, was correctly predicted by the two-class model. Regarding the cancer participants erroneously classified as having a complaint by the three-group model, five out of six received classifications that matched their report of language difficulties. In other words, the models predicted reported difficulties with language better than the presence of a general cognitive complaint.

From a clinical point of view, the FACT-Cog questionnaire is more relevant for assessing the presence of a cognitive complaint than a report of difficulties elicited by a closed-ended question. This questionnaire has been validated in an oncological population and offers good sensitivity and specificity through its multidimensional structure (Costa et al., 2018). Since it has been designed to assess the presence of a cognitive complaint in general, each of its dimensions cannot be interpreted independently. In contrast, the questions asked of participants in this study targeted language, which may explain why speech analysis matches the reported linguistic difficulties better than a general cognitive complaint.

Difficulty with word-finding is a very common cognitive problem encountered by cancer patients (Downie et al., 2006; Skoogh et al., 2012), and may have a direct impact on speech. Interestingly, two silent features were shared by both the machine learning models. We hypothesized that silence may be more reflective of lexical access deficits than other speech features. Generally, the presence of long silent pauses in discourse production could be a sign of cognitive difficulties in cancer survivors.

In addition to cognitive impairment, other factors can also influence speech production. For instance, fatigue, a well-known side effect of cancer (Hofman et al., 2007), possibly slows down speech rate by increasing pause duration. Reported fatigue appeared to be significantly higher for cancer survivors with a cognitive complaint than for other groups. Although it may impact cognition, no unidirectional causal relationship between this physiological factor and speech variables can be asserted. That is, these covariables mutually affect each other. Cancer-related fatigue is a long-lasting symptom that should be considered in the cognitive assessment of cancer patients to be more reflective of their abilities and impairment.

Depression is also a common confounding factor in CRCI. A study conducted on subjective CRCI showed for instance an association between self-reported fatigue and self-reported depression and cognitive difficulties (Veal et al., 2023). In line with these findings, the current study supports the idea that these cofactors could be strongly intertwined. The objective evaluation of CRCI should therefore consider psychological factors rather than isolate them. With this in mind, we claim that speech analysis provides a holistic approach that better grasp the patients' cognitive status in their everyday life. However, our results did not provide any evidence for the reported depression symptoms to affect speech variables in breast cancer survivors. First, we assessed level of reported depression

using a self-administered questionnaire. Although scores significantly differed between groups, they remained below clinical threshold alert. This questionnaire did not diagnose depressive symptoms in our population. Second, the Bayesian models showed only a trend for reported depression scores to influence speech features (95%CI [-0.36, -0.00]), along with a group effect. In other words, reported depression scores alone are not sufficient to explain speech features since reported cognitive impairment also significantly affected speech in cancer survivors.

Clinical implications

This study shows the interest in speech analysis for identifying reported cognitive impairment in cancer survivors, as this method offers several benefits. First, the data are easy and fast to obtain with minimal equipment. In France, the Working Group on Cognition in Oncology recommends a comprehensive battery to diagnose CRCI (Taillia et al., 2015). However, the large number of tests makes the evaluation time-consuming and tiring for patients who are likely to experience fatigue. In contrast, picture-based storytelling requires no more than a few minutes. Second, speech recording is not invasive. Patients with cancer undergo long evaluations throughout the treatment phase, which often involve heavy medical imaging, such as scanning or MRI. Speech production avoids the reminiscence of stressful times, which can induce changes in cognition (Andreotti et al., 2015; Hermelink et al., 2015). Finally, language is an everyday ecological tool that can complement the FACT-Cog questionnaire. Speech analysis provides implicit information while patients are unaware of the outcome measure (i.e., speech features).

Limitations and perspectives

First, this study is limited by the small number of participants. For this reason, we employed feature selection and a cross-validation scheme to minimize overfitting in the machine learning analyses. Although our results demonstrate the ability of speech features to identify Cancer-Related Cognitive Impairment, this finding should be replicated on a larger and more diverse cohort, while potentially including a larger number of speech features. Second, given the small sample size, we did not control for treatment. However, studies have shown that treatment type and duration influence the severity of CRCI. While the current study brings proof of concept for speech analysis as a diagnostic tool for CRCI, subsequent studies should include larger cohorts to distinguish the impact of treatment on cancer survivors' speech. This will help to identify at-risk profiles of developing long-lasting CRCI. Finally, this study focused only on patients with breast cancer. It is nevertheless likely that Cancer-Related Cognitive Impairment in other cancer types can also be captured through speech features, which might vary depending on the cancer type. This method proves efficient, and future studies should examine other cancer types.

Supplemental Material

See the supplementary material at https://osf.io/e6289/files/osfstorage/67825dd9cb25e08486428e46 for the comprehensive results of our statistical analyses. The appendices report the Bayesian model comparison, regression coefficients, and summary of posterior distribution for $\text{SpeechSilence}_{\text{ratio}}$ and $\text{SilentPause}_{\text{mean}}$, as well as the Bayesian model comparison for $\text{FilledPause}_{\text{mean}}$.

Authorship

Amélie B. Richard: Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – original draft, Writing – review and editing, Visualization, Project administration, Funding acquisition. **Alexandre Foncelle:** Formal analysis, Data Curation, Writing – review and editing. **Sophie Jacquin-Courtois:** Resources, Supervision. **Pauline Mouchès:** Methodology, Formal analysis, Data curation, Writing – review and editing, Supervision.

Acknowledgments

ABR was financed by a Ph.D. grant awarded by Paul-Valery Montpellier 3 University and funded by the French Ministry of Higher Education (Ministère de l'Enseignement Supérieur et de la Recherche — MESRI).

Data availability Statement

The data supporting the findings of this study are available from the Comité de Protection des Personnes Est-III. Restrictions apply to the availability of these data, which were used under the license for this study. Data are available from the authors upon reasonable request and with permission from the Comité de Protection des Personnes Est-III.

Comprendre l'anomie en contexte de CRCI grâce à l'oculométrie

1 Contribution de l'article à la thèse

Cet article répond au besoin de comprendre les mécanismes cognitifs impliqués dans le CRCI. Plus précisément, il se penche sur les processus sollicités lors d'un phénomène de manque du mot. Une technique d'oculométrie a permis de tester trois hypothèses quant aux mécanismes.

2 Résumé de l'article

2.1 Introduction

En 2022, l'Agence internationale de recherche sur le cancer a rapporté 2,26 millions de nouveaux cas de cancer du sein diagnostiqués dans le monde (Ferlay et al., 2024). Grâce à l'amélioration des soins, certains types de cancer, tel que le cancer du sein, présentent des chances de survie de plus en plus importantes. La qualité de vie des patients après les traitements fait désormais l'objet d'une attention particulière. Un de ces objectifs vise notamment à diminuer les effets secondaires du cancer et de ses traitements. Parmi eux, l'apparition de modifications cognitives gênantes, ou *Cancer-Related Cognitive Impairment* (CRCI), est une séquelle de cancer très répandue. Le CRCI fait référence à un éventail de difficultés cognitives rapportées par les patients en oncologie tout au long de leur parcours de soin et dans l'après-cancer. Ces difficultés recouvrent diverses composantes cognitives telles que la mémoire, l'attention, les fonctions exécutives et le langage. L'un des symptômes les plus répandus concerne le manque du mot, ou anomie. Pourtant, les outils psychométriques actuels parviennent difficilement à détecter l'anomie chez les survivants d'un cancer. L'absence d'outil adapté limite la compréhension des mécanismes cognitifs sous-jacents au manque du mot en contexte de CRCI. Autrement dit, les tests actuels ne permettent pas de déterminer si l'anomie rapportée par les patients est liée à une dégradation de la mémoire sémantique, à un déficit du contrôle exécutif ou à un amoindrissement des capacités attentionnelles.

La recherche lexicale modifie le flux de la parole continue. La production d'un discours est rarement linéaire dans la mesure où elle est souvent ponctuée de pauses. La récupération laborieuse

d'un mot, de son concept ou des phonèmes qui le compose, s'accompagne généralement d'une ou plusieurs interruptions de parole. À ce titre, l'analyse de la parole constitue une méthode sensible pour repérer et comprendre les difficultés de recherche lexicale.

Des études ont investigué la relation entre le langage et l'attention visuelle *via* l'analyse du suivi oculaire. Ces travaux suggèrent que la parole influence la direction du regard (Huettig & Altmann, 2005). En effet, les locuteurs anticiperaient les mots qu'ils prévoient de dire en fixant les éléments visuels qui se rapportent à ces mots (Griffin & Bock, 2000). Cette théorie « basée sur la cognition » entre en concurrence avec une autre théorie qui stipule que l'attention visuelle est guidée par des processus cognitifs de bas-niveau (c.-à-d. sensoriels-perceptifs). En d'autres termes, le regard serait attiré par des éléments visuels proéminents, tels qu'un contraste de couleur.

Pour tester ces deux théories, J. M. Henderson et Hayes (2017) et J. M. Henderson et al. (2018) ont mis au point une technique de cartes de chaleur, l'une, appelée carte « sémantique », générée à partir des éléments porteurs de sens (p. ex., un objet), et l'autre conçue à partir des éléments saillants de la scène visuelle (p. ex. une couleur vive), dite une carte de « saillance ». Cette méthode compare le niveau de correspondance entre les points de fixation sur une scène visuelle et sa carte sémantique au niveau de correspondance entre ces mêmes points de fixation et la carte de saillance de cette même scène. Selon ces travaux, la carte sémantique prédirait les points de fixation plus précisément que la carte de saillance. Appliquée à une population oncologique, cette technique pourrait renseigner quant aux difficultés de recherche lexicale rencontrées par les survivants d'un cancer.

2.2 Objectifs

L'objectif de cette étude est de tester trois hypothèses concernant l'origine du manque du mot chez des survivantes d'un cancer du sein : (i) une dégradation des connaissances sémantiques, (ii) un déficit de contrôle exécutif empêchant d'accès à la forme phonologique des mots, et (iii) des difficultés attentionnelles impliquées lors de la production de mots.

Selon Reilly et al. (2020), l'anomie résultant d'une atteinte de la mémoire se traduirait par une augmentation du nombre de fixations. En effet, les patients présentant une dégradation du système sémantique nécessitent plusieurs fixations avant de faire correspondre les traits d'un élément visuel à son concept. Pour confirmer notre première hypothèse, les survivantes d'un cancer du sein devraient produire plus de fixations sur un même élément que les sujets contrôles.

En ce qui concerne notre deuxième hypothèse, nous nous attendons à ce que le délai entre la première visualisation d'un élément et sa dénomination (*Eye-Voice Span*, EVS) soit plus long pour les survivants que pour les sujets témoins. Ce délai témoignera des difficultés d'accès aux phonèmes des mots recherchés.

Quant à notre dernière hypothèse, nous pensons que le regard des survivantes sera davantage

guidé par des éléments visuellement proéminents que par les éléments nécessaires à la structure de l'histoire (p. e.x, « palmier » qui est un élément informatif). Cette hypothèse est testée en comparant la correspondance entre les points de fixation des participantes et la carte de saillance, ainsi que la correspondance entre ces points et la carte de sémantique. La fixation excessive d'éléments proéminents pourrait témoigner d'un déficit visuo-attentionnel.

2.3 Méthode

Huit survivantes d'un cancer du sein et 13 femmes contrôles saines ont pris part à cette étude. Les critères d'inclusion et d'exclusion étaient les mêmes que ceux de l'étude précédente, « Dévoiler les marqueurs prosodiques du CRCI ». En résumé, les participantes ayant eu un cancer du sein (groupe *Breast Cancer*, BC) avaient toutes terminé la phase curative de leur traitement (c.-à-d. intervention chirurgicale, chimiothérapie et/ou une radiothérapie) depuis quatre à 52 semaines. Les participantes contrôles (groupe *Healthy Control*, HC) présentaient les mêmes critères d'éligibilité, excepté qu'elles devaient ne jamais avoir eu de cancer. L'ensemble des participantes a complété le questionnaire FACT-Cog pour déceler la présence d'une plainte cognitive, un test de dépistage cognitif (MoCA), un test de vitesse de traitement (SDMT), un questionnaire concernant l'anxiété et la dépression (HADS) et la fatigue (FACIT-F).

La procédure de cette étude suit la méthode décrite dans le chapitre Méthodologie. Lors de l'expérience, les participantes ont eu pour consigne de raconter une histoire à partir d'une séquence de cinq vignettes. Leur production orale a été enregistrée et transcrite de manière semi-automatique. Les pauses précédant immédiatement un item lexical de classe ouverte (c.-à.-d. nom, verbe, adverbe et adjectif) ont été annotées.

Le comportement oculomoteur des participantes, lors de la tâche de narration, a été enregistré par une technique d'oculométrie. Des régions d'intérêt (*Areas of Interest*, AoI) ont été définies pour délimiter les éléments visuels porteurs de sens (p. ex., « casquette », par contraste avec un trait ou une couleur vive). Les données de suivi oculaire ont été temporellement alignées sur les données vocales, dans le but de repérer les AoI correspondant aux items lexicaux de classe ouverte précédés d'une pause. Ces mots ont été considérés comme faisant l'objet d'une recherche lexicale laborieuse. Le délai entre l'apparition du premier regard sur un élément sémantique et le début de la dénomination de son concept, c'est-à-dire l'*Eye-Voice Span* (EVS), ainsi que le nombre de visites de ces éléments a été mesuré.

Des cartes de chaleur retraçant les fixations oculaires des participantes ont été générées à partir de scripts Python. Les fixations ont été définies comme une séquence de points sur une zone où la fovéa s'est stabilisée pendant au moins 100 ms.

Une carte de chaleur affichant les éléments visuellement proéminents de l'image globale (c.-à-d. l'image incluant les cinq vignettes) présentée à l'écran a été crée à partir de script Python.

Une carte de chaleur représentant les éléments porteurs de sens au sein de l'image a été conçue en suivant la procédure décrite par J. M. Henderson et Hayes (2017). Deux cent vingt cercles d'un rayon de 75 px, dont un chevauchement de 50 px entre chaque cercle, ont été extraits de l'image. Les cercles ont été présentés de manière aléatoire à des volontaires qui ont évalué le niveau d'information de chaque cercle à partir d'une échelle de Likert à 6 points (de « Très faible » à « Très élevé »). Les scores moyens de chaque cercle ont ensuite été calculés et pondérés afin d'augmenter les contrastes entre les cercles peu informatifs et les cercles très informatifs. Enfin, une carte sémantique (*meaning map*) a été générée à partir de ces notations à l'aide d'un script Python.

Les cartes de fixation des participants et la carte sémantique, ainsi que les cartes de fixation et la carte de saillance, ont été superposées pour calculer le nombre de pixels correspondant dans chaque paire de cartes. En d'autres termes, la valeur d'intensité d'un pixel dans une carte de fixation a été soustraite de la valeur du pixel correspondant, en termes de coordonnées spatiales, dans les cartes de saillance et de sémantique. Deux analyses ont été réalisées au moyen de cette méthode : l'une axée sur le chevauchement spatial et l'autre prenant en compte le temps passé sur chaque zone des cartes.

Les analyses statistiques ont essentiellement comporté des tests U de Mann-Whitney pour comparer les variables démographiques, prosodiques et oculomotrices des deux groupes.

2.4 Principaux résultats

La durée de l'EVS était significativement plus longue pour le groupe de survivantes (N = 8, M = 2.24 s) que pour le groupe contrôle (N = 13, M = 1.65 s) ; U = 3344.5, p = .002. Ce résultat suggère que les survivantes du cancer ont mis plus de temps que les sujets témoins à nommer les objets, depuis le premier regard avec l'intention de parler jusqu'à l'articulation du mot. Une analyse qualitative n'a pas révélé de différence entre les groupes concernant les mots touchés par la recherche lexicale.

Les fixations n'ont pas été plus nombreuses pour les survivantes d'un cancer que pour les contrôles ; U = 54, p = .92. Ce résultat suggère que ces participantes n'ont pas eu davantage besoin de regarder les éléments visuels porteurs de sens pour récupérer leur concept.

De manière générale, les fixations de l'ensemble des participantes ont davantage correspondu à la carte sémantique qu'à la carte de saillance. L'analyse temporelle des cartes de fixation et de la carte sémantique n'a pas révélé de différence significative entre les survivantes du cancer (BC) et les témoins sains (HC) ; U = 39, p = .373. De même, aucune différence significative n'a été observée entre les groupes concernant le pourcentage de pixels communs aux cartes de fixation et la carte de saillance ; U = 49, p = .860. L'analyse spatiale des cartes de fixation et de la carte sémantique n'a pas non montré de différence significative entre les groupes ; U = 45, p = .645, tout comme l'analyse spatiale des cartes de fixation et la carte de saillance U = 52, p = 1.00.

2.5 Discussion

Cette étude explore le manque du mot rencontré par des survivantes du cancer du sein L'ajout d'une technique d'oculométrie à l'analyse de la parole permet de mieux appréhender les mécanismes cognitifs sous-jacents à l'anomie. Comparé au groupe contrôle, les participantes ayant eu un cancer n'ont pas nécessité davantage d'exploration des éléments informatifs pour retrouver les concepts auxquels ils sont associés. Cependant, le délai de dénomination entre la première visualisation d'un item et sa dénomination était significativement plus long pour les survivantes d'un cancer du sein. Des recherches menées sur la vitesse de dénomination ont montré que les yeux se dirigent vers la cible à dénommer moins de 900 ms avant sa production verbale (Griffin, 2004 ; Griffin & Bock, 2000). Dans notre étude, seuls les mots immédiatement précédés de pauses ont été analysés. Ce choix s'explique par la volonté de sélectionner uniquement les mots faisant l'objet d'une recherche lexicale laborieuse. Il n'est donc pas étonnant de constater que le délai moyen des participantes a excédé le seuil généralement observé. Examinés ensemble, ces résultats semblent écarter la piste d'un manque du mot d'origine mnésique. En revanche, ils penchent en faveur de l'hypothèse d'un déficit exécutif sous-jacent, lequel serait caractérisé par la latence d'accès au lexique. Ces observations sont soutenues par une revue systématique mettant en avant la diminution du contrôle exécutif chez les survivants d'un cancer (Yao et al., 2017).

Notre étude a également étudié la possibilité d'un déficit attentionnel pouvant conduire à un manque du mot chez les participantes ayant eu un cancer du sein. Pour explorer cette piste, les fixations oculaires des participantes ont été comparées à des zones d'intérêt sémantique et à des éléments visuellement saillants. Il était attendu que le regard des survivantes soit plus attiré par les éléments saillants que celui des sujets contrôles. Contrairement à notre hypothèse, aucune différence significative n'a été observée entre les groupes. Les deux groupes ont en priorité fixé les éléments porteurs de sens. Pour autant, la piste d'un déficit attentionnel entraînant un manque du mot ne peut être rejetée. En effet, de plus en plus d'études sur le CRCI mettent en lumière l'implication des processus attentionnels dans les changements cognitifs liés au cancer. En outre, nous n'avons pas contrôlé le niveau de détails des vignettes présentées aux participantes. Il est possible que ces vignettes n'aient pas contenu suffisamment d'éléments sémantiques et visuellement complexes pour être suffisamment sensibles à un défaut subtil de l'attention. D'autres recherches sont requises pour mieux comprendre l'influence des processus cognitifs sur le délai de dénomination chez des survivants d'un cancer.

3 Apport de l'article à la recherche

Cette étude met en évidence les avantages des tâches écologiques et des méthodes d'investigation non invasives, comme l'oculométrie, pour explorer les processus cognitifs sous-jacents à la production de parole chez les survivants de cancer présentant des plaintes cognitives. Elle offre un éclairage sur

les processus cognitifs liés aux difficultés de dénomination chez les survivantes du cancer du sein. En particulier, à la lumière des résultats, il semble que ces difficultés de dénomination ne soient pas dues à un déficit mnésique, mais plutôt à un problème d'accès au lexique, qui serait sous-tendu par un déficit de contrôle exécutif. Ces résultats ouvrent de nouvelles perspectives thérapeutiques, puisque que la rééducation de l'anomie dépend des processus cognitifs dont elle découle.

Oculomotor behavior of breast cancer survivors reporting cognitive changes

Amélie B. Richard[a,b], Laure Pisella[b], Sophie Jacquin-Courtois[b,c], Maude Beaudoin[b]

[a] Université Paul-Valéry Montpellier 3, CNRS UMR5267; Praxiling, 34090 Montpellier, France [b] Université de Lyon, Université Lyon 1, INSERM U1028; CNRS UMR5292; Lyon Neuroscience Research Centre, Trajectoires Team, F-69676 Lyon, France [c] Physical Medicine and Rehabilitation Department, Henry-Gabrielle Hospital, Hospices Civils de Lyon, 69610 Pierre-Bénite, France

Correspondence to Amélie B. Richard: amelie.richard@inserm.fr.

Disclosure: The authors declare no competing financial or nonfinancial interests.

ABSTRACT

Purpose: Word-finding difficulties are very common in cancer survivors. This symptom can affect survivors' social interactions and considerably decrease their quality of life. However, this cognitive problem is rarely detected by verbal fluency tasks. Oculomotor behavior is a useful technique for investigating linguistic processes in connected speech. In this study, we tested whether word-finding difficulties in cancer survivors were caused by memory deficits, executive deficits, or attention deficits. **Method:** Eight women who had completed breast cancer treatments and 13 controls were asked to narrate a picture-based story. Speech production and oculomotor behavior were recorded and analyzed using semi-automated methods. Eye-Voice Span (i.e., the delay between the first gaze at a meaningful element and its naming) was measured for words preceded by a pause. The number of fixations was also measured. Fixation maps of the participants were overlapped with a saliency map (i.e., prominent features such as color contrast) and a meaning map (i.e., semantic elements such as "sailboard") for a pixel-matching analysis. **Result:** EVS duration was significantly longer for breast cancer survivors than for controls. This result suggests that survivors had more difficulties to retrieve their words than the controls. Survivors did not fixate more semantic items than the controls. Pixel-matching analyses revealed no difference between the way participants explored the maps. Generally, all participants gazed more at the semantic elements than at the saliency features. **Conclusions:** The analysis of oculomotor behavior in cancer survivors during a laborious lexical search indicates that their word-finding difficulties are caused by a deficit in lexical access rather than a memory deficit. More research is needed to investigate the attentional processes involved in the

speech production of cancer survivors.

Keywords: eye-tracking, lexical search, cognitive impairment, breast cancer

---

Main

In 2022, the International Agency for Research on Cancer reported 2.26 million new cases of breast cancer diagnosed worldwide (Ferlay et al., 2024). Despite being a leading cause of death in women, breast cancers lie among tumors that are best treated. The number of breast cancer survivors is increasing, henceforth quality of life after cancer becomes a challenge for public health. Alongside from hair loss and other well-documented side effects, cognitive impairment is a widespread sequela of cancer and its treatments (Hurria et al., 2007). Cancer-Related Cognitive Impairment (CRCI) refers to a set of self-reported cognitive difficulties that substantially impacts daily functioning of survivors without affecting their autonomy (Ahles & Root, 2018). For instance, patients often describe short-term memory loss, concentration problems, and forgetting their words (F. M. Henderson et al., 2019). Failure in naming is the most widespread symptom of CRCI, regardless of the cancer type (see for instance Downie et al. [2006] and Skoogh et al. [2012]). Also called anomia, word-finding difficulties is to the inability to retrieve a word in a given time. This phenomenon may occur several times a day over a long period, in which case it substantially interferes with communication. Depending on the brain region affected, anomia either reflects a failure of executive control to select the right word, or is the consequence of damage to the knowledge on word meaning (Caramazza & Hillis, 1990). Studies also showed the implication of attentional processes in word production (Jongman et al., 2020). The intensity of word-finding difficulties in cancer survivors is well captured in medical interviews (Bolton & Isaacs, 2018) and standardized questionnaires such as the Functional Assessment of Cancer Therapy – Cognitive Function (FACT-Cog v3 [Costa et al., 2018]). Yet current psychometric tools such as naming tasks and verbal fluency tasks fail to fully detect this widespread symptom of CRCI (Costa & Fardell, 2019). The discrepancy between self-reported and objective assessment highlights the need for more evidence about word-finding difficulties in a context of CRCI. In addition to hamper the understanding of the cognitive mechanisms underlying anomia (i.e., degradation of semantic memory, weak executive control or attentional deficit), the lack of tailored psychometric tool hinders cancer survivors from accessing care.

Word-finding difficulties are known to impact the flow of connected speech. Unlike isolated words, connected speech is characterized as a continuous stream of phonemes used to apprehend or deliver a message. Connected speech is rarely linear as it contains disfluencies such as pauses, repairs and repetitions. Especially, the processes of lexical search and semantic retrieval in memory modify the fluidity of language in use, and hesitation pauses often point out word-finding difficulties. However, the analysis of spoken language is a cumbersome task that requires time as well as strong knowledge in linguistics. Furthermore, semantic content, lexical choice and prosody (i.e., intonation,

rhythm and loudness) highly hinge on speakers. The large variability and unpredictability of these variables in speech samples preclude their comparison for generalization. For that reason, most studies on language production focus on single-word tasks or reading task to better control the conditions. Unfortunately, these tasks are far from reflecting daily life verbal ability. Moreover, studies report inconsistent outcomes in the detection of anomia reported by cancer survivors with the use of verbal fluency or naming tasks (see for instance Andryszak et al. [2017] and Janelsins et al. [2022]). Thanks to its high sensitivity, connected speech thus appears to be a better choice to understand the anomia reported by cancer patients.

Eyesight plays a major role in the way human focus their attention. The visual world paradigm states that visual environment and speech comprehension or production influence each other. Cooper (1974) found that listeners tend to gaze at objects resembling the semantic word contained in the spoken message (for instance, a cable when talking about snake [Huettig & Altmann, 2005]). On the production side, eye-tracking allow researchers to study speech generation in real-time. Studies have shown that speakers look at the semantic elements of visual scene in the order of the planned utterance. In other words, the focus of visual attention follows a similar time course that of message generation (Griffin & Bock, 2000). Moreover, speakers keep fixate a visual item during lexical search, and do not move forward until they are able to name it. The tracking of eye movement might therefore be particularly enlightening on lexical access processes. In addition to the visual word paradigm, two competing theories aim at explaining visual attentional guidance in real-world scenes. The image-based theory states that visual attention is guided by saliency features (e.g., luminance contrast), while the cognitive-based theory suggest that visual attention is driven by semantic features (e.g., informative elements). To investigate these theories, Henderson et al. (J. M. Henderson & Hayes, 2017; J. M. Henderson et al., 2018) developed meaning maps of visual scenes. These maps displayed a spectrum of hues corresponding to the level of informativeness based on raters' judgement. The authors then compared meaning and saliency maps to classical attention maps (i.e., heatmaps generated from the fixation point coordinates) to estimate which of the first two maps better matches the latter. They found that the meaning maps predicted the attention maps better than the saliency maps, providing evidence for the cognitive-based guidance theory. In light of this finding, the comparison of meaning versus saliency maps might be of interest to examine visual attentional processes that underlie lexical search in cancer survivors with cognitive impairment.

Previous studies conducted by our team showed that speech features related to anomia such as silence or filled pauses were able to identify cancer patients with cognitive impairment among controls (Richard, Foncelle, et al., under review). To further our investigation, we look into the relationship between these markers and cancer survivors' oculomotor behavior. The main goal of the current study is to shed light on a possible mechanism of anomia in breast cancer survivors who report a cognitive complaint. In other words, we aim to determine whether word-finding difficulties stem from underlying deficits in memory, executive function, or attention.

In the absence of clear lesions in the brain regions dedicated to semantic memory or precise

deficits in memory measured by neuropsychological tests, we hypothesize that anomia in cancer survivors may stem from executive or attentional deficits caused by a dysfunction in executive control or visual attentional difficulties.

According to Reilly et al. (2020), anomia resulting from memory damage would result in increased fixations. In fact, a higher number of fixations may suggest difficulty in mapping visual features to the abstract concept of an object.

Regarding a deficit in executive control, we expect that during a lexical search event, the delay between the first viewing of an item and its naming would be longer for survivors than for control subjects. This hypothesis is assessed by measuring Eye-Voice Span (i.e., the delay between the first gaze at a meaningful element and its naming).

As for the attentional hypothesis, contrary to those of the control subjects, the cancer survivors' visual attention would be more guided by visually salient elements than the elements necessary to story structure (e.g., "palm tree" which is an informative element). This assumption is tested comparing the match between the participants' fixation maps and the saliency heatmap as well as the meaning heatmap.

Methods

Ethical procedures

This research involving human participants will be performed in accordance with the ethical standard of the Declaration of Helsinki and its later amendments. The participants recruited in this study are part of a research project that focuses on cognitive impairment induced by cancer and its treatment in women who have had breast cancer (Discourse, Cancer and Cognition [DisCCo]), and was validated by the French ethics committee Comité de Protection des Personnes Est-III in August 2023 (national number: 2023-A00267-38). All participants provided written consent form to participate.

Population

Ten women post breast cancer and fourteen age- and education-matched healthy controls volunteered for the current study. To be included to the study, volunteers must meet the following criteria: (i) to be a woman aged between 30 and 65, (ii) to be a native French speaker or be fluent in French, (iii) to have normal or corrected sight and hearing. In addition, participants who had cancer must were (iv) have completed surgery, chemotherapy, and radiotherapy more than four weeks and less than one year ago. Any participant who meet the following criteria were not eligible: (i) a history of language disorders, (ii) a neurological or psychiatric history, (iii) the use of medications likely to cause cognitive changes.

Additionally, results of post-cancer participants (iv) presenting metastases were removed from the analyses. Participants declared no eyesight pathology. One post-cancer volunteer was not eligible due to the presence of metastases, and one control was not included for dyslexia.

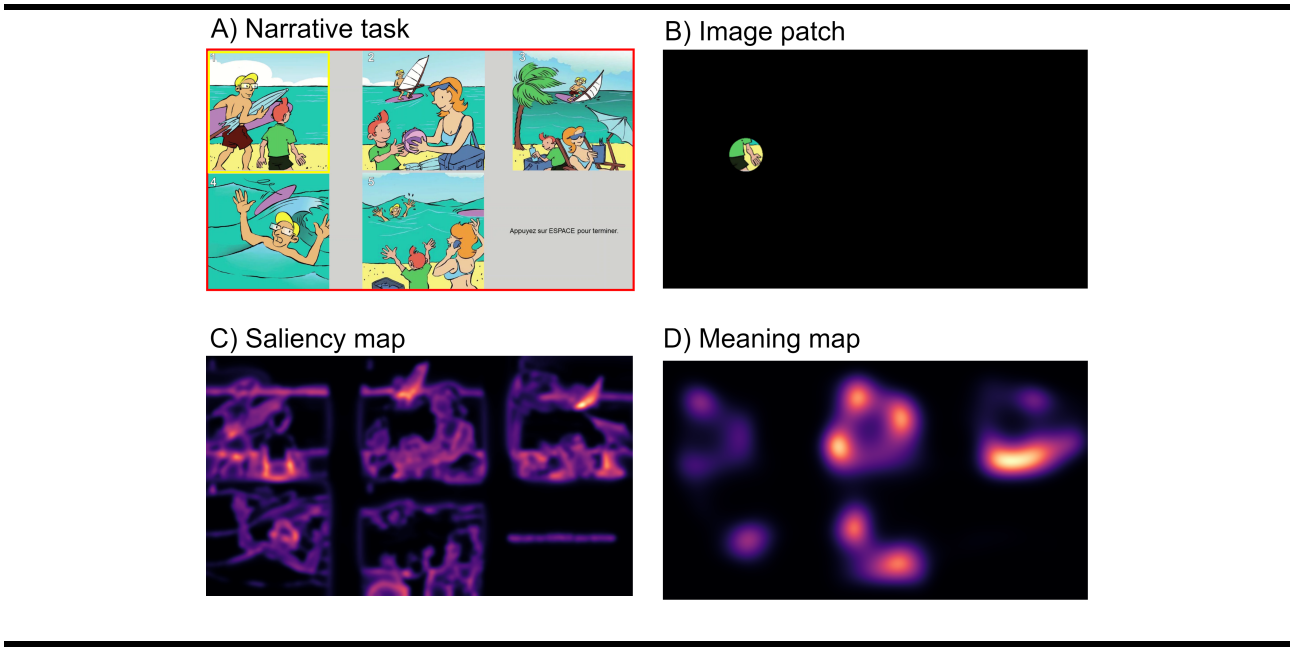
All participants filled the French version of the FACT-Cog v3 questionnaire (Joly et al., 2012). This questionnaire has been designed to detect the presence of a cognitive complaint in cancer survivors. A threshold of 55/72 or lower on the Perceived Cognitive Impairment (PCI) FACT-Cog subscale indicates the presence of a cognitive complaint (Van Dyk et al., 2020). To be included in the analyses, post-cancer participants must have a score below 55/72, while controls must have a score above 55/72. All participants meet these criteria but one post-cancer participant. Eventually, eight post-cancer participants (BC = 8) and eight age-matched manually picked up controls (HC = 8) were included in this study.

In addition, participants underwent a neuropsychological and psychological evaluation to control for cognitive performance and confounding factors. The Montreal Cognitive Assessment (MoCA) test (Nasreddine et al., 2005) was administered for screening cognition. A score below 26/30 indicates the presence of cognitive decline. The Symbol Digit Modalities Test (SDMT) (Strober et al., 2020) was used to measured processing speed. A standard deviation above 1.5 was deemed abnormal. Participants also filled the Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-F) questionnaire (Cella et al., 1993), designed to estimate fatigue in cancer patients. The higher the score, the greater the level of fatigue. There is no standardized threshold for this questionnaire. Participants also completed the Hospital Anxiety Depression Scale (HADS) (Zigmond & Snaith, 1983), which comprises two subscales designed to assess anxiety (HAD-A) and depression (HAD-D). A score above 11/21 on either subscale indicates the presence of anxiety or depression.

Task

Participants received written instruction to narrate the picture-based story (Figure 1A) that will appear on the screen, with no time limit (Bézy et al., 2016). The task was preceded by a 250 ms fixation cross displayed in the center of the screen. Both speech production and oculomotor behavior were recorded during the task.

Figure 1. A) Picture-based narrative task. Red square: whole image. Yellow square: picture 1. B) Example of a scene patch extracted from the whole image. C) Saliency map. D) Meaning map.



The experiment was conducted in a soundproof booth using a Tobii Pro Fusion eye-tracking bar of a 120 Hz resolution. The bar was mounted 7 cm below a 22 inches screen on a support made for the experiment. Participants were seated 70 cm from the screen. The task was displayed in 1920 x 1080 px format in full screen mode. In addition, speech production was recorded using a Shure® omnidirectional lapel microphone placed 15 cm below the participants' mouth and linked to an M-Audio interface. Eye-tracking data were pre-processed using Tobii Pro Lab software (Tobii AB, 2023) before extraction.

Speech analysis

Speech recordings were automatically transcribed using Whisper (Radford et al., 2022) and transcripts were manually checked. Speech phenomena interpreted as the speaker hesitation were annotated using SPPAS (Bigi, 2015) and Praat (Boersma & Weenink, 2024). Three hesitation phenomena were retained for this study: silent pauses (i.e., the absence of speech, cough, or yawn during 200 ms or more), filled pauses (i.e., non-semantic voiced segments such as "uh" and "uhm" produced for 200 ms or more with a flat fundamental frequency contour) and sustained vowels (i.e., sustained vowels produced for 200 ms flat fundamental frequency contour). Only hesitation phenomena preceding lexical items (i.e., open-class content words such as nouns, verbs, adjectives and adverbs) were retained for this study. These phenomena are assumed to reflect the lexical search processes.

Eye-Voice Span measure

Area of Interests (henceforth AoI) were defined for semantically relevant visual items such as "cap, "kid face" or "sunshade". Eye-tracking data were then converted to a `.TextGrid` format and time-aligned with speech data using a Python script. This allowed us to search for AoIs that temporally correspond to the more or less successfully retrieved word. Using ELAN (Max Planck Institute for Psycholinguistics, 2024), we calculated the delay between the onset of the first gaze at a meaningful element and the onset naming of its concept (for instance, the drawing of a sailboard can be referred to as "sailboard", "nautical activity" or "surf"), that is the Eye-Voice Span (henceforth EVS), along with the number of visits of these elements.

Oculomotor behavior

Eye-tracking data were extracted and were analyzed using Python packages for visualization. All scripts were developed for the purpose of this study using libraries OpenCV (Bradski, 2000), SciPy (Virtanen et al., 2020), PIL (Umesh, 2012) and scikit-image (Van Der Walt et al., 2014).

Fixation were defined as the sequence of points on an area where the fovea has stabilized during at least 100 ms. This threshold was deemed to be the minimum time required to extract and process visual information (Blignaut, 2009). We computed fixation maps from the coordinates of fixation points for each participant using a Python script.

A saliency map was generated from the original image used in the narrative task using a Python script (Figure 1C). The OpenCV's saliency detection extract prominent regions of an image based on the contrast between adjacent pixels.

A meaning map was created following the method develop by Henderson et al. (J. M. Henderson & Hayes, 2017). Two hundred and twenty scene patches with a radius of 75 px and an overlap of 50 px between each snippet were extracted from the original image (Figure 1B). Patches were randomly presented to 86 volunteers (62% female, median age = 37.5) who were asked to rate the level of information of each patch using a 6-point Likert scale (the French equivalent for "Very low", "Low", "Somewhat low", "Somewhat high", "High", "Very high"). Only ten volunteers completed the full survey. However, participation rate was around 20% for each patch, in such a way that each patch has been assessed by in average 15 raters. Mean scores were then calculated and weighted for each patch in a 2-base logic (i.e., from Very low = 0, to Very high = 32) to increases contrasts between low informative patches and very informative patches. Eventually, a meaning map was generated from these rating using a Python script (Figure 1D).

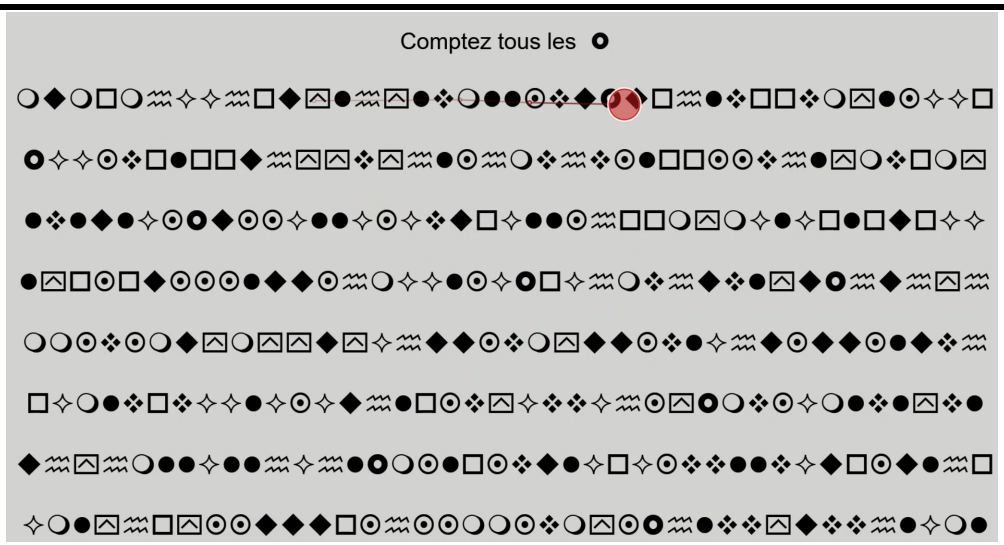
The participant fixation maps and the meaning map as well as the fixation maps and the saliency map were overlaid to calculate the number of pixels that matched in each pair of maps. In other words, the intensity value of a pixel in a fixation map according to its coordinates was subtracted from the

value of the corresponding pixel in the saliency and semantics maps. Two analyses were carried out using this method: one that focused on spatial overlap, and one that considered the time spend on each area of the maps.

Non-verbal control task

A simple non-verbal task was added to check whether speech production affected visual patterns in the BC participants. Participants were instructed to count the number of visual targets among distractors displayed in lines (Figure 2). The task was displayed during 40 s on screen before disappearing.

Figure 2. Non-verbal control task. Red dot: a fixation point recorded with the eye-tracking system.



Statistical analysis

Considering the exploratory nature of this study, the statistical analysis mostly consisted in descriptive statistics. The small group size led to report median values instead of means to better grasp the population characteristics. Likewise, Mann-Whitney U (Mann & Whitney, 1947) tests for independent non-parametric data was run to compare the demographic, speech and oculomotor variables across groups.

Results

Participants characteristics

Table 1 summarizes the population characteristics. Several Mann-Whitney U tests were performed to compare demographic and psychological variables in BCs and HCs. There was not a significant difference in age between the BC group and the HC group; U = 29.5, p = .109. There was not a significant difference between groups in cognitive performance assessed using the MoCA either; U = 63, p = .442, and the SDMT; U = 64.5, p = .379. Post-cancer participants reported significantly higher levels of anxiety; U = 64.5, p = .006, depression; U = 14, p = .005, and fatigue; U = 98.5, p < .001. However, their scores were below alert thresholds and are not deemed clinically relevant.

Table 1. Age, cognitive and psychological characteristics of the study population.

Variables	Healthy Control group Median (SD)	Breast Cancer group Median (SD)	*p*-value
N =	13	8	—
Age	42 (10.42)	50 (9.20)	.109
FACT-Cog PCI (/72)	68 (4.84)	28 (11.96)	< .001
MoCA (/30)	28 (2.39)	27.5 (2.18)	.442
SMDT (SD > 1.5)	0 (1)	-.025 (.8)	.378
FACIT-F (/52)	47 (7.13)	24 (9.72)	$1e^{-3}$ *
HAD-A (/21)	5 (4.19)	13.5 (4.45)	.006
HAD-D /21	2 (1.67)	6 (3.06)	.005

Lower oculomotor data quality in cancer survivors

Tobii Lab Pro only measures fixations and saccades if the x and y coordinates are available. Otherwise, fixations and saccades are considered incomplete (referred as partial fixations and saccades), and are removed from the visualizations. The absence of coordinates can occur for several reasons: 1. A poor calibration, 2. Participants looked away from the screen. A glance at the data validity revealed that BC participants produced more partial fixations than the HC participants. Subsequently, some data points are missing from the analyses. Calibration results were however correct. A possible explanation is that the BC participants' blinks were relatively frequent and long or participants often looked away from the screen.

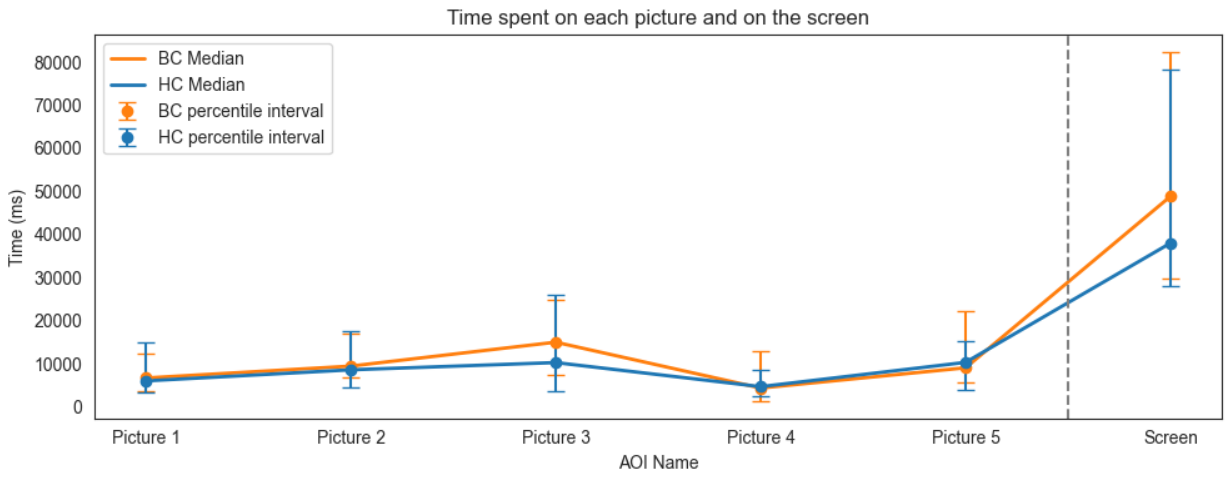
A Mann-Whitney U test was performed to compare time spend looking away from the screen in BCs and HCs. There was a trend for BC participants (M = 15.6 s, SD = 20.4 s) to look away from screen longer than HC participants (M = 11.0 s, SD = 48.1 s). However, the difference did not reach significance; U = 25, p = .054.

Visual strategy for exploring a scene

Median time to complete the speech task was higher for post-cancer participants (M = 60.91 s, SD = 20.16) than for control participants (M = 50.53 s, SD = 35.78 s). However, a Mann-Whitney test provided no evidence for a significant difference between groups; U = 42, p = .50.

Figure 3 displays the time participants spent to explore each picture and more generally on the screen. Two Mann-Whitney U tests were performed to compare the time spent on the pictures and on the screen in BCs and HCs. There was not a significant difference in the time spent on the pictures; U = 47, p = .75, and on the screen in BCs and HCs; U = 42, p = .50. From a qualitative perspective, BC participants spend more time on the screen than the HC participants. This suggests that cancer survivors looked more often outside the pictures (i.e., the grey background) than the controls.

Figure 3. Comparison of the duration BC participants spent on pictures and on the whole image to that of the HC group (medians and percentile interval error bars). The difference did not reach significance.



Finally, a Mann-Whitney U test was performed to compare the number of fixations on AoIs in BC group and HC group. There was not a significant difference in number of fixations on AoIs in BC group and HC group; U = 54, p = .92.

Non-verbal visual search task

Regarding the non-verbal task, all participants adopted a left-to-right search strategy. The time that participants took to reach the end of the symbol series was measured. A Mann-Whitney U test was performed to compare the time necessary to complete the non-verbal task in BCs and HCs. There was not a significant difference in percentage of matched pixels in BCs and HCs; U = 36.5, p = .274.

A descriptive analysis showed that 4/8 BC participants did not complete the series versus 2/13 HC participants. HCs were in average 3 second faster than BCs, even though two BCs performed more quickly than the HC group's mean duration. The number of fixations returned to the target was the same across all groups. Eventually, we noticed an effect of age on duration, both BC and HC older participants taking longer than younger ones to proceed.

Evidence for a meaning-based guidance of attention

Temporal fixation maps and spatial fixation maps of participants were overlaid on the meaning map and the saliency map. The percentage of matched pixels (excluding black pixels) between maps was calculated. Table 2 summarizes the results of the pixel overlap analysis. Generally, fixation points matched more the meaning map than the saliency map for participants of both groups.

Table 2. Summary of the overlap comparison.

% of overlapping	Healthy Control group Median (SD)	Breast Cancer group Median (SD)	*p*-value
Temporal fixation maps - Meaning map	9.03 (2.14)	10.55 (2.61)	.374
Temporal fixation maps - Saliency map	3.21 (1.74)	3.45 (1.15)	.860
Spatial fixation maps - Meaning map	26.4 (5.07)	27.52 (6.88)	.645
Spatial fixation maps - Saliency map	12.53 (3.68)	11.7 (2.15)	1.00

A Mann-Whitney U test was performed to compare the percentage of matched pixels in BCs and HCs the temporal fixation maps. There was not a significant difference in percentage of matched pixels in temporal fixation-meaning overlap in BCs and HCs; U = 39, p = .373. There was not a significant difference in percentage of matched pixels in temporal fixation-saliency overlap in BCs and HCs; U = 49, p = .860.

A Mann-Whitney U test was performed to compare the percentage of matched pixels in the spatial fixation maps in BCs and HCs. There was not a significant difference in percentage of matched pixels in spatial fixation-meaning overlap in BCs and HCs; U = 45, p = .645. There was not a significant difference in percentage of matched pixels in spatial fixation-saliency overlap in BCs and HCs; U = 52, p = 1.00.

Figure 4 shows a temporal fixation map for a typical control participant which was overlaid on the meaning map and the saliency map.

Figure 4. Example of the time overlaid results for a typical control participant. The figures show the overlapping pixels in the fixation map and the meaning map (A), as well as in the fixation map and the saliency map (B).

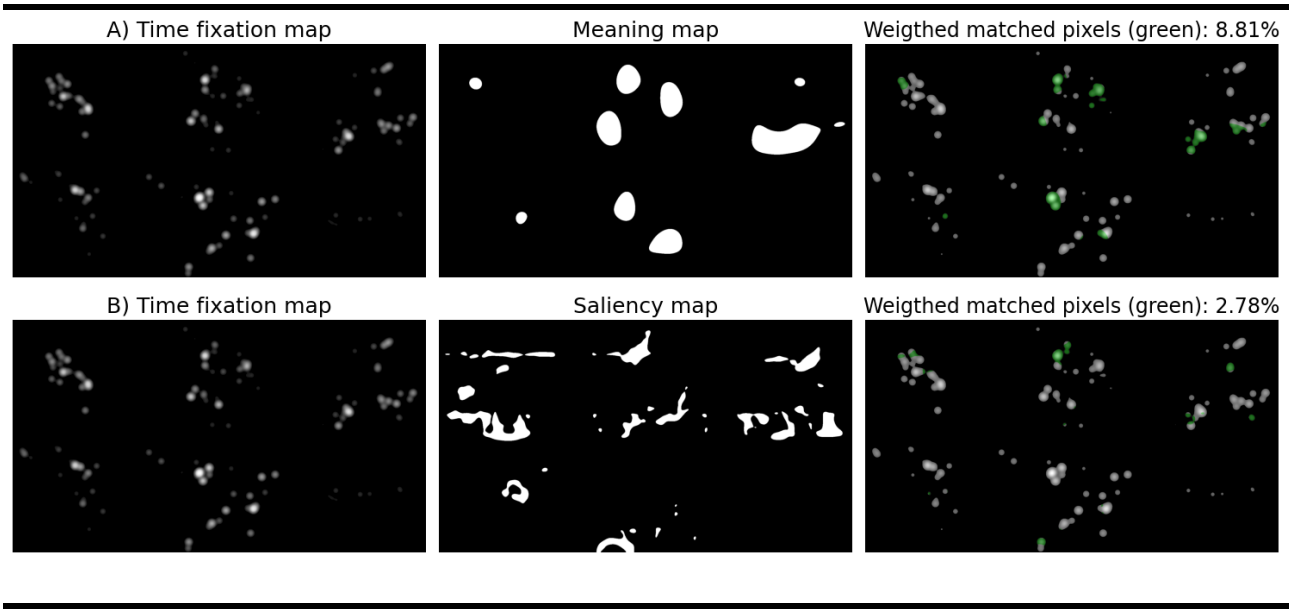
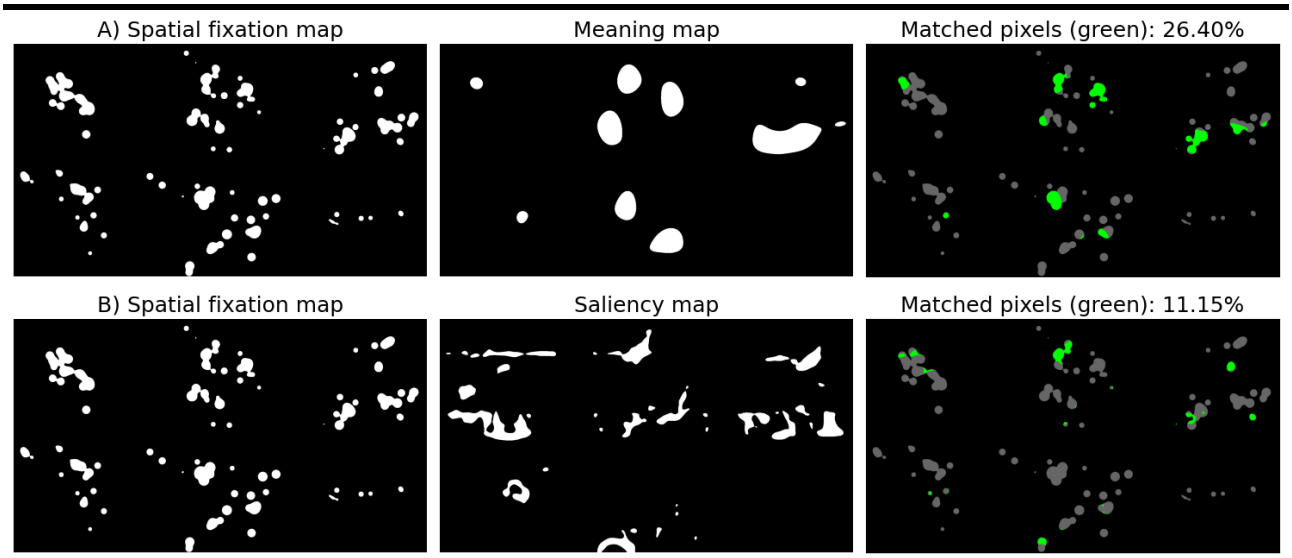


Figure 5 shows a spatial fixation map for a typical control participant which was overlaid on the meaning map and the saliency map.

Figure 5. Example of the spatial overlaid results for a typical control participant. The figures show the overlapping pixels in the fixation map and the meaning map (A), as well as in the fixation map and the saliency map (B).



Overview of lexical retrieval

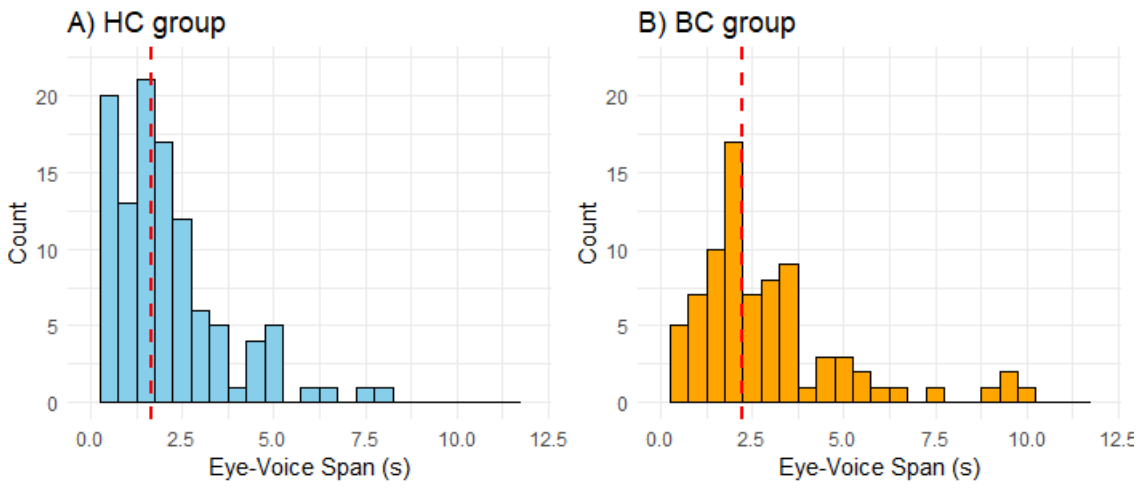
Table 3 (see the Appendices) offers an overview of a descriptive qualitative analysis of lexical search events. A Mann-Whitney U test was performed to compare the frequency of word retrieval in BCs and HCs. The BC group did not search for their words significantly more often than the control group (BC_{median} = 10 words, HC_{median} = 7 words); U = 58, p = .686. Table 4 (see the Appendices) reports the main and additional narrative features reported by the participants. Main narrative features are the elements which structure the story and enable its understanding. Additional features are details which do not contribute to the narrative thread. All main features were recognized by all the participants except for one HC participant who did not mention the windsurfing activity. The number of additional features were disparate and participant-dependent (BC_{median} = 3.5 [SD = 2.71], HC_{median} = 3.22 [SD = 4]).

Longer Eye-Voice Span in cancer survivors

A Mann-Whitney U test was performed to compare the EVS duration in BCs and HCs. The EVS duration was significantly longer in the BC group (N = 8, M = 2.24 s) than in the HC group (N = 13, M = 1.65 s); U = 3344.5, p = .002. This result suggests that cancer survivors took longer than controls to name items from the first gaze with the intent to speak to word articulation. Figure 6 shows the distribution of EVS duration (s) by groups. The BC group displayed longer duration than the HC group. A qualitative analysis revealed no substantial difference between groups regarding the word

impacted by the lexical retrieval.

Figure 6. Distribution of Eye-Voice Span duration (s) for (A) the control group and (B) the breast cancer survivors. The red dotted line indicates the median.



Discussion

This study opens new avenues for investigating word-finding difficulties reported by breast cancer survivors. Speech production was examined in combination with an eye-tracking technique to gain deeper insight into the underlying cognitive mechanisms of anomia. The main finding reveals that survivors experienced a significantly longer delay in accessing their mental lexicon and producing words compared to the control group. Additionally, the analysis of time spent on each image showed a tendency for post-cancer participants to take longer to complete the task. While this trend was observed, no statistically significant difference was found between the groups. Taken together, these results suggest that an underlying deficit in executive functioning may contribute to the observed anomia in cancer survivors.

Naming latency induced by executive dysfunction?

Overall, cancer survivors searched for words more frequently than the controls. The words they were looking for were mostly items visually represented in the image, which survivors focused on until retrieving the word's phonological form. This correlates with the EVS measure, which revealed a significantly longer latency between gazing at a visual item with the intent to speak and its subsequent naming. During this time, the BC and HC groups continued to stare at the object to be retrieved, with rare glances at other objects. Research on eye-movements made during speech production revealed that time spent on gazing at an object is highly correlated to the lexical retrieval and phonological

encoding of its associated word (Griffin & Bock, 2000). In non-prepared speech, speakers select and proceed words while speech unfolds, which increases the chance of interruptions in speech flow. According to (Griffin & Bock, 2000) and (Griffin, 2004), eye movements anticipate the naming of informative elements in speech production free from lexical difficulties, and the fixation of these elements occurs less than a second (around 900 ms) prior to their verbal output. Moreover, in a context of multiword production the speed of naming is correlated with executive control abilities, particularly those related to updating and inhibiting capacities (Shao et al., 2012). Hence, the longer duration of EVS seems to indicate a dysfunction in lexical and/or phonological processes (Pistono & Hartsuiker, 2021). Thus, cancer survivors who reported cognitive changes may experience an increased time to select and encode the phonological form of a word. This latency emphasizes the effort that survivors make to retrieve a word, even when they are able to produce the right word. The prolonged duration of EVS in BCs supports the hypothesis that word-finding difficulties are caused by a deficit in executive control that is not specific to language processing. On that point, a systematic review on executive function in breast cancer patients found clear evidence of impairments in the core components of executive functions (Yao et al., 2017). Interestingly, while the review reported disparity in neuropsychological measures, it highlighted that inhibition consistently emerged as a recurrent deficit across the studies.

Interplay of inhibition and attentional processes

So far, the findings of this study bolster the notion of an executive control deficit rather than deficits in memory or attention. No significant difference were observed in the number of fixations on AoIs across groups, indicating that semantic processing was intact for survivors. Likewise, the fixation map analysis showed that the BC group exhibited a pattern akin to the HC group, focusing more on meaningful elements than on salient ones. Hence, difficulties in lexical search might not be driven by issues with visual attention. While the memory deficit hypothesis can be reasonably dismissed, the attention-related hypothesis is harder to dismiss, as this study does not provide enough information to reject it.

Considering the close interplay between attention and inhibitory processes, a growing body of evidence points out that attentional deficits in breast cancer survivors may stem from upstream of executive functioning. Melara and colleagues' research suggests that cancer survivors face sensory gating deficits, which could lead to subsequent cognitive impairment. For instance, in a study using an auditory oddball paradigm, the team found that cancer patients demonstrate reduced inhibition of redundant information and memory dysfunction in encoding new information (Melara et al., 2023). Likewise, using a revised version of the Attention Network Test (which measures reaction time to the detection of visual stimuli), Melara et al. (2023) found significant Event-Related Potential (ERP) effects of cancer treatment on inhibitory control in breast cancer patients. Invalid cues that preceded the target resulted in more time for breast cancer patients to recognize that the cue was incorrect. According to the authors, attentional deficits found in cancer patients after treatment begin in "early

sensory/perceptual processing and extend to compensatory top-down executive processes." Consistent with these findings, Edelman et al. (2024) demonstrated that breast cancer patients performed worse than controls on an antisaccade task, particularly among older patients who had undergone adjuvant treatments, supporting the hypothesis of a decline in the efficiency of early attentional inhibitory processes. In this study, the analysis of the fixation map across groups revealed no evidence of disruption in bottom-up processes of attention. However, the prolonged duration of EVS may indicate difficulty with covert attention. Further investigation is needed to gain a better understanding of the attentional processes that may influence naming latency in breast cancer survivors.

Limits and future directions

While comparing meaning maps and saliency maps is a promising approach for investigating subtle attention deficits in breast cancer patients, the method used in this study has several limitations. First, the pictures were not controlled for low-level visual parameters. Thus, the salient features were too scattered to effectively capture the patients' attention. Additionally, the scene complexity was relatively simple. In contrast, a more detailed, picture-based story might increase the difficulty of lexical retrieval in breast cancer survivors. Second, parafoveal processing was not explored, although some cancer treatments may cause visual neurotoxicity (Anderson et al., 2020), and deficits in visual modalities have been observed in multiple studies (Oliva et al., 2024). In this study, participants successfully retrieved the main elements of the story, indicating that these elements were processed regardless of whether they had been fixated on or not. A qualitative examination of verbal semantics revealed that the attention to irrelevant details was more related to individual differences than to the group. However, an increase in the complexity of details may slow down visual perception in breast cancer survivors.

Conclusion

This study investigated three potential cognitive mechanisms underlying word-finding difficulties in breast cancer survivors: memory, attention, and executive deficits. The analysis of speech and oculomotor behavior in a picture-based task suggests that, despite patients' concerns about memory, anomia in the context of cancer is more likely linked to executive function deficits rather than memory issues. However, the employed method did not provide sufficient evidence to support an attentional explanation, indicating the need for further research in this area. While word-finding difficulties are most commonly documented in breast cancer survivors, they are not exclusive to this group. Anomia is a widespread issue across various cancer types, and this method could benefit a broader population by being applicable to other cancer groups.

Acknowledgement

The authors would like to thank Anne Cheylus for her help in developing the Python scripts.

Amélie B. Richard is financed by a PhD grant awarded by Paul-Valery Montpellier 3 University and funded by the French Ministry of Higher Education (Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation — MESRI).

Appendices

Table 3. Overview of a qualitative analysis of lexical search events. Data loss: no data were recorded in a short lap of time preceding the lexical search. Narrative items: words which are not present in the picture or which depict abstract concepts (e.g., "family"). Fixation duration below 100 ms: the target item was fixated less than 100 ms.

Variable	BC (N = 8) Total (mean)	HC (N = 13) Total (mean)
Number of lexical retrieval difficulties associated with a fixation	81 (10.25)	111 (8.54)
Number of words without fixation	22 (2.75)	28 (2.15)
Number of words without fixation – data loss	2	0
Number of words without fixation – narrative items	15	24
Number of words without fixation – looked elsewhere	4	3
Number of words without fixation – fixation duration below 100 ms	1	1
Including abstract concepts (narrative items)	10 (1.25)	22 (1.69)

Table 4. List of narrative features (i.e., Man and child on the beach – Woman presence – Windsurfing – Wind and problem – Phone call) not recognized, and additional features described by the participants.

Participant	Main narrative features not recognized	Additional narrative features described
Control group		
L02		Sailboard – Man Cap – Man Glasses – Ball – Cooler – Sunshade – Woman Sunglasses – Palm Tree – Beverage
L03		Sailboard – Ball – Book – Cooler
L04		Sunshade – Sailboard – Beverage – Cooler – Book
L05		Man Cap – Man Glasses – Man Short – Kid Short – Kid Shirt – Sunshade – Ball – Woman Sunglasses – Cooler
L06		Ball – Beverage – Cooler
L07		Man Glasses – Man Yellow Cap – Sailboard – Sunshade – Cooler – Ball – Palm Tree – Deckchair – Woman Sunglasses – Beverage – Woman Swimmer
L08		Ball
L09		Book – Beverage
L10	Windsurfing	Ball
L12		Ball – Book
L14		Ball – Book – Beverage – Cooler
L19		Ball
L23		Ball – Palm Tree – Man Cap
Breast Cancer group		

(Table continues)

(Continued)

Participant	Main narrative features not recognized	Additional narrative features described
L11		Sunshade – Sailboard – Ball – Woman Glasses – Woman Hair – Cooler – Sunshade on the beach – Palm Tree – Book – Woman Glasses Wore – Beverage
L13		Sailboard – Sunshade – Ball – Cooler
L15		Ball – Book
L16		Ball – Book – Beverage – Cooler
L17		Multiple Families – Book – Woman Sunglasses – Someone Is Drowning
L20		Sunshade – Sailboard – Beverage
L22		Sailboard – Ball – Palm Tree – Beverage
L24		Ball – Beverage

Discussion générale

L'objectif principal de ce manuscrit était de mieux comprendre les difficultés cognitives décrites par les survivants d'un cancer, et en particulier le manque du mot. Cette thèse a d'abord cherché à identifier des paramètres prosodiques pouvant indiquer la présence d'un trouble cognitif induit par un cancer et ses traitements (*Cancer-Related Cognitive Impairment*, CRCI). Elle a ensuite investigué les potentiels mécanismes cognitifs conduisant à un manque du mot en contexte de CRCI.

Dans un premier temps, cette discussion s'articule autour des marqueurs prosodiques du CRCI. Elle essaye de déterminer dans quelle mesure des facteurs confondants tels que la fatigue influencent ces difficultés cognitives, puis elle introduit la notion de continuum de CRCI. Dans un deuxième temps, elle explore les déficits cognitifs sous-jacents au manque du mot en contexte de CRCI. Elle passe ensuite en revue les thérapies expérimentées pour la rééducation du CRCI. Pour conclure, elle présente les biais et les limites rencontrés dans ces travaux, avant de proposer des pistes de recherche pour les travaux ultérieurs.

1 Indices prosodiques d'un CRCI

La première étude expérimentale de cette thèse visait à identifier les potentiels marqueurs prosodiques du CRCI. L'utilisation d'un modèle de classification automatique a révélé que la combinaison du ratio de parole sur silence, de la durée moyenne des pauses silencieuses et de la durée moyenne des pauses remplies permettait de différencier les survivantes d'un cancer présentant des difficultés cognitives de sujets témoins.

1.1 Interruptions silencieuses

Il est intéressant de souligner que deux de ces marqueurs de parole sont liés à la durée des interruptions silencieuses. Ces résultats s'inscrivent dans la continuité de travaux menés sur d'autres populations présentant des troubles cognitifs subtils dans des contextes pathologiques différents. À titre d'exemple, Vincze et al. (2021) ont comparé des individus présentant un *Mild Cognitive Impairment* (MCI) à des patients ayant un diagnostic de maladie d'Alzheimer ainsi qu'à des sujets témoins sains. Le MCI est un syndrome cognitif qui témoigne souvent d'un stade intermédiaire entre le vieillissement normal et le début d'une maladie d'Alzheimer (Petersen, 2016). Tout comme le

CRCI, le MCI se caractérise par une plainte cognitive concernant la mémoire et le manque du mot. L'expression du MCI est également discrète, amoindrissant la qualité de vie des personnes sans pour autant affecter leur autonomie. De fait, les tests neuropsychologiques actuels peinent à détecter ce syndrome. L'objectif de Vincze et son équipe était donc de repérer les individus présentant un MCI au moyen de paramètres de parole, notamment les pauses silencieuses et les pauses remplies. Ces pauses silencieuses, qualifiées « d'hésitations », étaient définies par l'absence de parole entre deux mots pendant plus de 30 ms. Selon leurs résultats, les personnes présentant un MCI produisaient davantage de pauses silencieuses, et ces pauses étaient de plus grande durée, par rapport aux témoins ainsi qu'aux personnes atteintes d'une maladie d'Alzheimer (Vincze et al., 2021). À l'instar de nos propres résultats, les variables liées aux silences se sont révélées les plus représentatives de la population souffrant d'un CRCI.

De manière similaire, une étude menée par K. M. Smith et al. (2018) a mesuré la durée des pauses silencieuses lors d'une tâche de description d'image, dans le but d'identifier les marqueurs de parole d'une anomie chez des individus atteints d'une maladie de Parkinson. Les patients présentaient des troubles moteurs mineurs et deux tiers d'entre eux rapportaient un manque du mot. Cependant, leurs scores aux tests cognitifs demeuraient dans les normes. La fréquence des pauses silencieuses (supérieures à 1 seconde) était plus élevée chez les patients que chez les témoins. De plus, le taux de pauses avant un verbe était corrélé à l'évaluation de l'anomie rapportée par les patients (K. M. Smith et al., 2018). De façon comparable, les données de nos études ont permis de prédire l'anomie décrite par les survivantes d'un cancer du sein avec plus de précision que le questionnaire FACT-Cog. Ce dernier, outil clinique de référence pour détecter la présence, au moins subjective, d'un CRCI, n'a cependant pas été conçu pour identifier un manque du mot. Le fait que la production de silence soit corrélée avec la plainte des patients souligne l'intérêt de cette méthode pour évaluer ces difficultés cognitives.

La plupart des études explorant la relation entre les troubles neurocognitifs mineurs et la parole portent sur des populations ayant des pathologies ou des syndromes neurodégénératifs liés à l'âge. Toutefois, l'intérêt pour les pauses silencieuses a également été observé chez des sujets plus jeunes, notamment chez des individus atteints de sclérose en plaques (SEP). La SEP est une maladie neurologique progressive affectant généralement des sujets jeunes et peut entraîner des troubles cognitifs subtils (Sumowski et al., 2018). Par exemple, une étude comparant des personnes atteintes de SEP à des sujets témoins a révélé que, parmi un ensemble de disfluences, seules les pauses silencieuses permettaient de distinguer les groupes (Bóna, 2024).

Les pauses silencieuses ont également démontré leur pouvoir discriminant dans le cadre d'aphasies d'origine vasculaire. Une étude a révélé que des personnes présentant des symptômes légers d'aphasie non fluente (caractérisée par un manque du mot) après un AVC produisaient des interruptions silencieuses significativement plus longues que les sujets témoins sains (Angelopoulou et al., 2018).

D'un point de vue linguistique, il est intéressant de souligner que les études précédemment citées

incluaient des locuteurs de hongrois (Bóna, 2024 ; Vincze et al., 2021), d'anglais (K. M. Smith et al., 2018), de grec (Angelopoulou et al., 2018) et de français (Richard, Lelandais & Jacquin-Courtois, under review). Malgré la diversité linguistique, les pauses silencieuses ont été identifiées comme un marqueur de trouble cognitif dans ces quatre langues. Par conséquent, la durée ou la fréquence des pauses silencieuses pourrait constituer un marqueur commun à plusieurs systèmes de langue. Ces comparaisons inter-linguistiques doivent toutefois être nuancées, car les pauses silencieuses font souvent l'objet de définitions et de mesures différentes selon les études. En effet, la revue systématique portant sur les marqueurs linguistiques de troubles cognitifs subtils que nous avons menée a mis en évidence la variété des paramètres de parole employés à travers les travaux sur cette thématique. Dès lors, il semble essentiel de recourir à des paramètres qui ne soient pas influencés par des postulats cognitifs ou grammaticaux.

Le ratio parole sur silence pourrait répondre à cette problématique. Parmi les paramètres que nous avons étudiés, ce ratio s'est distingué comme l'un des plus prometteurs pour l'identification de difficultés cognitives. En dehors du contexte clinique, ce paramètre a démontré son efficacité pour mesurer la fluence d'apprenants d'une deuxième langue (Cucchiarini et al., 2002 ; De Jong, 2023) ou pour comprendre la relation entre la taille du lexique et les capacités de fluences verbales (Jongman et al., 2021). Toutefois, cette mesure reste encore peu utilisée dans les études sur la détection des troubles cognitifs subtils (voir Pakhomov et al. [2010] sur la dégénérescence fronto-temporale lobaire). Il n'a d'ailleurs pas émergé de notre revue systématique (Richard, Lelandais et al., 2024). Ceci est probablement dû au fait que les études effectuent généralement une distinction entre les pauses silencieuses et omettent de leurs mesures celles qui seraient liées à une fonction grammaticale (p. ex. entre deux syntagmes).

Le calcul de ce ratio consiste à diviser la durée totale des silences par la durée totale de l'échantillon. Un ratio élevé indique la présence de moins de pauses silencieuses, et donc un discours perceptiblement plus fluide. Ce ratio peut être facilement obtenu avec un système de reconnaissance automatique de la parole (*Automatic Speech Recognition*, ASR) sans nécessiter de transcription (De Jong, 2018).

Contrairement à la moyenne ou à la médiane de la durée des pauses silencieuses, qui peuvent être influencées par des valeurs extrêmes, le ratio parole sur silence prend en compte l'intégralité des silences. Cette mesure est également moins sensible aux variations stylistiques du discours, car elle ne repose pas sur les aspects qualitatifs des pauses. En conséquence, elle limite les questions liées à la « nature » ou à la « fonction » des pauses.

Bien que ce ratio se révèle être un moyen efficace pour détecter les difficultés cognitives, il ne permet pas d'approfondir les hypothèses sur les troubles sous-jacents. L'ajout d'une autre technique d'analyse comportementale, telle que l'oculométrie, permet ainsi de mieux comprendre les processus cognitifs sous-jacents aux difficultés rencontrées par les survivantes d'un cancer du sein lors de la production de messages verbaux.

1.2 Relation entre silence et facteurs confondants

Nos résultats ont montré une tendance des variables de parole à être corrélées aux scores de fatigue et de dépression. Cependant, il semble prématuré de conclure à une relation unidirectionnelle entre les scores obtenus à ces tests et la production de parole pour plusieurs raisons.

La présence de symptômes dépressifs peut influencer la vitesse du débit de parole, notamment par l'allongement de la durée des pauses (Albuquerque et al., 2021). Les survivantes d'un cancer présentant une plainte cognitive ont obtenu des scores indiquant des niveaux d'anxiété et de dépression supérieurs à ceux du groupe témoin. Toutefois, interprétés isolément, ces scores ne révèlent pas la présence de ces symptômes chez toutes les survivantes. Par ailleurs, les scores obtenus aux questionnaires d'évaluation de l'anxiété et de la dépression demeurent en dessous des seuils d'alerte (Brennan et al., 2010; Mitchell et al., 2010; Zigmond & Snaith, 1983). De fait, la différence de niveaux d'anxiété et de dépression ne correspond pas à une réalité clinique, puisque ces derniers restent en dessous des seuils de référence.

En ce qui concerne la fatigue, il n'existe pas de seuil standardisé pour détecter la présence de fatigue significative chez des patientes atteintes d'un cancer du sein. La fatigue est un symptôme très fréquent chez les survivants d'un cancer (Hofman et al., 2007), ce qui rend possible la présence d'une fatigue significative. Néanmoins, la relation entre la fatigue et la diminution de la fluence est relativement peu documentée et n'est pas clairement établie à ce jour. Une relation avérée concerne l'influence de la fatigue sur le contrôle moteur de la parole (Hartelius et al., 2004). Or, les participantes ayant eu un cancer n'ont pas présenté un débit articulatoire plus lent que celui des participantes du groupe contrôle. Par conséquent, cette observation écarte l'hypothèse d'une influence de la fatigue sur les dimensions articulatoires chez les survivantes.

En somme, la dépression et la fatigue sont des facteurs fréquemment associés au CRCI (Gullett et al., 2019). Toutefois, ces facteurs sont souvent des éléments confondants liés aux traitements, notamment la chimiothérapie, qui peuvent exacerber le CRCI sans pour autant expliquer de manière directe les dysfonctionnements cognitifs observés (Meyers et al., 2005).

De plus, les questionnaires autoadministrés sont dits « explicites », en ce sens que les participants ont connaissance des critères sur lesquels ils sont évalués. La compréhension des critères de jugement peut dont influencer leurs résultats, par un biais d'effet lié à l'attente (*expectancy effect*). En effet, les réponses des participants sont conditionnées par leurs attentes concernant leur appartenance à un groupe spécifique (cancer ou groupe témoin). Ainsi, les survivantes de cancer pourraient être tentées d'amplifier les traits de leurs difficultés, tandis que les participantes du groupe témoin pourraient être enclines à surestimer l'efficience de leur cognition. À l'inverse, les variables prosodiques restent implicites, ce qui souligne l'intérêt de combiner ces deux approches.

Pour finir, la relation observée entre la production de parole et les scores de dépression ou de fatigue pourrait en réalité capturer la corrélation entre le FACT-Cog et ces mesures. En effet, le CRCI

participe d'un ensemble de symptômes secondaires associés au cancer (Kim et al., 2012 ; Starkweather et al., 2013), dont l'anxiété, la dépression et la fatigue font partie. Autrement dit, la relation entre la parole et les variables psychologiques pourrait être expliquée par le fait que les survivantes ayant signalé un trouble cognitif ont également rapporté des niveaux accrus d'anxiété, de dépression et de fatigue. Veal et al. (2023) ont observé que les survivants de cancer percevant davantage de déficits cognitifs étaient aussi ceux qui rapportaient le plus de fatigue, et que ceux présentant plus de fatigue et de dépression avaient également mentionné plus de problèmes de mémoire. Cette observation est confirmée par notre étude utilisant des modèles d'apprentissage automatique. Dans cette étude, la population a été répartie en trois groupes : un premier groupe de survivantes dont les scores FACT-Cog signalent une gêne cognitive, un deuxième groupe de survivantes sans plainte cognitive et un troisième groupe de sujets témoins ne présentant pas de plainte cognitive. Cependant, plusieurs survivantes sans plainte cognitive ont été classées par les modèles dans le groupe des personnes présentant un CRCI. Ces erreurs de classification semblent dues à des similitudes entre leurs caractéristiques prosodiques et celles des survivantes ayant une plainte cognitive.

1.3 Continuum des changements cognitifs induits par un cancer

Les erreurs de classification commises par nos modèles sont révélatrices. En règle générale, les modèles sont mieux parvenus à classer les participantes en fonction de leur difficulté à trouver leurs mots (rapportés de manière déclarative) que selon leur score FACT-Cog. Parmi les survivantes du corpus, un tiers d'entre elles a obtenu un score inférieur au seuil de détection d'une plainte cognitive selon le FACT-Cog. Cependant, six de ces onze participantes ont été incorrectement classées dans le groupe présentant un CRCI. Dans cinq de ces cas, cette classification erronée faisait écho aux difficultés linguistiques qu'elles percevaient. Le décalage entre le score au FACT-Cog et la production d'interruptions de parole pourrait indiquer un statut cognitif se situant sur un continuum de difficultés cognitives perçues et objectivées, entre cognition normale et troubles cognitifs manifestes.

De manière générale, la durée des pauses permet de distinguer clairement le groupe de survivantes avec plainte du groupe contrôle. En revanche, les variables du groupe sans plainte présentent une plus grande dispersion. D'une part, seules trois participantes sur les onze qui constituent ce groupe ont déclaré ne pas avoir de difficultés linguistiques. D'autre part, la majorité de leurs pauses silencieuses se rapprochent — en durée — de celles des survivantes avec une plainte, tandis que certaines se situent plutôt dans la fourchette des pauses observées chez les sujets témoins. Ainsi, le groupe de survivantes sans plainte cognitive a révélé une dispersion importante, qui mérite d'être davantage explorée. Cette variabilité pourrait être expliquée par plusieurs facteurs, notamment les traitements. Nous n'avons pas été en mesure de contrôler les traitements reçus par les participantes en raison de la diversité des traitements oncologiques et de la taille restreinte de notre échantillon. Or, les différences observées pourraient être liées à ces traitements, puisque certains d'entre eux, comme le tamoxifène (Ahles et al., 2010 ; Bender et al., 2006 ; Schilder et al., 2010) ou les chimiothérapies (Quesnel et al., 2009), ont un impact avéré sur les performances verbales. En outre, Nguyen et al.

(2013) ont trouvé que les traitements locaux (chirurgie et radiothérapie) altéraient davantage la cognition comparés à la chimiothérapie chez des survivantes de cancer âgées de plus de 65 ans, dix ans après la guérison. Toutefois, les personnes traitées par chimiothérapie présentaient elles-mêmes des performances amoindries contrairement aux contrôles. En revanche, Rick et al. (2018), après avoir contrôlé différents types de traitements, n'ont pas trouvé d'effet des traitements oncologiques sur la cognition. Une méta-analyse menée par Bernstein et al. (2017) révèle que les performances cognitives globales de patients traités par chimiothérapie diffèrent de celles des contrôles, mais pas de celles de patients ne recevant pas de chimiothérapie. Seuls les scores d'attention et de fonctions exécutives variaient entre les patients traités et ceux non traités par chimiothérapie. Toutefois, cette différence serait en partie dépendante des groupes observés et des mesures neuropsychologiques utilisées (Bernstein et al., 2017).

Dès lors, le CRCI pourrait représenter un continuum de difficultés cognitives, dont les répercussions varient non seulement en fonction de leur sévérité et de leur persistance (Ahles & Root, 2018), mais aussi selon des facteurs externes tels que les événements vécus au cours d'une journée. À l'occasion d'une récente présentation, Ahles (2025) a souligné le caractère fluctuant de la gêne cognitive induite par les cancers, dont les phénotypes varient d'un patient à l'autre. Le CRCI serait sujet à une variabilité interindividuelle. Ahles (2025) avance l'idée que des lésions sur les mêmes sites cérébraux peuvent entraîner des conséquences diverses qui dépendent des individus. De plus, chaque individu présente des variations de performances cognitives au fil du temps. Ces variations seraient intensifiées par le cancer et ses traitements. En d'autres termes, l'expression du CRCI pourrait être modulée par les expériences de vie quotidienne de chaque patient. Le CRCI pourrait ainsi s'inscrire sur un spectre variable de changements cognitifs. Dès lors, les marqueurs de parole pourraient refléter une facette spécifique de ce continuum. Une analyse complémentaire réalisée à différents points temporels apparaît donc essentielle pour valider cette hypothèse.

2 Corrélats cognitifs de l'anomie en contexte de CRCI

La dispersion des marqueurs de parole reflèterait le continuum de difficultés cognitives rencontrées par les patients. Quelles sont précisément la nature et l'étendue de ces difficultés ? L'analyse du comportement oculomoteur lors de la production de parole a permis de mieux comprendre certains des mécanismes sous-jacents au CRCI.

2.1 Délai de dénomination

Le manque du mot est un symptôme courant du CRCI, bien que ses corrélats cognitifs demeurent encore mal connus. De manière générale, les difficultés lexicales peuvent découler soit d'une dégradation du système sémantique, soit d'un déficit d'accès au lexique, souvent lié à une altération

du contrôle exécutif. Une troisième hypothèse suggère également que l'attention joue un rôle clé dans les processus de sélection lexicale.

Tout comme le manque du mot, la plainte mnésique est une source d'inquiétude qui ressort des études qualitatives recueillant les difficultés cognitives vécues par les patients (Downie et al., 2006). Cependant, comme le suggère Ahles et Root (2018), il est possible que les patients confondent leur perception de ces difficultés avec un déficit attentionnel ou exécutif. En effet, un manque d'attention ou un dysfonctionnement du contrôle exécutif pourrait enrayer l'étape d'encodage mnésique, ce qui se traduirait par un sentiment d'oubli (Mulligan, 1998).

Pour rappel, dans le cas d'un manque du mot d'origine sémantique, les individus confondent les concepts ou les ont oubliés (Caramazza & Hillis, 1990). En revanche, dans le cas d'un manque du mot d'origine exécutive, les individus ont accès au concept, mais pas à sa forme phonologique (c.-à-d. les sons des mots) (Caramazza & Hillis, 1990). À ce titre, le discours verbatim des patients offre un éclairage précieux sur les processus sous-jacents à l'anomie qu'ils rencontrent. Dans une étude menée par Kanaskie et Loeb (2015), des entretiens semi-dirigés avec des patientes post-cancer du sein révèlent les difficultés cognitives auxquelles elles sont confrontées. L'anomie, seul symptôme unanimement partagé, est mentionnée par les sept patientes de l'étude. Une patiente décrit une impression de vide lorsqu'elle parle ; elle visualise ce qu'elle veut dire, sans pour autant atteindre le mot adéquat. Comme le rapportent les auteurs, les patientes exprimaient la sensation de ne pas avoir accès aux mots. Ce phénomène est illustré par une participante d'une étude de Player et al. (2014) : « Je vais dire « cette plante violette là » mais je ne pourrai pas vous dire que c'est de la lavande… jusqu'à bien plus tard. »[1] (Player et al., 2014, p. 5). Dans une étude de Bolton et Isaacs (2018), une autre patiente décrit : « Oh, est-ce que tu peux me passer le truc pour écrire, le bâton pour écrire avec l'encre dedans ? Je sais ce que je veux, mais je n'arrive pas à trouver le mot. [2] (Bolton & Isaacs, 2018, p. 1226). Nos participantes rapportent des difficultés similaires : « C'est pas parce que je le connais pas, mais je cite l'autre mot parce que c'est plus facile. […] Je vais dire à mes enfants, va prendre ça sur le meuble blanc. Alors que je sais qu'il est brun par exemple. » ou encore « Pour euh p-pour essayer d'expliquer par exemple euh… pour euh… pallier à-à ce problème où je cherche mes mots, ben je peux passer par une définition, comme ça, ça se voit pas. » [3].

À la lumière de ces éléments, il semble que le manque du mot des patients ne résulte pas d'une dégradation du système sémantique. Cette hypothèse est soutenue par le fait que les études en neuroimagerie ne révèlent ni lésions apparentes ni activation anormale dans les zones associées à la mémoire sémantique des mots (Bernstein et al., 2021), telles que les régions temporales gauches (Vigneau et al., 2006). En somme, les résultats d'imagerie concernant les régions associées à la mémoire divergent sans toutefois être clairement établis (Sousa et al., 2020). En revanche, les études en neuroimagerie menées chez les patients atteints de cancer mettent clairement en évidence

[1] Phrase originale : « *I'll say 'that purple plant there' but I can't tell you it's lavender … until much later* »

[2] Phrase originale : « *Oh can you pass me the writing thing, the writing stick with the ink in it'. I know what I want but I can't find the word…* »

[3] Données de thèse non publiées

des lésions ou des activations anormales dans les régions du cerveau associées aux difficultés attentionnelles et exécutives. Kesler et al. (2009) montre, par exemple, une activation réduite des régions préfrontales lors des processus d'encodage mnésique dans le cadre d'une tâche de mémoire verbale (stimulus visuel). Bien qu'il n'y ait pas de différence significative concernant le nombre d'erreurs entre le groupe de patientes atteintes de cancer et les sujets témoins, le temps de réaction était plus lent pour les patientes.

Nos données d'oculométrie semblent confirmer l'idée d'un déficit exécutif plutôt que d'un déficit mnésique. D'une part, l'analyse spatiale et temporelle des points de fixation indique que les survivantes d'un cancer n'ont pas besoin de regarder davantage les images pour retrouver leurs mots. D'autre part, le calcul du délai de dénomination entre la première visualisation d'un item et sa dénomination montre que les survivantes présentant une plainte cognitive mettent plus de temps que les contrôles à rechercher leurs mots. Ensemble, ces résultats suggèrent un problème lié à la sélection lexicale ou à l'encodage phonologique des mots. Autrement dit, ils écarteraient la possibilité d'une dégradation du système sémantique comme cause primaire du manque du mot. Au contraire, ils pointeraient plutôt un problème lié à la forme du mot : les concepts sont préservés, mais leur forme phonologique reste temporairement inaccessible. Lorsqu'ils détectent une différence, les scores des tests de mémoire refléteraient davantage un défaut exécutif et attentionnel qu'un trouble mnésique à proprement parler.

2.2 Parole continue vs. fluence verbale

Le manque du mot est une plainte courante chez les patients atteints de cancer. Alors que les tâches de fluence verbale peinent à les révéler, nos mesures prosodiques semblent être plus précises pour indiquer ces difficultés lexicales.

Disparité des résultats de fluences verbales

Selon les recommandations de l'ICCTF (Wefel et al., 2011), l'évaluation du CRCI devrait inclure une tâche de fluence verbale. Cette tâche permet autant d'évaluer l'accès au lexique que les fonctions exécutives. La consigne consiste à demander au participant de trouver le plus grand nombre de mots en un temps donné (généralement une minute) commençant par une certaine lettre (fluences littérales ou phonémiques) ou appartenant à une catégorie spécifique (fluences catégorielles ou sémantiques). Les noms propres et les dérivés morphologiques ne sont pas autorisés et les répétitions ne sont pas comptées. La plupart des études cliniques calculent la fluence en termes de nombre de mots corrects (p. ex., Andryszak et al. [2017]), bien que certaines études incluent également le nombre d'erreurs (p. ex., Balogh et al. [2023]).

Plusieurs études sur l'évaluation du CRCI incluent des tests de fluence verbale, mais les résultats

sont divergents. Certaines d'entre elles confirment la présence d'une différence de fluence verbale entre les patients atteints de cancer et les témoins (Boele et al., 2015), tandis que d'autres n'observent aucune différence (Mehlsen et al., 2009 ; Nguyen et al., 2013). De même, certaines études soulignent l'impact de la chimiothérapie sur la performance en fluence verbale (Durán-Gómez et al., 2022 ; Janelsins et al., 2022), tandis que d'autres ne constatent aucun effet (Biglia et al., 2012 ; Hurria & Lachs, 2007). Une étude longitudinale présente même un résultat contre-intuitif : Andryszak et al. (2017) rapportent une amélioration des fluences phonémiques chez les contrôles ainsi que chez les patientes traitées par chimiothérapie au cours du temps. Cependant, l'amélioration des performances est attribuée par les auteurs à un effet d'entraînement. Par ailleurs, elle n'a pas été pas retrouvée pour les fluences sémantiques chez les patientes atteintes de cancer du sein (Andryszak et al., 2017). En outre, cette étude ne compare pas directement les groupes. Il est donc difficile de conclure que les patientes aient réalisé de moins bonnes performances que les témoins.

Fonctions cognitives sollicitées par les fluences verbales

En clinique, les tâches de fluences verbales sont administrées soit pour estimer les capacités d'accès au lexique, soit pour évaluer les fonctions exécutives. La fluence de mots isolés est sous-tendue par l'interaction entre les processus attentionnels, la mémoire verbale à long terme et les connaissances sémantiques (Ruff et al., 1997).

Shao et al. (2012) ont examiné l'association entre différentes mesures des fonctions exécutives (mise à jour, flexibilité et inhibition) et les performances à des tâches de fluence verbale. Les résultats ont montré une corrélation de la fluence verbale avec les capacités de mise à jour et d'inhibition, mais aucune corrélation avec la flexibilité. Dans une étude ultérieure, Shao et al. (2014) a démontré que le score moyen des fluences dépend principalement des capacités de mise à jour, plutôt que de la taille du vocabulaire ou de la vitesse d'accès lexical.

Whiteside et al. (2016) ont examiné si les fluences verbales littérales et catégorielles solliciteraient principalement les fonctions exécutives ou les compétences linguistiques, telles que les connaissances sémantiques. L'équipe a administré des tests de langage et des tests de fonctions exécutives, en plus des tâches de fluence verbale, à 304 patients adultes (moyenne d'âge de 45 ans) présentant des étiologies et des syndromes variés. Une analyse à deux facteurs (langage et fonctions exécutives) a révélé que les fluences verbales étaient principalement associées aux tests de dénomination, mais non aux tests des fonctions exécutives. Les résultats suggèrent que les dimensions sémantiques du langage constituent une composante essentielle des fluences verbales (Whiteside et al., 2016). Cependant, les résultats de cette étude ne permettent pas d'exclure l'éventuelle implication des fonctions exécutives dans la recherche lexicale.

En s'appuyant sur les travaux précédents, Aita et al. (2019) ont de nouveau testé cette hypothèse sur des étudiants en bonne santé (moyenne d'âge de 19 ans), mais en distinguant cette fois les trois composants des fonctions exécutives : raisonnement, mémoire de travail, alternance et mise à jour.

L'équipe a observé une association entre les fluences verbales (phonémiques et sémantiques) et les tests des fonctions exécutives, ainsi qu'entre les fluences et le test de vocabulaire. L'équipe conclut que les fonctions exécutives contribuent également aux fluences verbales phonémiques et sémantiques. En outre, elle attribue un rôle spécifique des fonctions exécutives dans le processus de recherche lexicale, à savoir le contrôle exécutif du langage (Aita et al., 2019).

Examinées ensemble, ces trois études suggèrent que les tâches de fluences verbales sollicitent autant des processus strictement linguistiques que les fonctions exécutives. Ces résultats sont confirmés par des données issues de neuroimagerie. Lors des fluences littérales et catégorielles, l'imagerie cérébrale révèle un chevauchement de l'activation des régions temporales et frontales. Toutefois, une plus grande mobilisation des régions frontales est observée lors des fluences phonémiques, tandis qu'une activation accrue des régions temporales est notée lors des fluences catégorielles (Katzev et al., 2013 ; Schmidt et al., 2019). Autrement dit, les fluences verbales nécessitent à la fois les fonctions exécutives et linguistiques : à ceci près que les fluences phonémiques recrutent davantage les fonctions exécutives, tandis que les fluences sémantiques mobilisent davantage les fonctions linguistiques.

Fluence de mots isolés et fluence en parole continue

Les fluences verbales ont une validité écologique moindre. D'une part, les niveaux de difficulté entre fluences catégorielles et fluences phonémiques ne s'équivalent pas. En effet, les connaissances de sémantique lexicale jouent un rôle plus déterminant dans la fluence catégorielle que dans la fluence littérale (Shao et al., 2014). De plus, la fluence catégorielle serait plus facile à réaliser que la fluence littérale, car la récupération d'un mot activerait d'autres concepts associés. D'autre part, les fluences verbales activeraient des réseaux cérébraux restreints. Birn et al. (2010) ont comparé les régions cérébrales activées lors des tâches de fluences catégorielles et littérales, ainsi que celles activées pendant les tâches de manipulation et de parole automatique. Les résultats ont indiqué une activation importante dans l'hémisphère gauche lors des deux tâches de fluences, tandis qu'une activation plus marquée dans l'hémisphère droit a été observée lors de la tâche de parole automatisée. En revanche, des processus distincts ont été mis en évidence entre les fluences littérales et catégorielles. En effet, la fluence littérale a été associée à une activation plus forte dans le gyrus prémoteur/inférieur frontal, tandis que les fluences catégorielles ont entraîné une activité plus importante dans le gyrus fusiforme gauche et le gyrus frontal moyen gauche. Les fluences verbales isolées activeraient principalement des processus cognitifs restreints et relativement identifiables (Birn et al., 2010).

A contrario, la parole continue repose sur l'activation complexe de processus cognitifs interagissant entre eux (Hagoort, 2019). La production d'un message verbal est donc plus difficile que celle de mots isolés. Chez les survivants d'un cancer, cette activation multiple pourrait prolonger la durée de la recherche lexicale. En effet, des mécanismes de compensation ont été observés dans des études de neuroimagerie lors de tâches psychométriques simples, tandis que les scores de ces

tâches demeuraient dans les normes (Andryszak et al., 2017). Ces mécanismes compensatoires se manifestent par une hyper- ou hypo-activation cérébrale, et ont principalement été retrouvés dans les régions frontales associées au réseau attentionnel et au contrôle exécutif (B. C. McDonald et al., 2012). L'activation de ces régions inhibe les stimuli qui interféreraient dans la production du langage (Dhooge & Hartsuiker, 2012). Cette découverte pourrait expliquer la différence entre les scores normaux obtenus lors des fluences verbales pour les personnes atteintes de cancer et les variables liées à la parole, car ces deux mesures sont distinctes. Ainsi, le délai de dénomination en parole continue, marqué par des interruptions silencieuses ou remplies, serait plus sensible que les fluences verbales pour détecter une anomie subtile.

2.3 L'anomie révèle un trouble plus large

En résumé, l'anomie constitue une caractéristique notable du CRCI. Elle représente ainsi un point d'entrée pour l'étude de ce syndrome. Bien que les patients se plaignent principalement de troubles de la mémoire, de plus en plus d'études suggèrent qu'il pourrait exister un problème en amont du processus d'encodage, notamment lié à un déficit attentionnel. L'anomie dans le CRCI pourrait être une conséquence de ce déficit attentionnel, particulièrement au moment de la recherche lexicale. La relation entre la production de parole et le CRCI mériterait d'être davantage explorée.

3 Interventions cliniques pour le traitement du CRCI

Cette section examine les options thérapeutiques qui peuvent être proposées pour le CRCI. Elle suggère ensuite les pistes d'intervention orthophonique pour les troubles cognitivo-linguistiques associés au CRCI.

À l'instar d'autres troubles cognitifs, la prise en charge clinique du CRCI repose sur des données probantes *Evidence-Based Practice*. Plusieurs types d'interventions destinées à traiter le CRCI sont régulièrement évalués lors d'essais cliniques. Dans une méta-analyse, Zeng et al. (2020), identifient, par exemple, parmi dix types d'interventions non pharmacologiques celles qui ont montré des effets positifs : les interventions de méditation et de pleine conscience pour réduire le stress, ainsi que l'entraînement cognitif, la rééducation cognitive et l'exercice physique. Cependant, une revue systématique menée par Merceur et al. (2024) révèle qu'environ une étude sur deux ne démontre pas l'efficacité des thérapies, y compris la rééducation cognitive, la rééducation physique et les thérapies combinées.

3.1 Thérapies actuelles

Bien que l'objet de cette thèse ne soit pas la prise en charge du CRCI, les résultats qui en sont issus enrichissent les connaissances existantes sur les mécanismes de ce trouble. L'originalité de ce travail repose sur l'apport des données comportementales, par l'analyse de la parole et des mouvements oculaires. Jusqu'à présent, les études portant sur l'évaluation objective du CRCI se sont appuyées sur des tests neuropsychologiques classiques, dont les scores n'étaient que faiblement associés aux difficultés perçues par les patients. Nos données comportementales, au contraire, semblent capturer l'étendue de ces déficits. De plus, nos résultats semblent introduire une continuité entre les modifications cérébrales observées et la manière dont elles se traduisent sur le comportement des survivantes. Ces analyses mettent en lumière des déficits de nature avant tout exécutive. Ces résultats, qui doivent néanmoins être confirmés par des recherches de plus large envergure, sont comparables à ceux de Edelman et al. (2024). À notre connaissance, cette étude est l'une des rares à inclure des données comportementales pour l'évaluation du CRCI. Ces éléments de compréhension quant aux mécanismes sous-jacents au CRCI peuvent orienter le choix d'interventions cliniques.

Traitements pharmacologiques

Les traitements pharmacologiques testés pour le CRCI incluent des molécules utilisées dans le traitement des déficits attentionnels et des démences (Karschnia et al., 2019). Par exemple, l'utilisation d'antagonistes dopaminergiques et noradrénergiques, tels que le méthylphénidate et le modafinil, a été testée chez des patients atteints de cancer pour réguler le réseau préfrontal impliqué dans les dysfonctionnements attentionnels. Cependant, ces traitements n'ont pas montré d'efficacité sur la réduction du CRCI. En ce qui concerne les médicaments anti-démence, l'inhibiteur de l'acétylcholinestérase, le donepezil, qui agit notamment sur l'hippocampe (centre de la mémoire), a montré des effets marginaux. Enfin, l'utilisation d'autres produits, tels que le ginkgo biloba ou l'érythropoétine (une hormone stimulant la production de globules rouges), n'a pas non plus révélé d'effets significatifs. En somme, les résultats des traitements pharmacologiques restent pour l'heure non concluants (Yang & Ah, 2024).

Thérapies cognitives

Les thérapies cognitives destinées à traiter le CRCI sont les interventions les mieux documentées et celles qui présentent les meilleurs taux d'efficacité (p. ex., [Ercoli et al., 2015]). (Bray et al., 2017) démontrent l'efficacité d'un programme de rééducation cognitive à distance auprès de 242 participants. Cette prise en charge a permis une diminution des symptômes cognitifs évalués par le FACT-Cog, bien qu'aucun effet n'ait été observé à travers les tests neuropsychologiques. Pour contourner ce problème, Huang et al. (2025) ont utilisé une technique d'oculométrie dans le but d'évaluer l'efficacité de deux interventions thérapeutiques : une intervention de *biofeedback* visant à

réduire le niveau de stress, ainsi qu'une intervention combinant *biofeedback* et réadaptation cognitive. Les techniques de *biofeedback* permettent au patient de contrôler ses fonctions corporelles (p. ex, rythme cardiaque) en obtenant un retour visuel de ses métriques (p. ex., nombre de pulsations cardiaques par minute) en temps réel. Les cinquante-deux participantes de l'étude, qui étaient des survivantes d'un cancer du sein décrivant des difficultés cognitives, ont bénéficié de l'une ou l'autre des interventions pendant six semaines. Comparé à la ligne de base, les fixations oculaires du groupe combinant *biofeedback* et réadaptation cognitive étaient plus stables et leurs saccades plus courtes que celles de l'autre groupe après l'intervention. Selon les auteurs, ces résultats suggèrent l'effet positif de la réadaptation cognitive sur le maintien attentionnel et le contrôle de l'inhibition chez cette population (Huang et al., 2025). De plus, cette étude montre l'intérêt de mesures comportementales implicites telles que l'*eye-tracking* pour l'évaluation du CRCI.

En ce qui concerne les thérapies cognitivo-comportementales, telles que la psychoéducation sur l'acceptation des pertes cognitives, celles-ci ont montré des effets sur l'amélioration des performances cognitives. Toutefois, des preuves supplémentaires sont nécessaires pour confirmer ces résultats (par exemple, [Ferguson et al., 2012]).

Les thérapies cognitives basées sur la pleine conscience, également axées sur l'acceptation des pertes cognitives, montreraient également des modifications fonctionnelles encourageantes au niveau cérébral lors des tâches de mémoire de travail (Melis et al., 2025). Cependant, une autre étude ne confirme les effets de ces thérapies comparés à une intervention d'activité physique ou une liste d'attente (Melis et al., 2023). Les thérapies cognitives basées sur la pleine conscience sont principalement utilisées pour réduire le stress. En revanche, d'autres techniques visant à diminuer le stress pour améliorer les performances cognitives, telles que le Qigong, le Tai Chi, le yoga, la thérapie de soutien, et l'acupuncture, ont montré peu (Yang & Ah, 2024) ou pas d'effets (Zeng et al., 2020).

Activité physique

Les effets de l'exercice physique sur le cerveau pourraient constituer une ressource intéressante quant à la prise en charge du CRCI. En effet, la pratique d'une activité physique favoriserait la neurogenèse (naissance de nouveaux neurones) dans l'hippocampe, une structure de l'encéphale impliquée dans la mémoire (Fabel & Kempermann, 2008). De plus, l'exercice a montré des effets significatifs concernant l'amélioration de l'état physique des survivants d'un cancer, ainsi que de leurs niveaux d'anxiété, de stress et de dépression. L'intérêt que l'activité physique suscite pour restaurer l'efficience cognitive chez des survivants d'un cancer est donc logique. Cependant, les bénéfices qu'auraient ce type d'interventions demeurent limités. Lors d'essais cliniques comprenant de l'activité physique, quelle qu'elle soit, les survivants ont constaté un changement positif de leur statut cognitif (Hartman et al., 2024). En revanche, l'exercice n'a pas eu d'impact sur les déficits mesurés au moyen de tests (Hartman et al., 2024). Dans une autre étude, la pratique de l'aérobie n'a

pas révélé de différence significative comparé au fait d'être sur liste d'attente (Brunet et al., 2025). L'activité physique, seule, s'avérerait donc inefficace pour traiter le CRCI (Campbell et al., 2020). L'interprétation de ces résultats se trouve néanmoins limité par la diversité des exercices proposés ainsi que par l'absence d'évaluation cognitive appropriée. Davantage d'études, plus rigoureusement contrôlées, sont requises pour tirer des conclusions fiables s'agissant des effets de l'activité physique sur le CRCI (Yang & Ah, 2024). Elle est toutefois valide pour atténuer les répercussions des facteurs confondants tels que la fatigue, l'anxiété et la dépression (Campbell et al., 2019). Pour cette raison, elle est préconisée pour améliorer l'état psychologique et physique des survivants du cancer (Campbell et al., 2019). Les recommandations concernant la pratique d'une activité physique régulière ne sont toutefois pas spécifiques au cancer. En effet, l'Organisation Mondiale de la Santé préconise de l'exercice quotidien pour l'ensemble de la population (*WHO Guidelines on Physical Activity and Sedentary Behaviour*, 2020). Tandis que la sédentarité et l'inactivité augmentent les risques de maladies cardiovasculaires, l'exercice physique protège de l'apparition de ces pathologies et réduit le risque de maladies psychiatriques (*WHO Guidelines on Physical Activity and Sedentary Behaviour*, 2020).

Approches combinées

Yang et Ah (2024) suggèrent que la combinaison des exercices physiques et de la rééducation cognitive pourrait avoir des effets prometteurs. Cependant, les échantillons de ces études étaient trop réduits pour permettre une généralisation des résultats. En dépit du manque de preuves solides, une étude Delphi sur la détection et la rééducation du CRCI indique que l'approche multimodale est néanmoins recommandée par le panel d'experts consultés (Duivon et al., 2024).

3.2 Pistes d'intervention en orthophonie

Étant donné la prévalence des difficultés linguistiques associées au CRCI, le rôle de l'orthophoniste dans sa prise en charge est largement justifié. Une étude menée par Geels et al. (2024b) a évalué les connaissances des orthophonistes américains concernant le CRCI et leur pratique clinique. Trois quarts des participants ont estimé que la prise en charge du CRCI relevait des compétences des orthophonistes, et 65% ont indiqué avoir déjà prodigué des soins à cette population. En France, les orthophonistes sont les professionnels de santé référents pour la rééducation et la réadaptation des troubles cognitivo-linguistiques, c'est-à-dire des troubles cognitifs affectant le langage, la parole et la communication. Chez des individus ayant un diagnostic de cancer (hors tumeur du cerveau), les troubles du langage et de la communication pourraient davantage être liés à un déficit exécutif-attentionnel qu'à un trouble du langage en tant que tel. Néanmoins, l'anomie observée chez ces patients génère de l'anxiété durant leurs interactions sociales (Geels et al., 2024a ; Kanaskie & Loeb, 2015). Dès lors, il paraît pertinent d'intégrer une intervention orthophonique aux soins de support oncologiques.

Toutefois, aucune étude scientifique n'a encore exploré les effets d'une prise en charge orthophonique dans le cadre du CRCI. Il est possible que le développement de méthodes orthophoniques soit ralentit par le manque de documentation sur les troubles linguistiques induits par un cancer et ses traitements. Par conséquent, il n'existe pas à ce jour de rééducation orthophonique validée pour cette population. Compte tenu des mécanismes cognitifs et des processus inflammatoires sous-jacents au CRCI, il pourrait être intéressant d'appliquer des méthodes de soin orthophonique dont l'efficacité a été démontrée sur des populations présentant des difficultés similaires qui sont associées à des inflammations cérébrales, telles que des personnes atteintes de sclérose en plaques ou de COVID long. Une intervention précoce longitudinale, combinant rééducation cognitive et réadaptation, pourrait alors être envisagée. L'objectif serait de favoriser au mieux le transfert des résultats, étant donné que le trouble se manifeste principalement dans la vie quotidienne. Compte tenu de la complexité du CRCI et de ses co-facteurs, cette rééducation devrait s'inscrire dans une approche clinique holistique, c'est-à-dire impliquer les différents acteurs de soin des patients.

4 Limites et perspectives

Cette section présente les principales limites rencontrées au cours de cette thèse, ainsi que les perspectives pour les recherches à venir. Ces travaux préliminaires ouvrent la voie à d'autres recherches sur les caractéristiques de parole et du comportement oculomoteur du CRCI.

D'abord, il convient de souligner le caractère exploratoire de cette étude. Cette recherche visait un objectif écologique, c'est-à-dire une analyse de la parole et de ses corrélats cognitifs dans un contexte aussi naturel que possible. Toutefois, cet objectif s'accompagne de certaines limites méthodologiques, notamment le faible contrôle des paramètres temporels, tels que la durée du discours, et des variations linguistiques dans la production discursive elle-même. En effet, contrairement à des paradigmes expérimentaux plus contrôlés, tels que ceux qui analysent des mots isolés ou de la lecture, la parole en un contexte de discours peut être influencée par de multiples facteurs (émotions, environnement, etc.). Le manque de contrôle des paramètres discursifs pourrait introduire une variabilité supplémentaire qui complexifie l'interprétation des mécanismes cognitifs sous-tendant les paramètres prosodiques.

Par ailleurs, la taille de la cohorte constitue un autre facteur limitant majeur. En effet, cette étude pilote repose sur un nombre restreint de participants, ce qui empêche toute généralisation des résultats à une population plus large de survivants du cancer. Cette petite taille d'échantillon ne permet pas de tirer des conclusions définitives. Il n'est notamment pas possible d'obtenir une taille d'effet pour des variables prosodiques avec un nombre limité de participants. Pour renforcer la fiabilité et l'applicabilité de nos résultats, cette recherche mériterait d'être conduite en incluant un plus grand nombre de participants et en obtenant plus d'échantillons de parole par participant.

De plus, cette thèse s'est exclusivement concentrée sur les survivantes d'un cancer du sein. Cette approche rend difficile l'extrapolation des résultats à d'autres types de cancers. Les différences

cognitives observées dans cette population spécifique peuvent être influencés par des facteurs propres à cette pathologie, tels que le type de traitement ou les caractéristiques physiopathologiques du cancer du sein. Le manque de diversité clinique limite pour l'instant la portée des conclusions de nos travaux. Des recherches futures devraient inclure des participants ayant suivi divers types de traitements pour mieux comprendre les variations entre les différents cancers. L'objectif sera de déterminer si les marqueurs prosodiques identifiés dans cette thèse sont également présents dans des populations atteintes d'un autre type de cancer.

Enfin, une limite importante de cette étude concerne l'absence de contrôle des traitements administrés aux participantes ayant eu un cancer du sein. Le petit nombre de participantes a rendu la prise en compte de cette variable difficile, pourtant essentielle, car les effets des traitements sur les fonctions cognitives peuvent être multiples et variés. L'impact des chimiothérapies, radiothérapies ou autres traitements spécifiques pourrait interférer avec les symptômes cognitifs observés, mais il est difficile de démêler ces effets en l'absence d'un échantillon plus large et de données plus précises sur les traitements administrés. Pour contourner cette limite, il serait pertinent de mener une étude multicentrique impliquant un plus grand nombre de participantes. Une telle approche permettrait de collecter des données plus diversifiées et de mieux contrôler les effets des différents traitements sur les troubles cognitifs, tout en augmentant la représentativité des résultats.

Compte tenu de la chronologie du cancer, il semble nécessaire d'évaluer la parole et les mouvements oculaires à plusieurs points temporels tout au long de l'évolution de la maladie et de sa rémission ou guérison. Une étude en cours menée par notre équipe évalue la parole de participantes atteintes d'un cancer du sein à deux moments temporels : immédiatement après l'arrêt des traitements et à distance, plus d'un an après l'arrêt des traitements, afin d'observer d'éventuelles modifications. À long terme, cette technique pourra aider à identifier les patients à risque de développer des troubles persistants.

Un objectif à long terme consiste à développer un outil combinant la reconnaissance automatique de la parole et un modèle de classification pour permettre une identification rapide du CRCI et des différents profils de patients. Il est fréquent que les scores obtenus par les patients ne se situent pas dans la plage des seuils pathologiques. Cependant, une détérioration cognitive peut être observée lorsque ces scores sont comparés au fil du temps (Stewart et al., 2008). L'établissement d'une ligne de base avant le début des traitements est essentiel pour évaluer l'étendue des changements cognitifs induits par le cancer et ses traitements (Janelsins et al., 2011). Cette démarche peut être réalisée par le recueil d'un échantillon de parole. L'analyse de la parole pourrait ainsi faciliter le suivi des évolutions cognitives. Cet outil, facile et rapide à utiliser, pourrait être employé par divers professionnels de soins, tels que l'oncologue, le médecin en rééducation fonctionnelle, les infirmiers et les rééducateurs. Il serait également possible de l'introduire à différents points temporels, avant les traitements, voire avant même le diagnostic, par exemple avant la biopsie.

Conclusion

Les cancers sont des maladies hétérogènes et largement répandues. Grâce à l'amélioration des connaissances et des techniques médicales, certains types de cancer présentent de meilleures chances de survie. La qualité de vie après cancer constitue désormais un enjeu majeur. En effet, les cancers peuvent provoquer de nombreuses séquelles dermatologiques, psychologiques, mais aussi cognitives. Ce dernier trouble est dénommé *Cancer-Related Cognitive Impairment* (CRCI).

Le CRCI est un effet secondaire des cancers et de leurs traitements. Les patients rapportent des difficultés d'ordre mnésique, attentionnel, exécutif et linguistique. Ces troubles peuvent affecter leur vie quotidienne de manière considérable. Aux côtés des problèmes de mémoire, le manque du mot est l'un des symptômes cognitifs les plus fréquents. Le CRCI limite la reprise du travail et les interactions sociales d'une partie de la population oncologique.

Le CRCI est généralement perçu pendant ou juste après les traitements (c.-à-d., chirurgie, chimiothérapie, radiothérapie, hormonothérapie). Il peut parfois se déclarer avant l'administration des traitements. La plupart du temps, les symptômes diminuent au cours des premières années suivant l'arrêt des traitements. Cependant, certains patients présentent des symptômes persistants qui peuvent durer jusqu'à dix ans après la fin des traitements.

L'origine physiopathologique de ce trouble demeure mal identifiée. Les recherches récentes suggèrent que le cancer lui-même provoquerait des inflammations neurologiques qui entraîneraient à leur tour des difficultés cognitives. Le rôle des traitements, tels que la chimiothérapie, la radiothérapie et l'hormonothérapie, est bien établi dans l'aggravation de ce trouble cognitif. D'autres facteurs, comme la fatigue, l'anxiété et la dépression, sont souvent associés au CRCI et contribuent à la détérioration de ces symptômes. Toutefois, ils ne peuvent à eux seuls expliquer le développement de ce trouble.

Ces difficultés sont souvent mises en évidence lors des entretiens avec les patients, ainsi que par des questionnaires standardisés comme le FACT-Cog. Ces questionnaires montrent un large éventail de difficultés se manifestant au quotidien. Bien qu'utiles pour détecter la présence d'une plainte cognitive, ils ne permettent pas de préciser la nature de ces difficultés ni d'identifier leur origine cognitive.

Au contraire, les tests neuropsychologiques indiquent des résultats plus mitigés. En effet, les scores des patients se situent généralement dans la fourchette des normes. Ce décalage entre les difficultés perçues et les scores normaux peut être dû à un manque de sensibilité des tests ou

à une divergence entre les performances évaluées de manière subjective et objective. Le manque d'outils diagnostiques adaptés empêche certains patients d'accéder à des soins appropriés et ralentit la compréhension des mécanismes du CRCI. De nouvelles approches sont donc nécessaires pour mieux comprendre ce trouble.

Le langage est une fonction cognitive complexe qui interagit avec d'autres fonctions, comme la mémoire, l'attention, les fonctions exécutives et les fonctions motrices. Si l'un des mécanismes nécessaires au fonctionnement normal du langage est déficitaire, la production de parole s'en retrouvera affectée. De fait, la parole continue, qui consiste en l'enchaînement de phonèmes, comporte de nombreux indices prosodiques pouvant révéler des difficultés cognitives.

La production de la parole est rarement linéaire. Elle est souvent interrompue par des pauses plus ou moins longues, qui perturbent sa fluidité. Toutefois, toutes ces interruptions ne signalent pas nécessairement des difficultés cognitives. L'analyse détaillée des interruptions anormale permet de repérer certaines populations présentant des troubles cognitifs. La recherche lexicale peut être laborieuse chez l'ensemble de la population générale. Dans ce cas, elle entraîne souvent des pauses avant les mots recherchés. Lorsque ces difficultés lexicales sont relativement fréquentes, elles sont appellées manque du mot ou anomie. L'anomie peut résulter de trois déficits cognitifs distincts : d'abord la mémoire et la dégradation du système sémantique, comme chez les patients atteints de la maladie d'Alzheimer ; ensuite, les fonctions exécutives et les difficultés d'accès à la forme phonologique du mot ; enfin, un déficit attentionnel qui perturberait l'inhibition des concepts co-activés lors de la recherche lexicale. Un des défis actuels de l'analyse de la parole réside dans l'absence de méthode ou de critères linguistiques faisant consensus pour évaluer les troubles cognitifs. Bien que l'analyse de la parole puisse indiquer des variations cognitives, il est nécessaire d'utiliser une technique complémentaire pour identifier les processus impliqués dans le manque du mot et en comprendre l'origine cognitive.

À ce titre, l'oculométrie est une excellente technique complémentaire pour étudier la parole et la cognition. En effet, des recherches ont montré que le regard se dirige généralement vers des éléments ayant un lien sémantique avec les mots entendus. Sur un versant production, le paradigme des mondes visuels a révélé le processus incrémental de la parole. Autrement dit, un locuteur commence à articuler les phonèmes d'un mot avant d'avoir complètement encodé la suite de son message. Ce processus expliquerait la présence d'interruptions dans la parole. De plus, un mot moins fréquent entraîne un temps de récupération et de recherche lexicale plus long. Enfin, les éléments visuels représentant un événement de sens (ou une action) sont regardés en suivant l'ordre de leur prononciation. Le calcul du délai entre les fixations sur un élément visuel et sa dénomination pourrait ainsi renseigner sur la facilité de récupération de ce mot.

La vision joue un rôle majeur dans l'orientation de l'attention. Deux théories principales s'opposent à ce sujet. L'une propose que notre regard est attiré par les éléments proéminents d'une scène visuelle, comme une couleur vive (théorie *image-guidance*). L'autre suggère que la vision serait guidée par notre recherche de sens (théorie *cognitive-guidance*). La création de cartes de

chaleur basées sur la saillance et la sémantique, comparées aux fixations d'un participant, soutient l'hypothèse de la recherche de sens. Cette méthode pourrait être utile pour étudier d'éventuels troubles attentionnels chez des patients atteints de cancer.

Le **premier article** de cette thèse, Investiguer les troubles subtils de la cognition, a passé en revue la littérature sur les marqueurs potentiels des troubles cognitifs subtils. L'objectif était de recenser les paramètres actuellement utilisés et d'évaluer leur potentiel pour détecter un CRCI.

Le **deuxième article**, Dévoiler les marqueurs prosodiques du CRCI, a estimé le pouvoir discriminant de quatorze paramètres prosodiques présélectionnés. L'objectif était de déterminer la combinaison optimale de paramètres pour classer les participantes atteintes de cancer présentant un CRCI, les survivantes sans plainte et les témoins, à l'aide de modèles d'apprentissage machine.

Le premier objectif de cette thèse était de repérer les paramètres de parole caractéristiques du CRCI. Parmi un ensemble de variables linguistiques, la combinaison de trois marqueurs prosodiques a permis d'identifier les survivantes de cancer présentant une plainte cognitive, celles sans plainte cognitive et les contrôles avec une précision de 73,7%. Les erreurs du modèle ont principalement concerné les survivantes sans plainte cognitive. Ces patientes ont obtenu un score FACT-Cog dans les normes, bien qu'elles aient tout de même signalé des difficultés. Une explication possible est que le CRCI représente un continuum de difficultés, sujet à des variations de performances qui n'ont pas été capturées par le FACT-Cog, mais qui l'ont été par l'analyse de la parole. Deux des trois marqueurs prosodiques étaient liés au silence. Le silence pourrait, plus que les autres paramètres de la parole, refléter les difficultés cognitives, comme cela peut être observé dans d'autres pathologies ou syndromes.

Le **dernier article**, Comprendre l'anomie en contexte de CRCI grâce à l'oculométrie, a testé trois hypothèses quant à l'origine possible du manque du mot dans un contexte de CRCI. La première hypothèse suggérait que le manque du mot serait lié à une dégradation du système sémantique. Si tel était le cas, les patientes auraient passé plus de temps que les contrôles à fixer les éléments visuels dont la dénomination s'est avérée compliquée. La deuxième hypothèse a proposé que le manque du mot était dû à un problème d'accès à sa forme phonologique. Dans ce cas, les patientes auraient présenté un délai de dénomination plus long entre la première fixation d'un élément et sa dénomination, comparé aux contrôles. La troisième hypothèse a envisagé que le manque du mot pourrait être lié à un déficit attentionnel. Les patientes auraient été davantage attirées par les éléments visuellement saillants des images que par ceux porteurs de sens, par rapport aux contrôles.

Le deuxième objectif était de comprendre les mécanismes sous-jacents au CRCI à travers l'étude de l'anomie. Les résultats ont suggéré que le délai de dénomination indiquerait une anomie d'origine exécutive plutôt que mnésique. Quant à l'hypothèse d'un déficit attentionnel, celle-ci n'est pas validée par nos données. Cependant, il semble difficile de l'écarter, compte tenu des recherches indiquant des déficits attentionnels chez les patients atteints de cancer. Les supports imagés utilisés dans cette thèse ont probablement manqué de complexité pour capturer ces difficultés. D'autres recherches sur le sujet

sont nécessaires pour rejeter ou infirmer cette hypothèse.

Grâce à la combinaison de l'analyse de la parole et d'une technique d'oculométrie, cette thèse contribue également à l'identification des patients à risque de développer un CRCI. De plus, elle apporte des éléments de compréhension des mécanismes quant à l'origine du manque du mot dans le cadre du CRCI.

D'un point de vue linguistique, notre thèse souligne l'absence de marqueurs homogènes, dont la grande diversité de paramètres et le manque de consensus limite la reproductibilité des études. Elle propose néanmoins une nomenclature qui pourrait servir de référence pour les futures études. Elle offre également une méthodologie rigoureuse concernant l'analyse de la parole qui peut être appliquée en contexte clinique.

Pour conclure, cette thèse pose une première pierre au développement d'outils automatisés de diagnostic du CRCI. Les recherches amorcées dans ce travail devront être poursuivies par d'autres études sur des populations cliniques déclarant des troubles subtils de la cognition. Il serait, par exemple, pertinent d'étudier d'autres populations oncologiques (cancer du poumon, cancer du testicule, etc.), ainsi que d'autres pathologies ou syndromes présentant des symptômes similaires (séquelles d'AVC mineurs, séquelles de COVID long, etc.).

Bibliographie

Aaronson, N. K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N. J., Filiberti, A., Flechtner, H., Fleishman, S. B., Haes, J. C. J. M. D., Kaasa, S., Klee, M., Osoba, D., Razavi, D., Rofe, P. B., Schraub, S., Sneeuw, K., Sullivan, M., & Takeda, F. (1993). The European Organization for Research and Treatment of Cancer QLQ-C30 : A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology. *JNCI Journal of the National Cancer Institute*, *85*(5), 365-376. https://doi.org/10.1093/jnci/85.5.365

Ahles, T. A. (2025, juin). Application of Complex Systems Theory to Cancer-Related Cognitive Change (CRCC) in Patients with Non- CNS Cancers.

Ahles, T. A., & Root, J. C. (2018). Cognitive Effects of Cancer and Cancer Treatments. *Annual Review of Clinical Psychology*, *14*(1), 425-451. https://doi.org/10.1146/annurev-clinpsy-050817-084903

Ahles, T. A., Saykin, A. J., McDonald, B. C., Furstenberg, C. T., Cole, B. F., Hanscom, B. S., Mulrooney, T. J., Schwartz, G. N., & Kaufman, P. A. (2008). Cognitive function in breast cancer patients prior to adjuvant treatment. *Breast Cancer Research and Treatment*, *110*(1), 143-152. https://doi.org/10.1007/s10549-007-9686-5

Ahles, T. A., Saykin, A. J., McDonald, B. C., Li, Y., Furstenberg, C. T., Hanscom, B. S., Mulrooney, T. J., Schwartz, G. N., & Kaufman, P. A. (2010). Longitudinal Assessment of Cognitive Changes Associated With Adjuvant Treatment for Breast Cancer : Impact of Age and Cognitive Reserve. *Journal of Clinical Oncology*, *28*(29), 4434-4440. https://doi.org/10.1200/JCO.2009.27.0827

Ahles, T. A., Saykin, A. J., Noll, W. W., Furstenberg, C. T., Guerin, S., Cole, B., & Mott, L. A. (2003). The relationship of APOE genotype to neuropsychological performance in long-term cancer survivors treated with standard dose chemotherapy. *Psycho-Oncology*, *12*(6), 612-619. https://doi.org/10.1002/pon.742

Aita, S. L., Beach, J. D., Taylor, S. E., Borgogna, N. C., Harrell, M. N., & Hill, B. D. (2019). Executive, language, or both ? An examination of the construct validity of verbal fluency measures. *Applied Neuropsychology : Adult*, *26*(5), 441-451. https://doi.org/10.1080/23279095.2018.1439830

Albuquerque, L., Valente, A. R. S., Teixeira, A., Figueiredo, D., Sa-Couto, P., & Oliveira, C. (2021). Association between acoustic speech features and non-severe levels of anxiety and depression symptoms across lifespan (E. Van Exel, Éd.). *PLOS ONE*, *16*(4), e0248842. https://doi.org/10.1371/journal.pone.0248842

American Psychiatric Association. (2022, mars). *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5-TR). https://doi.org/10.1176/appi.books.9780890425787

Amidi, A., & Wu, L. M. (2019). Structural brain alterations following adult non-CNS cancers : a systematic review of the neuroimaging literature. *Acta Oncologica*, *58*(5), 522-536. https://doi.org/10.1080/0284186X.2018.1563716

Anderson, D. E., Holstein, S. A., & Kedar, S. (2020). Visual Pathway Degeneration in Chemotherapy-Related Neurotoxicity : A Review and Directions for Future Research. *Neuro-Ophthalmology*, *44*(3), 139-147. https://doi.org/10.1080/01658107.2019.1702703

Andreotti, C., Root, J. C., Ahles, T. A., McEwen, B. S., & Compas, B. E. (2015). Cancer, coping, and cognition : a model for the role of stress reactivity in cancer-related cognitive decline : Cancer, coping, and cognition. *Psycho-Oncology*, *24*(6), 617-623. https://doi.org/10.1002/pon.3683

Andryszak, P., Wiłkość, M., Żurawski, B., & Izdebski, P. (2017). Verbal fluency in breast cancer patients treated with chemotherapy. *Breast Cancer*, *24*(3), 376-383. https://doi.org/10.1007/s12282-016-0713-4

Angelopoulou, G., Kasselimis, D., Goutsos, D., & Potagas, C. (2024). A Methodological Approach to Quantifying Silent Pauses, Speech Rate, and Articulation Rate across Distinct Narrative Tasks : Introducing the Connected Speech Analysis Protocol (CSAP). *Brain Sciences*, *14*(5), 466. https://doi.org/10.3390/brainsci14050466

Angelopoulou, G., Kasselimis, D., Makrydakis, G., Varkanitsa, M., Roussos, P., Goutsos, D., Evdokimidis, I., & Potagas, C. (2018). Silent pauses in aphasia. *Neuropsychologia*, *114*, 41-49. https://doi.org/10.1016/j.neuropsychologia.2018.04.006

Antonucci, S., & Reilly, J. (2008). Semantic Memory and Language Processing : A Primer. *Seminars in Speech and Language*, *29*(1), 005-017. https://doi.org/10.1055/s-2008-1061621

Areklett, E. W., Andersson, S., Fagereng, E., Bruheim, K., Stubberud, J., & Lindemann, K. (2024). Cognitive impairment in cervical cancer survivors—Exploring the discrepancy between subjective and objective assessment. *Psycho-Oncology*, *33*(2), e6300. https://doi.org/10.1002/pon.6300

Armitage, S. G. (1946). Trail Making Test. https://doi.org/10.1037/t00757-000

Arnaboldi, P., Lucchiari, C., Santoro, L., Sangalli, C., Luini, A., & Pravettoni, G. (2014). PTSD symptoms as a consequence of breast cancer diagnosis : clinical implications. *SpringerPlus*, *3*(1), 392. https://doi.org/10.1186/2193-1801-3-392

Arnaboldi, P., Riva, S., Crico, C., & Pravettoni, G. (2017). A systematic literature review exploring the prevalence of post-traumatic stress disorder and the role played by stress and traumatic stress in breast cancer diagnosis and trajectory. *Breast Cancer : Targets and Therapy*, *Volume 9*, 473-485. https://doi.org/10.2147/BCTT.S111101

Ash, S., & Grossman, M. (2015, février). Why Study Connected Speech Production ? In R. M. Willems (Éd.), *Cognitive Neuroscience of Natural Language Use* (1re éd., p. 29-58). Cambridge University Press. https://doi.org/10.1017/CBO9781107323667.003

Aspelund, S. G., Halldorsdottir, T., Agustsson, G., Sigurdardottir Tobin, H. R., Wu, L. M., Amidi, A., Johannsdottir, K. R., Lutgendorf, S. K., Telles, R., Daly, H. F., Sigurdardottir, K.,

Valdimarsdottir, H. B., & Baldursdottir, B. (2024). Biological and psychological predictors of cognitive function in breast cancer patients before surgery. *Supportive Care in Cancer*, *32*(1), 88. https://doi.org/10.1007/s00520-023-08282-5

Astésano, C. (2016, juillet). Prosodic characteristics of Reference French. In S. Detey, J. Durand, B. Laks & C. Lyche (Éd.), *Varieties of Spoken French* (1re éd., p. 68-85). Oxford University PressOxford. https://doi.org/10.1093/acprof:oso/9780199573714.003.0006

Baddeley, A. (2003). Working memory and language : an overview. *Journal of Communication Disorders*, *36*(3), 189-208. https://doi.org/10.1016/S0021-9924(03)00019-4

Balogh, R., Imre, N., Gosztolya, G., Hoffmann, L., Pákáski, M., & Kálmán, J. (2023). The Role of Silence in Verbal Fluency Tasks – A New Approach for the Detection of Mild Cognitive Impairment. *Journal of the International Neuropsychological Society*, *29*(1), 46-58. https://doi.org/10.1017/S1355617721001454

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Baum, S. R., & Pell, M. D. (1999). The neural bases of prosody : Insights from lesion studies and neuroimaging. *Aphasiology*, *13*(8), 581-608. https://doi.org/10.1080/026870399401957

Beattie, G. W., & Butterworth, B. (1979). Contextual Probability and Word Frequency as Determinants of Pauses and Errors in Spontaneous Speech. *Language and Speech*, *22*(3), 201-211. https://doi.org/10.1177/002383097902200301

Becker, J. T., Boller, F., Lopez, O. L., Saxton, J., & McGonigle, K. L. (1994). The Natural History of Alzheimer's Disease : Description of Study Cohort and Accuracy of Diagnosis. *51*(6), 585-594.

Bell, M. J., Terhorst, L., & Bender, C. M. (2013). Psychometric Analysis of the Patient Assessment of Own Functioning Inventory in Women With Breast Cancer. *Journal of Nursing Measurement*, *21*(2), 320-334. https://doi.org/10.1891/1061-3749.21.2.320

Belz, M., & Trouvain, J. (2019). Are 'Silent' Pauses Always Silent ? *9th International Congress of Phonetic Sciences (ICPhS)*.

Belz, M. (2023). Defining Filler Particles : A Phonetic Account of the Terminology, Form, and Grammatical Classification of "Filled Pauses". *Languages*, *8*(1), 57. https://doi.org/10.3390/languages8010057

Ben-David, B. M., Multani, N., Shakuf, V., Rudzicz, F., & Van Lieshout, P. H. H. M. (2016). Prosody and Semantics Are Separate but Not Separable Channels in the Perception of Emotional Speech : Test for Rating of Emotions in Speech. *Journal of Speech, Language, and Hearing Research*, *59*(1), 72-89. https://doi.org/10.1044/2015_JSLHR-H-14-0323

Bender, C. M., Sereika, S. M., Berga, S. L., Vogel, V. G., Brufsky, A. M., Paraska, K. K., & Ryan, C. M. (2006). Cognitive impairment associated with adjuvant therapy in breast cancer. *Psycho-Oncology*, *15*(5), 422-430. https://doi.org/10.1002/pon.964

Benedict, R. H. B., Schretlen, D., Groninger, L., & Brandt, J. (1991). Hopkins Verbal Learning Test–Revised. https://doi.org/10.1037/t10851-000

Benton, A. L., Hamsher, K. D., & Sivan, A. B. (1994). Multilingual Aphasia Examination.

Bernstein, L. J., Edelstein, K., Sharma, A., & Alain, C. (2021). Chemo-brain : An activation likelihood estimation meta-analysis of functional magnetic resonance imaging studies. *Neuroscience & Biobehavioral Reviews*, *130*, 314-325. https://doi.org/10.1016/j.neubiorev.2021.08.024

Bernstein, L. J., McCreath, G. A., Komeylian, Z., & Rich, J. B. (2017). Cognitive impairment in breast cancer survivors treated with chemotherapy depends on control group type and cognitive domains assessed : A multilevel meta-analysis. *Neuroscience & Biobehavioral Reviews*, *83*, 417-428. https://doi.org/10.1016/j.neubiorev.2017.10.028

Betz, S., Bryhadyr, N., Türk, O., & Wagner, P. (2023). Cognitive Load Increases Spoken and Gestural Hesitation Frequency. *Languages*, *8*(1), 71. https://doi.org/10.3390/languages8010071

Bézy, C., Renard, A., & Pariente, J. (2016). *GREMOTS Batterie d'évaluation Des Troubles Du Langage Dans Les Maladies Neurodégénératives*. De Boeck Supérieur.

Bigi, B. (2015). SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech. *the Phonetician" - International Society of Phonetic Science*, *111–112*(2015-1-11), 5-69.

Bigi, B., & Priego-Valverde, B. (2022). The Automatic Search for Sounding Segments of SPPAS : Application to Cheese ! Corpus. In Z. Vetulani, P. Paroubek & M. Kubis (Éd.), *Human Language Technology. Challenges for Computer Science and Linguistics* (p. 16-27, T. 13212). Springer International Publishing. https://doi.org/10.1007/978-3-031-05328-3_2

Biglia, N., Bounous, V., Malabaila, A., Palmisano, D., Torta, D., D'Alonzo, M., Sismondi, P., & Torta, R. (2012). Objective and self-reported cognitive dysfunction in breast cancer women treated with chemotherapy : a prospective study : Chemotherapy-induced cognitive dysfunction. *European Journal of Cancer Care*, *21*(4), 485-492. https://doi.org/10.1111/j.1365-2354.2011.01320.x

Birn, R. M., Kenworthy, L., Case, L., Caravella, R., Jones, T. B., Bandettini, P. A., & Martin, A. (2010). Neural systems supporting lexical search guided by letter and semantic category cues : A self-paced overt response fMRI study of verbal fluency. *NeuroImage*, *49*(1), 1099-1107. https://doi.org/10.1016/j.neuroimage.2009.07.036

Blignaut, P. (2009). Fixation identification : The optimum threshold for a dispersion algorithm. *Attention, Perception, & Psychophysics*, *71*(4), 881-895. https://doi.org/10.3758/APP.71.4.881

Boele, F. W., Schilder, C. M., De Roode, M.-L., Deijen, J. B., & Schagen, S. B. (2015). Cognitive functioning during long-term tamoxifen treatment in postmenopausal women with breast cancer. *Menopause*, *22*(1), 17-25. https://doi.org/10.1097/GME.0000000000000271

Boersma, P., & Weenink, D. (2024). Praat : Doing Phonetics by Computer.

Bolton, G., & Isaacs, A. (2018). Women's experiences of cancer-related cognitive impairment, its impact on daily life and care received for it following treatment for breast cancer. *Psychology, Health & Medicine*, *23*(10), 1261-1274. https://doi.org/10.1080/13548506.2018.1500023

Bóna, J. (2024). Pausing and fluency in speech of patients with relapsing-remitting multiple sclerosis. *Clinical Linguistics & Phonetics*, *38*(4), 332-344. https://doi.org/10.1080/02699206.2023.2223347

Bonomi, A. E., Cella, D. F., Hahn, E. A., Bjordal, K., Sperner-Unterweger, B., Gangeri, L., Bergman, B., Willems-Groot, J., Hanquet, P., & Zittoun, R. (1996). Multilingual translation of the Functional Assessment of Cancer Therapy (FACT) quality of life measurement system. *Quality of Life Research*, *5*(3), 309-320. https://doi.org/10.1007/BF00433915

Boomer, D. S. (1965). Hesitation and Grammatical Encoding. *Language and Speech*, *8*(3), 148-158. https://doi.org/10.1177/002383096500800302

Boomer, D. S. (1970). Psycholinguistics ; experiments in spontaneous speech. *Lingua*, *25*, 152-164. https://doi.org/10.1016/0024-3841(70)90028-8

Borghesani, V., Pedregosa, F., Buiatti, M., Amadon, A., Eger, E., & Piazza, M. (2016). Word meaning in the ventral visual path : a perceptual to conceptual gradient of semantic coding. *NeuroImage*, *143*, 128-140. https://doi.org/10.1016/j.neuroimage.2016.08.068

Bortfeld, H., Leon, S., Bloom, J., Schober, M., & Brennan, S. (2001). Disfluency Rates in Conversation : Effects of Age, Relationship, Topic, Role, and Gender. *Language and Speech*, *44*(2), 123-147.

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*. https://doi.org/10.3389/fpsyg.2017.00269

Boué, A., Joly, F., Lequesne, J., & Lange, M. (2024). Does hormone therapy impact cognition in patients with prostate cancer ? A systematic review and meta-analysis. *Cancer*, *130*(12), 2120-2138. https://doi.org/10.1002/cncr.35210

Bourke, R. S., West, C. R., Chheda, G., & Tower, D. B. (1973). Kinetics of Entry and Distribution of 5-Fluorouracil in Cerebrospinal Fluid and Brain Following Intravenous Injection in a Primate. *Cancer Research*, *33*(7), 1735-1746.

Boykoff, N., Moieni, M., & Subramanian, S. K. (2009). Confronting chemobrain : an in-depth look at survivors' reports of impact on work, social networks, and health care response. *Journal of Cancer Survivorship*, *3*(4), 223. https://doi.org/10.1007/s11764-009-0098-x

Bradski, G. (2000). The OpenCV Library. *Doctor Dobbs Journal of Software Tools*, *11*(25).

Bray, V. J., Dhillon, H. M., Bell, M. L., Kabourakis, M., Fiero, M. H., Yip, D., Boyle, F., Price, M. A., & Vardy, J. L. (2017). Evaluation of a Web-Based Cognitive Rehabilitation Program in Cancer Survivors Reporting Cognitive Symptoms After Chemotherapy. *Journal of Clinical Oncology*, *35*(2), 217-225. https://doi.org/10.1200/JCO.2016.67.8201

Brennan, C., Worrall-Davies, A., McMillan, D., Gilbody, S., & House, A. (2010). The Hospital Anxiety and Depression Scale : A diagnostic meta-analysis of case-finding ability. *Journal of Psychosomatic Research*, *69*(4), 371-378. https://doi.org/10.1016/j.jpsychores.2010.04.006

Brookshire, R., & Nicholas, L. (1994). Test-Retest Stability of Measures of Connected Speech in Aphasia. *Clinical Aphasiology*, *22*, 119-133.

Brown, J. S., Amend, S. R., Austin, R. H., Gatenby, R. A., Hammarlund, E. U., & Pienta, K. J. (2023). Updating the Definition of Cancer. *Molecular Cancer Research*, *21*(11), 1142-1147. https://doi.org/10.1158/1541-7786.MCR-23-0411

Brunet, J., Sharma, S., Zadravec, K., Taljaard, M., LeVasseur, N., Srikanthan, A., Bland, K. A., Sabri, E., Collins, B., Hayden, S., Simmons, C., Smith, A. M., & Campbell, K. L. (2025). Aerobic exercise and CogniTIVe functioning in women with breAsT cancEr (ACTIVATE) : A randomized controlled trial. *Cancer*, *131*(1), e35540. https://doi.org/10.1002/cncr.35540

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic Analysis of Discourse in Aphasia : A Review of the Literature. *Clinical Linguistics & Phonetics*, *30*(7), 489-518. https://doi.org/10.3109/02699206.2016.1145740

Bryant, L., Spencer, E., & Ferguson, A. (2017). Clinical Use of Linguistic Discourse Analysis for the Assessment of Language in Aphasia. *Aphasiology*, *31*(10), 1105-1126. https://doi.org/10.1080/02687038.2016.1239013

Burguin, A., Diorio, C., & Durocher, F. (2021). Breast Cancer Treatments : Updates and New Challenges. *Journal of Personalized Medicine*, *11*(8), 808. https : / / doi . org / 10 . 3390 / jpm11080808

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue : What causes word finding failures in young and older adults ? *Journal of Memory and Language*, *30*(5), 542-579. https://doi.org/10.1016/0749-596X(91)90026-G

Bürkner, P.-C. (2021). Bayesian Item Response Modeling in *R* with **brms** and *Stan*. *Journal of Statistical Software*, *100*(5). https://doi.org/10.18637/jss.v100.i05

Campbell, K. L., Winters-Stone, K. M., Wiskemann, J., May, A. M., Schwartz, A. L., Courneya, K. S., Zucker, D. S., Matthews, C. E., Ligibel, J. A., Gerber, L. H., Morris, G. S., Patel, A. V., Hue, T. F., Perna, F. M., & Schmitz, K. H. (2019). Exercise Guidelines for Cancer Survivors : Consensus Statement from International Multidisciplinary Roundtable. *Medicine & Science in Sports & Exercise*, *51*(11), 2375-2390. https://doi.org/10.1249/MSS.0000000000002116

Campbell, K. L., Zadravec, K., Bland, K. A., Chesley, E., Wolf, F., & Janelsins, M. C. (2020). The Effect of Exercise on Cancer-Related Cognitive Impairment and Applications for Physical Therapy : Systematic Review of Randomized Controlled Trials. *Physical Therapy*, *100*(3), 523-542. https://doi.org/10.1093/ptj/pzz090

Campione, E., & Véronis, J. (2002). A Large-Scale Multilingual Study of Silent Pause Duration. *SPeech Prosody*.

Campione, E., & Véronis, J. (2004). Pauses d'hésitation En Français Spontané. *XXV JEP*.

Campione, E., & Véronis, J. (2005). Pauses and Hesitations in French Spontaneous Speech. *Disfluency in Spontaneous Speech (DiSS 2005)*, 43-46.

Candea, M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané. Etude sur un corpus de récits en classe de français.* [Thèse de doctorat]. Université de la Sorbonne nouvelle - Paris III.

Caramazza, A., & Hillis, A. E. (1990). Where Do Semantic Errors Come From ? *Cortex*, *26*(1), 95-122. https://doi.org/10.1016/S0010-9452(13)80077-9

Caramazza, A., & Hillis, A. E. (1991). Lexical organization of nouns and verbs in the brain. *Nature*, *349*(6312), 788-790. https://doi.org/10.1038/349788a0

Carroll, J. E., Van Dyk, K., Bower, J. E., Scuric, Z., Petersen, L., Schiestl, R., Irwin, M. R., & Ganz, P. A. (2019). Cognitive performance in survivors of breast cancer and markers of biological aging. *Cancer*, *125*(2), 298-306. https://doi.org/10.1002/cncr.31777

Cella, D. F., Tulsky, D. S., Gray, G., Sarafian, B., Linn, E., Bonomi, A., Silberman, M., Yellen, S. B., Winicour, P., & Brannon, J. (1993). The Functional Assessment of Cancer Therapy scale : development and validation of the general measure. *Journal of Clinical Oncology*, *11*(3), 570-579. https://doi.org/10.1200/JCO.1993.11.3.570

Chelune, G. J., Heaton, R. K., & Lehman, R. A. W. (1986). Neuropsychological and Personality Correlates of Patients' Complaints of Disability. In G. Goldstein & R. E. Tarter (Éd.), *Advances in Clinical Neuropsychology* (p. 95-126). Springer US. https://doi.org/10.1007/978-1-4613-2211-5_4

Cimprich, B., Visovatti, M., & Ronis, D. L. (2011). The Attentional Function Index—a self-report cognitive measure. *Psycho-Oncology*, *20*(2), 194-202. https://doi.org/10.1002/pon.1729

Clark, H. (2002). Using Uh and Um in Spontaneous Speaking. *Cognition*, *84*(1), 73-111. https://doi.org/10.1016/S0010-0277(02)00017-3

Clarke, N., Barrick, T. R., & Garrard, P. (2021). A Comparison of Connected Speech Tasks for Detecting Early Alzheimer's Disease and Mild Cognitive Impairment Using Natural Language Processing and Machine Learning. *Frontiers in Computer Science*, 44.

Cole, J. (2015). Prosody in context : a review. *Language, Cognition and Neuroscience*, *30*(1-2), 1-31. https://doi.org/10.1080/23273798.2014.963130

Commission Nationale de l'Informatique et des Libertés. (2016). Le règlement général sur la protection des données - RGPD.

Conroy, S. K., McDonald, B. C., Smith, D. J., Moser, L. R., West, J. D., Kamendulis, L. M., Klaunig, J. E., Champion, V. L., Unverzagt, F. W., & Saykin, A. J. (2013). Alterations in brain structure and function in breast cancer survivors : effect of post-chemotherapy interval and relation to oxidative DNA damage. *Breast Cancer Research and Treatment*, *137*(2), 493-502. https://doi.org/10.1007/s10549-012-2385-x

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language. *Cognitive Psychology*, *6*(1), 84-107. https://doi.org/10.1016/0010-0285(74)90005-X

Corley, M., & Hartsuiker, R. J. (2011). Why Um Helps Auditory Word Recognition : The Temporal Delay Hypothesis (M. Perc, Éd.). *PLoS ONE*, *6*(5), e19792. https://doi.org/10.1371/journal.pone.0019792

Costa, D., & Fardell, J. E. (2019). Why Are Objective and Perceived Cognitive Function Weakly Correlated in Patients With Cancer ? *Journal of Clinical Oncology*, *37*(14), 1154-1158. https://doi.org/10.1200/JCO.18.02363

Costa, D., Loh, V., Birney, D., Dhillon, H., Fardell, J., Gessler, D., & Vardy, J. (2018). The Structure of the FACT-Cog v3 in Cancer Patients, Students, and Older Adults. *Journal of Pain and Symptom Management*, *55*(4), 1173-1178. https://doi.org/10.1016/j.jpainsymman.2017.12.486

Crible, L., Degand, L., & Gilquin, G. (2017). The clustering of discourse markers and filled pauses : A corpus-based French-English study of (dis)fluency. *Languages in Contrast*, *17*(1), 69-95. https://doi.org/10.1075/lic.17.1.04cri

Crible, L., Didirková, I., Dodane, C., & Kosmala, L. (2022). Towards an Inclusive System for the Annotation of (Dis)Fluency in Typical and Atypical Speech. *Clinical Linguistics & Phonetics*, *38*(4), 381-398. https://doi.org/10.1080/02699206.2022.2126330

Cristofori, I., Cohen-Zimerman, S., & Grafman, J. (2019). Executive functions. In *Handbook of Clinical Neurology* (p. 197-219, T. 163). Elsevier. https://doi.org/10.1016/B978-0-12-804281-6.00011-2

Cucchiarini, C., Strik, H., & Boves, L. (2002). Quantitative assessment of second language learners' fluency : Comparisons between read and spontaneous speech. *The Journal of the Acoustical Society of America*, *111*(6), 2862-2873. https://doi.org/10.1121/1.1471894

Cuenca, M. J., & Crible, L. (2019). Co-occurrence of discourse markers in English : From juxtaposition to composition. *Journal of Pragmatics*, *140*, 171-184. https://doi.org/10.1016/j.pragma.2018.12.001

Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication*, *71*, 10-49. https://doi.org/10.1016/j.specom.2015.03.004

De Jong, N. H. (2018). Fluency in Second Language Testing : Insights From Different Disciplines. *Language Assessment Quarterly*, *15*(3), 237-254. https://doi.org/10.1080/15434303.2018.1477780

De Jong, N. H. (2023). Fluency in Speaking as a Dynamic Construct. *Language Teaching Research Quarterly*, *37*, 179-187. https://doi.org/10.32038/ltrq.2023.37.09

De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*(2), 385-390. https://doi.org/10.3758/BRM.41.2.385

de La Fuente Garcia, S., Ritchie, C. W., & Luz, S. (2020). Artificial Intelligence, Speech, and Language Processing Approaches to Monitoring Alzheimer's Disease : A Systematic Review. *Journal of Alzheimer's Disease*, *78*(4), 1547-1574. https://doi.org/10.3233/JAD-200888

De Looze, C., Dehsarvi, A., Crosby, L., Vourdanou, A., Coen, R. F., Lawlor, B. A., & Reilly, R. B. (2021). Cognitive and Structural Correlates of Conversational Speech Timing in Mild Cognitive Impairment and Mild-to-Moderate Alzheimer's Disease : Relevance for Early Detection Approaches. *Frontiers in Aging Neuroscience*, *13*, 637404. https://doi.org/10.3389/fnagi.2021.637404

De Ruiter, M. B., Reneman, L., Boogerd, W., Veltman, D. J., Van Dam, F. S., Nederveen, A. J., Boven, E., & Schagen, S. B. (2011). Cerebral hyporesponsiveness and cognitive impairment 10 years after chemotherapy for breast cancer. *Human Brain Mapping*, *32*(8), 1206-1219. https://doi.org/10.1002/hbm.21102

Debien, V., De Caluwé, A., Wang, X., Piccart-Gebhart, M., Tuohy, V. K., Romano, E., & Buisseret, L. (2023). Immunotherapy in breast cancer : an overview of current strategies and perspectives. *npj Breast Cancer*, *9*(1), 7. https://doi.org/10.1038/s41523-023-00508-3

Deeken, J. F., & Löscher, W. (2007). The Blood-Brain Barrier and Cancer : Transporters, Treatment, and Trojan Horses. *Clinical Cancer Research*, *13*(6), 1663-1674. https://doi.org/10.1158/1078-0432.CCR-06-2854

Degano, G., Donhauser, P. W., Gwilliams, L., Merlo, P., & Golestani, N. (2024). Speech prosody enhances the neural processing of syntax. *Communications Biology*, *7*(1), 748. https://doi.org/10.1038/s42003-024-06444-7

Delis, D. C., Freeland, J., Kramer, J. H., & Kaplan, E. (1987). California Verbal Learning Test. https://doi.org/10.1037/t48844-000

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283-321. https://doi.org/10.1037/0033-295X.93.3.283

Delrieu, L., Bouaoun, L., Fatouhi, D. E., Dumas, E., Bouhnik, A.-D., Noelle, H., Jacquet, E., Hamy, A.-S., Coussy, F., Reyal, F., Heudel, P.-E., Bendiane, M.-K., Fournier, B., Michallet, M., Fervers, B., Fagherazzi, G., & Pérol, O. (2021). Patterns of Sequelae in Women with a History of Localized Breast Cancer : Results from the French VICAN Survey. *Cancers*, *13*(5), 1161. https://doi.org/10.3390/cancers13051161

Demos-Davies, K., Lawrence, J., & Seelig, D. (2024). Cancer Related Cognitive Impairment : A Downside of Cancer Treatment. *Frontiers in Oncology*, *14*, 1387251. https://doi.org/10.3389/fonc.2024.1387251

Deprez, S., Amant, F., Smeets, A., Peeters, R., Leemans, A., Van Hecke, W., Verhoeven, J. S., Christiaens, M.-R., Vandenberghe, J., Vandenbulcke, M., & Sunaert, S. (2012). Longitudinal Assessment of Chemotherapy-Induced Structural Changes in Cerebral White Matter and Its Correlation With Impaired Cognitive Functioning. *Journal of Clinical Oncology*, *30*(3), 274-281. https://doi.org/10.1200/JCO.2011.36.8571

Deprez, S., Vandenbulcke, M., Peeters, R., Emsell, L., Smeets, A., Christiaens, M.-R., Amant, F., & Sunaert, S. (2014). Longitudinal Assessment of Chemotherapy-Induced Alterations in Brain Activation During Multitasking and Its Relation With Cognitive Complaints. *Journal of Clinical Oncology*, *32*(19), 2031-2038. https://doi.org/10.1200/JCO.2013.53.6219

Dhooge, E., & Hartsuiker, R. J. (2012). Lexical selection and verbal self-monitoring : Effects of lexicality, context, and time pressure in picture-word interference. *Journal of Memory and Language*, *66*(1), 163-176. https://doi.org/10.1016/j.jml.2011.08.004

Di Stasi, L. L., McCamy, M. B., Macknik, S. L., Mankin, J. A., Hooft, N., Catena, A., & Martinez-Conde, S. (2014). Saccadic Eye Movement Metrics Reflect Surgical Residents' Fatigue. *Annals of Surgery*, *259*(4), 824-829. https://doi.org/10.1097/SLA.0000000000000260

DiCanio, C. (2011). Text Replacement Script for Praat.

Dijkshoorn, A. B. C., Van Stralen, H. E., Sloots, M., Schagen, S. B., Visser-Meily, J. M. A., & Schepers, V. P. M. (2021). Prevalence of cognitive impairment and change in patients with

breast cancer : A systematic review of longitudinal studies. *Psycho-Oncology*, *30*(5), 635-648. https://doi.org/10.1002/pon.5623

Don, S., & Lickley, R. (2015). Uh – I Forgot What I Was Going to Say : How Memory Affects Fluency. *DiSS 2015*.

Downie, F. P., Mar Fan, H. G., Houédé-Tchen, N., Yi, Q., & Tannock, I. F. (2006). Cognitive function, fatigue, and menopausal symptoms in breast cancer patients receiving adjuvant chemotherapy : evaluation with patient interview after formal assessment : PSYCHOSOCIAL SIDE EFFECTS OF ADJUVANT CHEMOTHERAPY FOR BREAST CANCER. *Psycho-Oncology*, *15*(10), 921-930. https://doi.org/10.1002/pon.1035

Duez, D. (1982). Silent and Non-Silent Pauses in Three Speech Styles. *Language and Speech*, *25*(1), 11-28. https://doi.org/10.1177/002383098202500102

Duez, D. (1985). Perception of Silent Pauses in Continuous Speech. *Language and Speech*, *28*(4), 377-389. https://doi.org/10.1177/002383098502800403

Duffau, H. (2003). The role of dominant premotor cortex in language : a study using intraoperative functional mapping in awake patients. *NeuroImage*, *20*(4), 1903-1914. https://doi.org/10.1016/S1053-8119(03)00203-9

Duivon, M., Lange, M., Binarelli, G., Lefel, J., Hardy-Léger, I., Kiasuwa-Mbengi, R., Méric, J.-B., Charles, C., & Joly, F. (2024). Improve the management of cancer-related cognitive impairment in clinical settings : a European Delphi study. *Journal of Cancer Survivorship*, *18*(6), 1974-1997. https://doi.org/10.1007/s11764-023-01436-8

Dumas, A., Vaz Luis, I., Bovagnet, T., El Mouhebb, M., Di Meglio, A., Pinto, S., Charles, C., Dauchy, S., Delaloge, S., Arveux, P., Coutant, C., Cottu, P., Lesur, A., Lerebours, F., Tredan, O., Vanlemmens, L., Levy, C., Lemonnier, J., Mesleard, C., … Menvielle, G. (2020). Impact of Breast Cancer Treatment on Employment : Results of a Multicenter Prospective Cohort Study (CANTO). *Journal of Clinical Oncology*, *38*(7), 734-743. https://doi.org/10.1200/JCO.19.01726

Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums. *Technometrics*, *6*(3), 241-252. https://doi.org/10.1080/00401706.1964.10490181

Durán-Gómez, N., López-Jurado, C. F., Nadal-Delgado, M., Pérez-Civantos, D., Guerrero-Martín, J., & Cáceres, M. C. (2022). Chemotherapy-Related Cognitive Impairment in Patients with Breast Cancer Based on Functional Assessment and NIRS Analysis. *Journal of Clinical Medicine*, *11*(9), 2363. https://doi.org/10.3390/jcm11092363

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, *24*(6), 409-436. https://doi.org/10.1007/BF02143160

Edelman, J. A., Ahles, T. A., Estelle, M. C., Mohr, I., Li, Y., Melara, R., & Root, J. C. (2024). The effect of cancer and cancer treatment on attention control : evidence from anti-saccade performance. *Journal of Cancer Survivorship*. https://doi.org/10.1007/s11764-024-01711-2

Ercoli, L. M., Petersen, L., Hunter, A. M., Castellon, S. A., Kwan, L., Kahn-Mills, B. A., Embree, L. M., Cernin, P. A., Leuchter, A. F., & Ganz, P. A. (2015). Cognitive rehabilitation group

intervention for breast cancer survivors : results of a randomized clinical trial : Cognitive rehabilitation intervention for breast cancer survivors. *Psycho-Oncology*, *24*(11), 1360-1367. https://doi.org/10.1002/pon.3769

Fabel, K., & Kempermann, G. (2008). Physical Activity and the Regulation of Neurogenesis in the Adult and Aging Brain. *NeuroMolecular Medicine*, *10*(2), 59-66. https://doi.org/10.1007/s12017-008-8031-4

Feng, Y., Zhang, X. D., Zheng, G., & Zhang, L. J. (2019). Chemotherapy-induced brain changes in breast cancer survivors : evaluation with multimodality magnetic resonance imaging. *Brain Imaging and Behavior*, *13*(6), 1799-1814. https://doi.org/10.1007/s11682-019-00074-y

Ferguson, R. J., McDonald, B. C., Rocque, M. A., Furstenberg, C. T., Horrigan, S., Ahles, T. A., & Saykin, A. J. (2012). Development of CBT for chemotherapy-related cognitive change : results of a waitlist control trial. *Psycho-Oncology*, *21*(2), 176-186. https://doi.org/10.1002/pon.1878

Ferguson, R. J., Patel, J. V., Orosco, N. M., & Sinagra, C. (2024). Adoption of the International Cognition and Cancer Task Force recommendations for harmonizing studies of cancer-related cognitive impairment : A systematic review of published literature. *Journal of Clinical Oncology*, *42*(16_suppl), e24043-e24043. https://doi.org/10.1200/JCO.2024.42.16_suppl.e24043

Ferlay, J., Ervik, M., Lam, F., Laversanne, M., Colombet, M., Mery, L., Piñeros, M., Znaor, A., Soerjomataram, I., & Bray, F. (2024). Global Cancer Observatory : Cancer Today. Lyon, France : International Agency for Research on Cancer.

Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review*, *100*(2), 233-253. https://doi.org/10.1037/0033-295X.100.2.233

Ferreira, F., & Swets, B. (2002). How Incremental Is Language Production ? Evidence from the Production of Utterances Requiring the Computation of Arithmetic Sums. *Journal of Memory and Language*, *46*(1), 57-84. https://doi.org/10.1006/jmla.2001.2797

Filiou, R.-P., Bier, N., Slegers, A., Houzé, B., Belchior, P., & Brambati, S. M. (2020). Connected speech assessment in the early detection of Alzheimer's disease and mild cognitive impairment : a scoping review. *Aphasiology*, *34*(6), 723-755. https://doi.org/10.1080/02687038.2019.1608502

Fleming, B., Edison, P., & Kenny, L. (2023). Cognitive impairment after cancer treatment : mechanisms, clinical characterization, and management. *BMJ*, e071726. https://doi.org/10.1136/bmj-2022-071726

Fox Tree, J. (1995). The Effects of False Starts and Repetitions on the Processing of Subsequent Words in Spontaneous Speech. *Journal of Memory and Language*, *34*, 709-738.

Frank, E., Hall, M. A., & Witten, I. H. (2016). Online Appendix for "Data Mining : Practical Machine Learning Tools and Techniques". In *The WEKA Workbench* (4th). Morgan Kaufmann.

Galetto, V., Andreetta, S., Zettin, M., & Marini, A. (2013). Patterns of Impairment of Narrative Language in Mild Traumatic Brain Injury. *JOURNAL OF NEUROLINGUISTICS*, *26*(6), 649-661. https://doi.org/10.1016/j.jneuroling.2013.05.004

Ganz, P. A., Kwan, L., Castellon, S. A., Oppenheim, A., Bower, J. E., Silverman, D. H. S., Cole, S. W., Irwin, M. R., Ancoli-Israel, S., & Belin, T. R. (2013). Cognitive Complaints After Breast Cancer Treatments : Examining the Relationship With Neuropsychological Test Performance. *JNCI : Journal of the National Cancer Institute*, *105*(11), 791-801. https://doi.org/10.1093/jnci/djt073

Ganz, P. A., Petersen, L., Castellon, S. A., Bower, J. E., Silverman, D. H., Cole, S. W., Irwin, M. R., & Belin, T. R. (2014). Cognitive Function After the Initiation of Adjuvant Endocrine Therapy in Early-Stage Breast Cancer : An Observational Cohort Study. *Journal of Clinical Oncology*, *32*(31), 3559-3567. https://doi.org/10.1200/JCO.2014.56.1662

Geels, P. L., Flynn, K., Fogarty, K. J., & Lyerla, R. (2024a). Self-Report of Changes in Cognitive-Communication Function and Social Engagement Among Adults With Cancer-Related Cognitive Impairment. *American Journal of Speech-Language Pathology*, *33*(5), 2387-2398. https://doi.org/10.1044/2024_AJSLP-24-00072

Geels, P. L., Flynn, K., Fogarty, K. J., & Lyerla, R. (2024b). Cancer-Related Cognitive Impairment : A Survey of Speech-Language Pathologists' Knowledge, Training, and Attitudes. *Perspectives of the ASHA Special Interest Groups*, *9*(6), 1571-1579. https://doi.org/10.1044/2024_PERSP-24-00054

Giammarile, F., Vidal-Sicart, S., Paez, D., Pellet, O., Enrique, E.-L., Mikhail-Lette, M., Morozova, O., Maria Camila, N. M., Diana Ivonne, R. S., Delgado Bolton, R. C., Valdés Olmos, R. A., & Mariani, G. (2022). Sentinel Lymph Node Methods in Breast Cancer. *Seminars in Nuclear Medicine*, *52*(5), 551-560. https://doi.org/10.1053/j.semnuclmed.2022.01.006

Gibaldi, A., & Sabatini, S. P. (2021). The saccade main sequence revised : A fast and repeatable tool for oculomotor analysis. *Behavior Research Methods*, *53*(1), 167-187. https://doi.org/10.3758/s13428-020-01388-2

Goldman, J.-P., Roekhaut, S., & Simon, A.-C. (2010). Étude Statistique de La Durée Pausale Dans Différents Styles de Parole. *Actes Des 28èmes Journées d'étude Sur La Parole (JEP)*, 161-164.

Goldman-Eisler, F. (1958). The Predictability of Words in Context and the Length of Pauses in Speech. *Quart. J. exp. Psychol.*, *10*, 96-106. https://doi.org/10.1177/002383095800100308

Goldman-Eisler, F. (1961). The Distribution of Pause Durations in Speech. *Language and Speech*, *4*(4), 232-237. https://doi.org/10.1177/002383096100400405

Grice, M., & Kügler, F. (2021). Prosodic Prominence – A Cross-Linguistic Perspective. *Language and Speech*, *64*(2), 253-260. https://doi.org/10.1177/00238309211015768

Gries, S. T. (2015). Some Current Quantitative Problems in Corpus Linguistics and a Sketch of Some Solutions. *Language and Linguistics*, *16*(1), 93-117. https://doi.org/10.1177/1606822X14556606

Griffin, Z. M. (2004). Why Look ? Reasons for Eye Movements Related to Language Production. In J. M. Henderson & F. Ferreira (Éd.), *The Interface of Language, Vision, and Action : Eye Movements and the Visual World* (p. 213-247). Psychology Press.

Griffin, Z. M., & Bock, K. (2000). What the Eyes Say About Speaking. *Psychological Science*, *11*(4), 274-279. https://doi.org/10.1111/1467-9280.00255

Griffin, Z. M., & Spieler, D. H. (2006). Observing the what and when of language production for different age groups by monitoring speakers' eye movements. *Brain and Language*, *99*(3), 272-288. https://doi.org/10.1016/j.bandl.2005.08.003

Grosjean, F., & Deschamps, A. (1972). Analyse Contrastive Des Variables Temporelles En Français et En Anglais. *Phonetica*, *31*, 144-184.

Groupe de Travail Traumatisme Crânio-Cérébral Léger en Ile-de-France. (2021). *Rapport Du Groupe de Travail Traumatisme Crânio-Cérébral Léger En Ile-de-France* (rapp. tech.). Agence Régionale de Santé Île-de-France. ¨Île-de-France.

Gullett, J. M., Cohen, R. A., Yang, G. S., Menzies, V. S., Fieo, R. A., Kelly, D. L., Starkweather, A. R., Jackson-Cook, C. K., & Lyon, D. E. (2019). Relationship of fatigue with cognitive performance in women with early-stage breast cancer over 2 years. *Psycho-Oncology*, *28*(5), 997-1003. https://doi.org/10.1002/pon.5028

Hagoort, P. (2017). The core and beyond in the language-ready brain. *Neuroscience & Biobehavioral Reviews*, *81*, 194-204. https://doi.org/10.1016/j.neubiorev.2017.01.048

Hagoort, P. (2019). The neurobiology of language beyond single-word processing. *Science*, *366*(6461), 55-58. https://doi.org/10.1126/science.aax0289

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data : An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23-34. https://doi.org/10.20982/tqmp.08.1.p023

Hartelius, L., Burge, Å., Johansson, A., Ljungsfors, A., Mattsson, A., Winkworth, A., & Andersen, O. (2004). How Does Fatigue Affect Communication ? *International Journal of MS Care*, *6*(2), 39-51. https://doi.org/10.7224/1537-2073-6.2.39

Hartman, S. J., Marinac, C. R., Natarajan, L., & Patterson, R. E. (2015). Lifestyle factors associated with cognitive functioning in breast cancer survivors : Breast cancer survivors' cognition and lifestyle factors. *Psycho-Oncology*, *24*(6), 669-675. https://doi.org/10.1002/pon.3626

Hartman, S. J., Zablocki, R. W., Tam, R. M., Palmer, B. W., Parker, B. A., Sears, D. D., Ahles, T. A., & Natarajan, L. (2024). Relationship of Physical Activity and Cognitive Functioning among Breast Cancer Survivors : A Cross-Sectional Analysis. *Frontiers in Cognition*, *3*, 1332960. https://doi.org/10.3389/fcogn.2024.1332960

Hartsuiker, R. J., & Notebaert, L. (2010). Lexical Access Problems Lead to Disfluencies in Speech. *Experimental Psychology*, *57*(3), 169-177. https://doi.org/10.1027/1618-3169/a000021

Haute Autorité de Santé. (2018). Patients présentant un trouble neurocognitif associé à la maladie d'Alzheimer ou à une maladie apparentée.

Haywood, D., Dauer, E., Baughman, F., Lawrence, B., Rossell, S., Hart, N., & O'Connor, M. (2023). "Is My Brain Ever Going to Work Fully Again ?" : Challenges and Needs of Cancer Survivors with Persistent Cancer-Related Cognitive Impairment. *Cancers*, *15*(22), 5331. https://doi.org/10.3390/cancers15225331

He, R., Chapin, K., Al-Tamimi, J., Bel, N., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J. P., Alegret, M., Sanabria, A., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (2023). Automated Classification of Cognitive Decline and Probable Alzheimer's Dementia Across Multiple

Speech and Language Domains. *American Journal of Speech-Language Pathology*, *32*(5), 2075-2086. https://doi.org/10.1044/2023_AJSLP-22-00403

Henderson, F. M., Cross, A. J., & Baraniak, A. R. (2019). 'A new normal with chemobrain' : Experiences of the impact of chemotherapy-related cognitive deficits in long-term breast cancer survivors. *Health Psychology Open*, *6*(1), 2055102919832234. https://doi.org/10.1177/2055102919832234

Henderson, J. M., & Hayes, T. R. (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. *Nature Human Behaviour*, *1*(10), 743-747. https://doi.org/10.1038/s41562-017-0208-0

Henderson, J. M., Hayes, T. R., Rehrig, G., & Ferreira, F. (2018). Meaning Guides Attention during Real-World Scene Description. *Scientific Reports*, *8*(1), 13504. https://doi.org/10.1038/s41598-018-31894-5

Henneghan, A., Haley, A. P., & Kesler, S. (2020). Exploring Relationships Among Peripheral Amyloid Beta, Tau, Cytokines, Cognitive Function, and Psychosomatic Symptoms in Breast Cancer Survivors. *Biological Research For Nursing*, *22*(1), 126-138. https://doi.org/10.1177/1099800419887230

Henneghan, A., Van Dyk, K., Kaufmann, T., Harrison, R., Gibbons, C., Heijnen, C., & Kesler, S. R. (2021). Measuring Self-Reported Cancer-Related Cognitive Impairment : Recommendations From the Cancer Neuroscience Initiative Working Group. *JNCI : Journal of the National Cancer Institute*, *113*(12), 1625-1633. https://doi.org/10.1093/jnci/djab027

Hermelink, K., Voigt, V., Kaste, J., Neufeld, F., Wuerstlein, R., Buhner, M., Munzel, K., Rjosk-Dendorfer, D., Grandl, S., Braun, M., Von Koch, F. E., Hartl, K., Hasmuller, S., Bauerfeind, I., Debus, G., Herschbach, P., & Harbeck, N. (2015). Elucidating Pretreatment Cognitive Impairment in Breast Cancer Patients : The Impact of Cancer-related Post-traumatic Stress. *JNCI Journal of the National Cancer Institute*, *107*(7), djv099-djv099. https://doi.org/10.1093/jnci/djv099

Hermelink, K., Bühner, M., Sckopke, P., Neufeld, F., Kaste, J., Voigt, V., Münzel, K., Wuerstlein, R., Ditsch, N., Hellerhoff, K., Rjosk-Dendorfer, D., Braun, M., Von Koch, F. E., Härtl, K., Hasmüller, S., Bauerfeind, I., Debus, G., Herschbach, P., Mahner, S., & Harbeck, N. (2017). Chemotherapy and Post-traumatic Stress in the Causation of Cognitive Dysfunction in Breast Cancer Patients. *JNCI : Journal of the National Cancer Institute*, *109*(10). https://doi.org/10.1093/jnci/djx057

Hermelink, K., Küchenhoff, H., Untch, M., Bauerfeind, I., Lux, M. P., Bühner, M., Manitz, J., Fensterer, V., & Münzel, K. (2010). Two different sides of 'chemobrain' : determinants and nondeterminants of self-perceived cognitive dysfunction in a prospective, randomized, multicenter study. *Psycho-Oncology*, *19*(12), 1321-1328. https://doi.org/10.1002/pon.1695

Hofman, M., Ryan, J. L., Figueroa-Moseley, C. D., Jean-Pierre, P., & Morrow, G. R. (2007). Cancer-Related Fatigue : The Scale of the Problem. *The Oncologist*, *12*(S1), 4-10. https://doi.org/10.1634/theoncologist.12-S1-4

Holle, H., & Rein, R. (2013). The Modified Cohen's Kappa : Calculating Interrater Agreement for Segmentation and Annotation. In H. Lausberg (Éd.), *Understanding Body Movement : A Guide to Empirical Research on Nonverbal Behaviour (With an Introduction to the NEUROGES Coding System)* (p. 261-275). Peter Lang Verlag.

Holliday, N. (2021). Prosody and Sociolinguistic Variation in American Englishes. *Annual Review of Linguistics*, *7*(1), 55-68. https://doi.org/10.1146/annurev-linguistics-031220-093728

Holte, R. C. (1993). Very Simple Classification Rules Perform Well on Most Commonly Used Datasets. *Machine Learning*, *11*(1), 63-90. https://doi.org/10.1023/A:1022631118932

Hommel, M., Miguel, S. T., Naegele, B., Gonnet, N., & Jaillard, A. (2009). Cognitive determinants of social functioning after a first ever mild to moderate stroke at vocational age. *Journal of Neurology, Neurosurgery & Psychiatry*, *80*(8), 876-880. https://doi.org/10.1136/jnnp.2008.169672

Huang, Y., Yu, Z., Wang, Y., Liu, Q., Zhao, K., & Feng, W. (2025). Pilot randomized study of adaptive cognitive training's effect on breast cancer survivors assessed through eye tracking. *Journal of Cancer Survivorship*. https://doi.org/10.1007/s11764-025-01783-8

Huestegge, L. (2011). The role of saccades in multitasking : towards an output-related view of eye movements. *Psychological Research*, *75*(6), 452-465. https://doi.org/10.1007/s00426-011-0352-5

Huettig, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation : semantic competitor effects and the visual world paradigm. *Cognition*, *96*(1), B23-B32. https://doi.org/10.1016/j.cognition.2004.10.003

Huettig, F., & Hartsuiker, R. J. (2008). When you name the pizza you look at the coin and the bread : Eye movements reveal semantic activation during word production. *Memory & Cognition*, *36*(2), 341-360. https://doi.org/10.3758/MC.36.2.341

Huettig, F., & Hartsuiker, R. J. (2010). Listening to yourself is like listening to others : External, but not internal, verbal self-monitoring is based on speech perception. *Language and Cognitive Processes*, *25*(3), 347-374. https://doi.org/10.1080/01690960903046926

Hurria, A., & Lachs, M. (2007). Is cognitive dysfunction a complication of adjuvant chemotherapy in the older patient with breast cancer ? *Breast Cancer Research and Treatment*, *103*(3), 259. https://doi.org/10.1007/s10549-006-9383-9

Hurria, A., Somlo, G., & Ahles, T. (2007). Renaming "Chemobrain". *Cancer Investigation*, *25*(6), 373-377. https://doi.org/10.1080/07357900701506672

Hutchinson, A., Hosking, J., Kichenadasse, G., Mattiske, J., & Wilson, C. (2012). Objective and Subjective Cognitive Impairment Following Chemotherapy for Cancer : A Systematic Review. *Cancer Treatment Reviews*, *38*, 926-934. https://doi.org/http://dx.doi.org/10.1016/j.ctrv.2012.05.002

Inagaki, M., Yoshikawa, E., Matsuoka, Y., Sugawara, Y., Nakano, T., Akechi, T., Wada, N., Imoto, S., Murakami, K., Uchitomi, Y., & and The Breast Cancer Survivors' Brain MRI Database Group. (2007). Smaller regional volumes of brain gray and white matter demonstrated in breast cancer

survivors exposed to adjuvant chemotherapy. *Cancer*, *109*(1), 146-156. https://doi.org/10.1002/cncr.22368

Institut National du Cancer. (2018, juin). *La Vie Cinq Ans Après Un Diagnostic de Cancer* (rapp. tech.).

Institut National du Cancer. (2021). *Stratégie Décennale de Lutte Contre Les Cancers 2021-2030* (rapp. tech.).

Institut National du Cancer. (2024). *Panorama Des Cancers En France* (rapp. tech.).

Ivanova, O., Martínez-Nicolás, I., & Meilán, J. J. G. (2023). Speech changes in old age : Methodological considerations for speech-based discrimination of healthy ageing and Alzheimer's disease. *International Journal of Language & Communication Disorders*, *59*(1), 13-37. https://doi.org/10.1111/1460-6984.12888

Janelsins, M. C., Kohli, S., Mohile, S., Usuki, K., Ahles, T., & Morrow, G. (2011). An Update On Cancer- And Chemotherapy-Related Cognitive Dysfunction : Current Status. *Semin Oncol*, *38*(3), 431-438. https://doi.org/doi:10.1053/j.seminoncol.2011.03.014

Janelsins, M. C., Heckler, C. E., Peppone, L. J., Ahles, T. A., Mohile, S. G., Mustian, K. M., Palesh, O., O'Mara, A. M., Minasian, L. M., Williams, A. M., Magnuson, A., Geer, J., Dakhil, S. R., Hopkins, J. O., & Morrow, G. R. (2018). Longitudinal Trajectory and Characterization of Cancer-Related Cognitive Impairment in a Nationwide Cohort Study. *Journal of Clinical Oncology*, *36*(32), 3231-3239. https://doi.org/10.1200/JCO.2018.78.6624

Janelsins, M. C., Kesler, S. R., Ahles, T. A., & Morrow, G. R. (2014). Prevalence, mechanisms, and management of cancer-related cognitive impairment. *International Review of Psychiatry*, *26*(1), 102-113. https://doi.org/10.3109/09540261.2013.864260

Janelsins, M. C., Lei, L., Netherby-Winslow, C., Kleckner, A. S., Kerns, S., Gilmore, N., Belcher, E., Thompson, B. D., Werner, Z. A., Hopkins, J. O., Long, J., Cole, S., & Culakova, E. (2022). Relationships between cytokines and cognitive function from pre- to post-chemotherapy in patients with breast cancer. *Journal of Neuroimmunology*, *362*, 577769. https://doi.org/10.1016/j.jneuroim.2021.577769

JASP Team. (2023). JASP.

Jeffreys, H. (1998, août). *Theory of Probability*. Oxford University PressOxford. https://doi.org/10.1093/oso/9780198503682.001.0001

Jim, H. S., Phillips, K. M., Chait, S., Faul, L. A., Popa, M. A., Lee, Y.-H., Hussin, M. G., Jacobsen, P. B., & Small, B. J. (2012). Meta-Analysis of Cognitive Functioning in Breast Cancer Survivors Previously Treated With Standard-Dose Chemotherapy. *Journal of Clinical Oncology*, *30*(29), 3578-3587. https://doi.org/10.1200/JCO.2011.39.5640

Joanette, Y., Ska, B., & Côté, H. (2004). *Protocole Montréal d'Evaluation de La Communication*. Isbergues, France : Ortho Edition.

John, G. H., & Langley, P. (2013). Estimating Continuous Distributions in Bayesian Classifiers. https://doi.org/10.48550/ARXIV.1302.4964

Joly, F., Lange, M., Rigal, O., Correia, H., Giffard, B., Beaumont, J., Clisant, S., & Wagner, L. (2012). French Version of the Functional Assessment of Cancer Therapy–Cognitive Function (FACT-

Cog) Version 3. *Support Care Cancer*, *20*, 3297-3305. https://doi.org/10.1007/s00520-012-1439-2

Jongman, S. R., Khoe, Y. H., & Hintz, F. (2021). Vocabulary Size Influences Spontaneous Speech in Native Language Users : Validating the Use of Automatic Speech Recognition in Individual Differences Research. *Language and Speech*, *64*(1), 35-51. https://doi.org/10.1177/0023830920911079

Jongman, S. R., Roelofs, A., & Lewis, A. G. (2020). Attention for Speaking : Prestimulus Motor-cortical Alpha Power Predicts Picture Naming Latencies. *Journal of Cognitive Neuroscience*, *32*(5), 747-761. https://doi.org/10.1162/jocn_a_01513

Jung, S. O., Kim, J. E. E., & Kim, H.-J. (2023). Assessing objective cognitive impairments in cancer survivors : Features and validity of measures for research and clinical applications. *Asia-Pacific Journal of Oncology Nursing*, *10*(11), 100309. https://doi.org/10.1016/j.apjon.2023.100309

Kachkovskaia, T., & Kocharov, D. (2024). Pausing strategies in dialogue speech : the interlocutor factor. *Speech Prosody 2024*, 612-616. https://doi.org/10.21437/SpeechProsody.2024-124

Kanaskie, M. L., & Loeb, S. J. (2015). The experience of cognitive change in women with breast cancer following chemotherapy. *Journal of Cancer Survivorship*, *9*(3), 375-387. https://doi.org/10.1007/s11764-014-0387-x

Kang, Y.-E., Yoon, J.-H., Park, N.-h., Ahn, Y.-C., Lee, E.-J., & Son, C.-G. (2023). Prevalence of cancer-related fatigue based on severity : a systematic review and meta-analysis. *Scientific Reports*, *13*(1), 12815. https://doi.org/10.1038/s41598-023-39046-0

Karschnia, P., Parsons, M. W., & Dietrich, J. (2019). Pharmacologic management of cognitive impairment induced by cancer therapy. *The Lancet Oncology*, *20*(2), e92-e102. https://doi.org/10.1016/S1470-2045(18)30938-0

Katzev, M., Tüscher, O., Hennig, J., Weiller, C., & Kaller, C. P. (2013). Revisiting the Functional Specialization of Left Inferior Frontal Gyrus in Phonological and Semantic Fluency : The Crucial Role of Task Demands and Individual Ability. *The Journal of Neuroscience*, *33*(18), 7837-7845. https://doi.org/10.1523/JNEUROSCI.3147-12.2013

Kehler, A. (2006, janvier). Discourse Coherence. In L. R. Horn & G. Ward (Éd.), *The Handbook of Pragmatics* (1re éd., p. 241-265). Wiley. https://doi.org/10.1002/9780470756959.ch11

Kemmerer, D. (2022, septembre). *Cognitive Neuroscience of Language* (2e éd.). Routledge. https://doi.org/10.4324/9781138318427

Kerr, A. J., Dodwell, D., McGale, P., Holt, F., Duane, F., Mannu, G., Darby, S. C., & Taylor, C. W. (2022). Adjuvant and neoadjuvant breast cancer treatments : A systematic review of their effects on mortality. *Cancer Treatment Reviews*, *105*, 102375. https://doi.org/10.1016/j.ctrv.2022.102375

Kesler, S. R. (2011). Prefrontal Cortex and Executive Function Impairments in Primary Breast Cancer. *Archives of Neurology*, *68*(11), 1447. https://doi.org/10.1001/archneurol.2011.245

Kesler, S. R., Adams, M., Packer, M., Rao, V., Henneghan, A. M., Blayney, D. W., & Palesh, O. (2017). Disrupted brain network functional dynamics and hyper-correlation of structural and

functional connectome topology in patients with breast cancer prior to treatment. *Brain and Behavior*, *7*(3), e00643. https://doi.org/10.1002/brb3.643

Kesler, S. R., Bennett, F. C., Mahaffey, M. L., & Spiegel, D. (2009). Regional Brain Activation during Verbal Declarative Memory in Metastatic Breast Cancer. *Clinical Cancer Research*, *15*(21), 6665-6673. https://doi.org/10.1158/1078-0432.CCR-09-1227

Kim, H.-J., Barsevick, A. M., Fang, C. Y., & Miaskowski, C. (2012). Common Biological Pathways Underlying the Psychoneurological Symptom Cluster in Cancer Patients. *Cancer Nursing*, *35*(6), E1-E20. https://doi.org/10.1097/NCC.0b013e318233a811

Knudsen, E. I. (2007). Fundamental Components of Attention. *Annual Review of Neuroscience*, *30*(1), 57-78. https://doi.org/10.1146/annurev.neuro.30.051606.094256

Kohler, C., Chang, M., Allemann-Su, Y.-Y., Vetter, M., Jung, M., Jung, M., Conley, Y., Paul, S., Kober, K. M., Cooper, B. A., Smoot, B., Levine, J. D., Miaskowski, C., & Katapodi, M. C. (2020). Changes in Attentional Function in Patients From Before Through 12 Months After Breast Cancer Surgery. *Journal of Pain and Symptom Management*, *59*(6), 1172-1185. https://doi.org/10.1016/j.jpainsymman.2020.01.001

Kruskal, W. H., & Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, *47*(260), 583-621. https://doi.org/10.1080/01621459.1952.10483441

Labov, W. (2001). Uncovering the Event Structure of Narrative. In D. Tannen & J. Alatis (Éd.), *Georgetown University Round Table on Languages and Linguistics 2001* (p. 63-83). Georgetown University Press.

Lange, M., Heutte, N., Noal, S., Rigal, O., Kurtz, J.-E., Lévy, C., Allouache, D., Rieux, C., Lefel, J., Clarisse, B., Leconte, A., Veyret, C., Barthélémy, P., Longato, N., Tron, L., Castel, H., Eustache, F., Giffard, B., & Joly, F. (2019). Cognitive Changes After Adjuvant Treatment in Older Adults with Early-Stage Breast Cancer. *The Oncologist*, *24*(1), 62-68. https://doi.org/10.1634/theoncologist.2017-0570

Lange, M., Heutte, N., Rigal, O., Noal, S., Kurtz, J.-E., Lévy, C., Allouache, D., Rieux, C., Lefel, J., Clarisse, B., Veyret, C., Barthélémy, P., Longato, N., Castel, H., Eustache, F., Giffard, B., & Joly, F. (2016). Decline in Cognitive Function in Older Adults With Early-Stage Breast Cancer After Adjuvant Treatment. *The Oncologist*, *21*(11), 1337-1348. https://doi.org/10.1634/theoncologist.2016-0014

Lanzi, A. M., Saylor, A. K., Fromm, D., Liu, H., MacWhinney, B., & Cohen, M. L. (2023). DementiaBank : Theoretical Rationale, Protocol, and Illustrative Analyses. *American Journal of Speech-Language Pathology*, *32*(2), 426-438. https://doi.org/10.1044/2022_AJSLP-22-00281

Leigh, R. J., & Zee, D. S. (2015, juin). *The Neurology of Eye Movements* (5e éd.). Oxford University Press. https://doi.org/10.1093/med/9780199969289.001.0001

Lepage, C., Smith, A. M., Moreau, J., Barlow-Krelina, E., Wallis, N., Collins, B., MacKenzie, J., & Scherling, C. (2014). A prospective study of grey matter and cognitive function alterations in

chemotherapy-treated breast cancer patients. *SpringerPlus*, *3*(1), 444. https://doi.org/10.1186/2193-1801-3-444

Levelt, W. J. M. (1983). Monitoring and Self-Repair in Speech. *Cognition*, *14*(1), 41-104. https://doi.org/10.1016/0010-0277(83)90026-4

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(01). https://doi.org/10.1017/S0140525X99001776

Lewis, J., Chapparo, C., Mackenzie, L., & Ranka, J. (2016). Work after breast cancer : Identification of cognitive difficulties using the Perceive, Recall, Plan, and Perform (PRPP) System of Task Analysis. *British Journal of Occupational Therapy*, *79*(5), 323-332. https://doi.org/10.1177/0308022616639983

Lewis, J., & Mackenzie, L. (2022). Cognitive changes after breast cancer : a scoping review to identify problems encountered by women when returning to work. *Disability and Rehabilitation*, *44*(18), 5310-5328. https://doi.org/10.1080/09638288.2021.1919216

Li, M., & Caeyenberghs, K. (2018). Longitudinal assessment of chemotherapy-induced changes in brain and cognitive functioning : A systematic review. *Neuroscience & Biobehavioral Reviews*, *92*, 304-317. https://doi.org/10.1016/j.neubiorev.2018.05.019

Lickley, R. (2015, mars). Fluency and Disfluency. In M. A. Redford (Éd.), *The Handbook of Speech Production* (1re éd., p. 445-474). Wiley. https://doi.org/10.1002/9781118584156.ch20

Lickley, R. (2017). Disfluency in typical and stuttered speech. In C. Bertini, C. Celata, G. Lenoci, C. Meluzzi & I. Ricci (Éd.), *Social and biological factors in speech variation* (p. 373-387, T. 3). Officinaventuno.

Lickley, R., & Bard, E. G. (1998). When Can Listeners Detect Disfluency in Spontaneous Speech? *Language and Speech*, *41*(2), 203-226. https://doi.org/10.1177/002383099804100204

Liu, H., MacWhinney, B., Fromm, D., & Lanzi, A. (2023). Automation of Language Sample Analysis. *Journal of Speech, Language, and Hearing Research*, *66*(7), 2421-2433. https://doi.org/10.1044/2023_JSLHR-22-00642

Lo Vercio, L., Amador, K., Bannister, J. J., Crites, S., Gutierrez, A., MacDonald, M. E., Moore, J., Mouches, P., Rajashekar, D., Schimert, S., Subbanna, N., Tuladhar, A., Wang, N., Wilms, M., Winder, A., & Forkert, N. D. (2020). Supervised Machine Learning Tools : A Tutorial for Clinicians. *Journal of Neural Engineering*, *17*(6), 062001. https://doi.org/10.1088/1741-2552/abbff2

Lowit, A., Thies, T., Steffen, J., Scheele, F., Roheger, M., Kalbe, E., & Barbe, M. (2022). Task-based profiles of language impairment and their relationship to cognitive dysfunction in Parkinson's disease. *PloS One*, *17*(10), e0276218. https://doi.org/10.1371/journal.pone.0276218

Lu, Y., Wiltshire, C. E., Watkins, K. E., Chiew, M., & Goldstein, L. (2022). Characteristics of articulatory gestures in stuttered speech : A case study using real-time magnetic resonance imaging. *Journal of Communication Disorders*, *97*, 106213. https://doi.org/10.1016/j.jcomdis.2022.106213

Lyon, D. E., Cohen, R., Chen, H., Kelly, D. L., McCain, N. L., Starkweather, A., Ahn, H., Sturgill, J., & Jackson-Cook, C. K. (2016). Relationship of systemic cytokine concentrations

to cognitive function over two years in women with early stage breast cancer. *Journal of Neuroimmunology*, *301*, 74-82. https://doi.org/10.1016/j.jneuroim.2016.11.002

Maass, S., Roorda, C., Berendsen, A., Verhaak, P., & De Bock, G. (2015). The prevalence of long-term symptoms of depression and anxiety after breast cancer treatment : A systematic review. *Maturitas*, *82*(1), 100-108. https://doi.org/10.1016/j.maturitas.2015.04.010

Mahanama, B., Jayawardana, Y., Rengarajan, S., Jayawardena, G., Chukoskie, L., Snider, J., & Jayarathna, S. (2022). Eye Movement and Pupil Measures : A Review. *Frontiers in Computer Science*, *3*, 733531. https://doi.org/10.3389/fcomp.2021.733531

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition : A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, *33*(3), 503-535. https://doi.org/10.1037/0278-7393.33.3.503

Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether One of Two Random Variables Is Stochastically Larger than the Other. *Annals of Mathematical Statistics*, *18*(1), 50-60. https://doi.org/10.1214/aoms/1177730491

Martínez-Nicolás, I., Llorente, T. E., Martínez-Sánchez, F., & Meilán, J. J. G. (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. *Frontiers in Psychology*, *12*, 620251. https://doi.org/10.3389/fpsyg.2021.620251

Mateo, S. (2020). Procédure pour conduire avec succès une revue de littérature selon la méthode PRISMA. *Kinésithérapie, la Revue*, *20*(226), 29-37. https://doi.org/10.1016/j.kine.2020.05.019

Max Planck Institute for Psycholinguistics. (2024). ELAN.

Mayo, S. J., Lustberg, M., M. Dhillon, H., Nakamura, Z. M., Allen, D. H., Von Ah, D., C. Janelsins, M., Chan, A., Olson, K., Tan, C. J., Toh, Y. L., Oh, J., Grech, L., Cheung, Y. T., Subbiah, I. M., Petranovic, D., D'Olimpio, J., Gobbo, M., Koeppen, S., … Peters, K. B. (2021). Cancer-related cognitive impairment in patients with non-central nervous system malignancies : an overview for oncology providers from the MASCC Neurological Complications Study Group. *Supportive Care in Cancer*, *29*(6), 2821-2840. https://doi.org/10.1007/s00520-020-05860-9

McDonald, B. C., Conroy, S. K., Ahles, T. A., West, J. D., & Saykin, A. J. (2012). Alterations in Brain Activation During Working Memory Processing Associated With Breast Cancer and Treatment : A Prospective Functional Magnetic Resonance Imaging Study. *Journal of Clinical Oncology*, *30*(20), 2500-2508. https://doi.org/10.1200/JCO.2011.38.5674

McDonald, W. M. (2017). Overview of Neurocognitive Disorders. *FOCUS*, *15*(1), 4-12. https://doi.org/10.1176/appi.focus.20160030

McMurray, B. (2022). The myth of categorical perception. *The Journal of the Acoustical Society of America*, *152*(6), 3819-3842. https://doi.org/10.1121/10.0016614

McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix : Capturing Linguistic Features of Cohesion. *Discourse Processes*, *47*(4), 292-330. https://doi.org/10.1080/01638530902959943

Mehlsen, M., Pedersen, A. D., Jensen, A. B., & Zachariae, R. (2009). No indications of cognitive side-effects in a prospective study of breast cancer patients receiving adjuvant chemotherapy. *Psycho-Oncology*, *18*(3), 248-257. https://doi.org/10.1002/pon.1398

Melara, R. D., Ahles, T. A., Prashad, N., Fernbach, M., Edelman, J. A., & Root, J. (2023). Longitudinal Effects of Breast Cancer Treatment on Neural Correlates of Attention. *Archives of Clinical Neuropsychology*, *38*(1), 12-24. https://doi.org/10.1093/arclin/acac051

Melis, M., Blommaert, J., Van Der Gucht, K., Smeets, A., McDonald, B. C., Sunaert, S., Smith, A., & Deprez, S. (2025). The impact of mindfulness on working memory-related brain activation in breast cancer survivors with cognitive complaints. *Journal of Cancer Survivorship*, *19*(2), 507-517. https://doi.org/10.1007/s11764-023-01484-0

Melis, M., Schroyen, G., Leenaerts, N., Smeets, A., Sunaert, S., Van Der Gucht, K., & Deprez, S. (2023). The impact of mindfulness on cancer-related cognitive impairment in breast cancer survivors with cognitive complaints. *Cancer*, *129*(7), 1105-1116. https://doi.org/10.1002/cncr.34640

Menning, S., De Ruiter, M. B., Veltman, D. J., Boogerd, W., Oldenburg, H. S. A., Reneman, L., & Schagen, S. B. (2018). Changes in brain white matter integrity after systemic treatment for breast cancer : a prospective longitudinal study. *Brain Imaging and Behavior*, *12*(2), 324-334. https://doi.org/10.1007/s11682-017-9695-x

Merceur, M., Reilly, K. T., Bonan, I., Holé, J., Hummel, E., Cogné, M., & Jacquin-Courtois, S. (2024). A systematic review of rehabilitation programs for cognitive impairment related to breast cancer : Different programs at different times ? *Annals of Physical and Rehabilitation Medicine*, *67*(5), 101832. https://doi.org/10.1016/j.rehab.2024.101832

Meskal, I., Gehring, K., Rutten, G.-J. M., & Sitskoorn, M. M. (2016). Cognitive functioning in meningioma patients : a systematic review. *Journal of Neuro-Oncology*, *128*(2), 195-205. https://doi.org/10.1007/s11060-016-2115-z

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. (1998). Viewing and naming objects : eye movements during noun phrase production. *Cognition*, *66*(2), B25-B33. https://doi.org/10.1016/S0010-0277(98)00009-2

Meyers, C. A., Albitar, M., & Estey, E. (2005). Cognitive impairment, fatigue, and cytokine levels in patients with acute myelogenous leukemia or myelodysplastic syndrome. *Cancer*, *104*(4), 788-793. https://doi.org/10.1002/cncr.21234

Miller, J. L., Grosjean, F., & Lomanto, C. (1984). Articulation Rate and Its Variability in Spontaneous Speech : A Reanalysis and Some Implications. *Phonetica*, *41*(4), 215-225. https://doi.org/10.1159/000261728

Mitchell, A. J., Meader, N., & Symonds, P. (2010). Diagnostic validity of the Hospital Anxiety and Depression Scale (HADS) in cancer and palliative settings : A meta-analysis. *Journal of Affective Disorders*, *126*(3), 335-348. https://doi.org/10.1016/j.jad.2010.01.067

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions : Four General Conclusions. *Current Directions in Psychological Science*, *21*(1), 8-14. https://doi.org/10.1177/0963721411429458

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses : The PRISMA Statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mulligan, N. W. (1998). The role of attention during encoding in implicit and explicit memory. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, *24*(1), 27-47. https://doi.org/10.1037/0278-7393.24.1.27

Munir, F., Burrows, J., Yarker, J., Kalawsky, K., & Bains, M. (2010). Women's perceptions of chemotherapy-induced cognitive side affects on work ability : a focus group study. *Journal of Clinical Nursing*, *19*(9-10), 1362-1370. https://doi.org/10.1111/j.1365-2702.2009.03006.x

Myers, J. S. (2012). Chemotherapy-Related Cognitive Impairment : The Breast Cancer Experience. *Oncology Nursing Forum*, *39*(1), E31-E40. https://doi.org/10.1188/12.ONF.E31-E40

Myers, J. S., Von Ah, D., He, J., Kim, J., Asakura, Y., & Williams, K. (2022). Psycholinguistic Assessment of Cognitive Function in Breast Cancer Survivors. *Oncology Nursing Forum*. https://doi.org/10.1188/22.ONF.565-570

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA : A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, *53*(4), 695-699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nemeth, D., Janacsek, K., Csifcsak, G., Szvoboda, G., Howard, J. H., & Howard, D. V. (2011). Interference between Sentence Processing and Probabilistic Implicit Sequence Learning (M. Perc, Éd.). *PLoS ONE*, *6*(3), e17577. https://doi.org/10.1371/journal.pone.0017577

Ng, T., Teo, S. M., Yeo, H. L., Shwe, M., Gan, Y. X., Cheung, Y. T., Foo, K. M., Cham, M. T., Lee, J. A., Tan, Y. P., Fan, G., Yong, W. S., Preetha, M., Loh, W.-J. K., Koo, S.-L., Jain, A., Lee, G. E., Wong, M., Dent, R., … Chan, A. (2016). Brain-derived neurotrophic factor genetic polymorphism (rs6265) is protective against chemotherapy-associated cognitive impairment in patients with early-stage breast cancer. *Neuro-Oncology*, *18*(2), 244-251. https://doi.org/10.1093/neuonc/nov162

Nguyen, C. M., Yamada, T. H., Beglinger, L. J., Cavanaugh, J. E., Denburg, N. L., & Schultz, S. K. (2013). Cognitive features 10 or more years after successful breast cancer survival : comparisons across types of cancer interventions. *Psycho-Oncology*, *22*(4), 862-868. https://doi.org/10.1002/pon.3086

Oliva, G., Giustiniani, A., Danesin, L., Burgio, F., Arcara, G., & Conte, P. (2024). Cognitive impairment following breast cancer treatments : an umbrella review. *The Oncologist*, *29*(7), e848-e863. https://doi.org/10.1093/oncolo/oyae090

Olson, B., & Marks, D. L. (2019). Pretreatment Cancer-Related Cognitive Impairment—Mechanisms and Outlook. *Cancers*, *11*(5), 687. https://doi.org/10.3390/cancers11050687

Organisation mondiale de la santé (Éd.). (2019). *Breast tumours* (5th ed). OMS.

Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice : A review on eye movements in decision making. *Acta Psychologica*, *144*(1), 190-206. https://doi.org/10.1016/j.actpsy.2013.06.003

O'Shaughnessy, D. (2024). Trends and developments in automatic speech recognition research. *Computer Speech & Language*, *83*, 101538. https://doi.org/10.1016/j.csl.2023.101538

Osterrieth, P. A. (1944). Le Test de Copie d'une Figure Complexe ; Contribution à l'étude de La Perception et de La Mémoire. Test of Copying a Complex Figure ; Contribution to the Study of Perception and Memory. *Arch Psychol*, *30*, 206-356.

Pakhomov, S. V. S., Smith, G. E., Chacon, D., Feliciano, Y., Graff-Radford, N., Caselli, R., & Knopman, D. S. (2010). Computerized Analysis of Speech and Language to Identify Psycholinguistic Correlates of Frontotemporal Lobar Degeneration. *Cognitive and Behavioral Neurology*, *23*(3), 165-177. https://doi.org/10.1097/WNN.0b013e3181c5dde3

Pallaud, B., Bertrand, R., Blache, P., Prévot, L., & Rauzy, S. (2019). Suspensive and Disfluent Self Interruptions in French Language Interactions. In L. Degand, G. Gilquin, L. Meurant & A.-C. Simon (Éd.), *Fluency and Disfluency across Languages and Language Varieties* (Corpora and Language in Use – Proceedings 4, p. 109-138).

Parjane, N., Cho, S., Ash, S., Cousins, K. A. Q., Shellikeri, S., Liberman, M., Shaw, L. M., Irwin, D. J., Grossman, M., & Nevler, N. (2021). Digital Speech Analysis in Progressive Supranuclear Palsy and Corticobasal Syndromes. *Journal of Alzheimer's disease : JAD*, *82*(1), 33-45. https://doi.org/10.3233/JAD-201132

Patchell, R. A. (2003). The management of brain metastases. *Cancer Treatment Reviews*, *29*(6), 533-540. https://doi.org/10.1016/S0305-7372(03)00105-1

Patti, F., Nicoletti, A., Messina, S., Bruno, E., Fermo, S. L., Quattrocchi, G., Chisari, C. G., Maimone, D., Cilia, S., & Zappia, M. (2015). Prevalence and incidence of cognitive impairment in multiple sclerosis : a population-based survey in Catania, Sicily. *Journal of Neurology*, *262*(4), 923-930. https://doi.org/10.1007/s00415-015-7661-3

Peacock, C. E., Singh, P., Hayes, T. R., Rehrig, G., & Henderson, J. M. (2023). Searching for meaning : Local scene semantics guide attention during natural visual search in scenes. *Quarterly Journal of Experimental Psychology*, *76*(3), 632-648. https://doi.org/10.1177/17470218221101334

Pélisson, D., Alahyane, N., Panouillères, M., & Tilikete, C. (2010). Sensorimotor adaptation of saccadic eye movements. *Neuroscience & Biobehavioral Reviews*, *34*(8), 1103-1120. https://doi.org/10.1016/j.neubiorev.2009.12.010

Pembroke, L., Sherman, K. A., Dhillon, H. M., Francis, H., Gurney, H., & Gillatt, D. (2024). What is the nature and impact of cognitive difficulties following hormonal treatments for prostate cancer ? : An interpretative phenomenological analysis. *Supportive Care in Cancer*, *32*(8), 534. https://doi.org/10.1007/s00520-024-08749-z

Petersen, R. (2016). Mild Cognitive Impairment. *Continuum*, *22*(2), 404-418. https://doi.org/10.1212/CON.0000000000000313

Pistono, A., & Hartsuiker, R. J. (2021). Eye-movements can help disentangle mechanisms underlying disfluency. *Language, Cognition and Neuroscience*, *36*(8), 1038-1055. https://doi.org/10.1080/23273798.2021.1905166

Pistono, A., & Hartsuiker, R. J. (2023). Can object identification difficulty be predicted based on disfluencies and eye-movements in connected speech ? (M. Corley, Éd.). *PLOS ONE*, *18*(3), e0281589. https://doi.org/10.1371/journal.pone.0281589

Player, L., Mackenzie, L., Willis, K., & Loh, S. Y. (2014). Women's experiences of cognitive changes or 'chemobrain' following treatment for breast cancer : A role for occupational therapy ? *Australian Occupational Therapy Journal*, *61*(4), 230-240. https://doi.org/10.1111/1440-1630.12113

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, *21*(6), 322-334. https://doi.org/10.1038/s41583-020-0304-4

Quesnel, C., Savard, J., & Ivers, H. (2009). Cognitive impairments associated with breast cancer treatments : results from a longitudinal study. *Breast Cancer Research and Treatment*, *116*(1), 113-123. https://doi.org/10.1007/s10549-008-0114-2

R Core Team. (2021). A Language and Environment for Statistical Computing.

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust Speech Recognition via Large-Scale Weak Supervision. https://doi.org/10.48550/ARXIV.2212.04356

Rapoeye, K., Hartsuiker, R. J., & Pistono, A. (2023). Semantic interference affects speech production by increasing disfluencies, not errors. *Royal Society Open Science*, *10*(6), 230006. https://doi.org/10.1098/rsos.230006

Reilly, J., Flurie, M., & Ungrady, M. B. (2020). Eyetracking during picture naming predicts future vocabulary dropout in progressive anomia. *Neuropsychological Rehabilitation*, 1-19. https://doi.org/10.1080/09602011.2020.1835676

Reilly, J., Peelle, J., Antonucci, S., & Grossman, M. (2011). Anomia as a Marker of Distinct Semantic Memory Impairments in Alzheimer's Disease and Semantic Dementia. *Neuropsychology*, *25*(4), 413-426. https://doi.org/10.1037/a0022738

Reilly, J., Shain, C., Borghesani, V., Kuhnke, P., Vigliocco, G., Peelle, J. E., Mahon, B. Z., Buxbaum, L. J., Majid, A., Brysbaert, M., Borghi, A. M., De Deyne, S., Dove, G., Papeo, L., Pexman, P. M., Poeppel, D., Lupyan, G., Boggio, P., Hickok, G., … Vinson, D. (2025). What we mean when we say semantic : Toward a multidisciplinary semantic glossary. *Psychonomic Bulletin & Review*, *32*(1), 243-280. https://doi.org/10.3758/s13423-024-02556-7

Ren, X., Boriero, D., Chaiswing, L., Bondada, S., St. Clair, D. K., & Butterfield, D. A. (2019). Plausible biochemical mechanisms of chemotherapy-induced cognitive impairment ("chemobrain"), a condition that significantly impairs the quality of life of many cancer survivors. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1865*(6), 1088-1097. https://doi.org/10.1016/j.bbadis.2019.02.007

Rey, A. (1941). L'examen Psychologique Dans Les Cas d'encéphalopathie Traumatique. The Psychological Examination in Cases of Traumatic Encepholopathy. *Arch Psychol*, *28*, 215-285.

Rey, A. (1958). Rey Auditory Verbal Learning Test. https://doi.org/10.1037/t27193-000

Richard, A. B., Foncelle, A., Hirsch, F., Jacquin-Courtois, S., Reilly, K. T., & Lelandais, M. (2024). Speech markers of Cancer-Related Cognitive Impairment : A pilot study. *Speech Prosody 2024*, 76-80. https://doi.org/10.21437/SpeechProsody.2024-16

Richard, A. B., Foncelle, A., Jacquin-Courtois, S., & Mouchès, P. (under review). Using Machine Learning to Identify Speech Markers of Reported Cognitive Impairment in Breast Cancer Survivors.

Richard, A. B., Hirsch, F., Jacquin-Courtois, S., & Reilly, K. T. (2023). Self-Interruptions in Breast Cancer Patients Who Complain of Anomia. In R. Skarnitzl & J. Volín (Éd.), *Proceedings of the 20th International Congress of Phonetic Sciences* (p. 4012-4016).

Richard, A. B., Lelandais, M., & Jacquin-Courtois, S. (under review). Assessing Cancer-Related Cognitive Impairment with Speech Analysis.

Richard, A. B., Lelandais, M., Reilly, K. T., & Jacquin-Courtois, S. (2024). Linguistic Markers of Subtle Cognitive Impairment in Connected Speech : A Systematic Review. *Journal of Speech, Language, and Hearing Research*, *67*(12), 4714-4733. https://doi.org/10.1044/2024_JSLHR-24-00274

Richard, A. B., Reilly, K. T., & Jacquin-Courtois, S. (2022). Que Révèlent Les Disfluences Sur Le Manque Du Mot Rapporté Par Les Patientes Ayant Un Cancer Du Sein. *Actes Des 34e Journées d'Etudes Sur La Parole*. https://doi.org/10.21437/JEP.2022-38

Rick, O., Gerhardt, A., & Schilling, G. (2024). Cancer-Related Cognitive Dysfunction : A Narrative Review for Clinical Practice. *Oncology Research and Treatment*, *47*(5), 218-223. https://doi.org/10.1159/000538277

Rick, O., Reuss-Borst, M., Dauelsberg, T., Hass, H. G., König, V., Caspari, R., Götz-Keil, G., Pfitzner, J., Kerschgens, C., Fliessbach, K., & Hoppe, C. (2018). NeuroCog FX study : A multicenter cohort study on cognitive dysfunction in patients with early breast cancer. *Psycho-Oncology*, *27*(8), 2016-2022. https://doi.org/10.1002/pon.4763

Röhr, S., Pabst, A., Riedel-Heller, S. G., Jessen, F., Turana, Y., Handajani, Y. S., Brayne, C., Matthews, F. E., Stephan, B. C. M., Lipton, R. B., Katz, M. J., Wang, C., Guerchet, M., Preux, P.-M., Mbelesso, P., Ritchie, K., Ancelin, M.-L., Carrière, I., Guaita, A., … Sachdev, P. S. (2020). Estimating prevalence of subjective cognitive decline in and across international cohort studies of aging : a COSMIC study. *Alzheimer's Research & Therapy*, *12*(1), 167. https://doi.org/10.1186/s13195-020-00734-y

Rosas, J. C., Aguado-Barrera, M. E., Azria, D., Briers, E., Elliott, R., Farcy-Jacquet, M.-P., Giraldo, A., Gutiérrez-Enríquez, S., Rancati, T., Rattay, T., Reyes, V., Rosenstein, B., De Ruysscher, D., Sperk, E., Stobart, H., Talbot, C., Vega, A., Taboada-Valladares, B., Veldeman, L., … the REQUITE Consortium. (2023). (Pre)treatment risk factors for late fatigue and fatigue trajectories following radiotherapy for breast cancer. *International Journal of Cancer*, *153*(9), 1579-1591. https://doi.org/10.1002/ijc.34640

Roy, M., Fowler, A. M., Ulaner, G. A., & Mahajan, A. (2023). Molecular Classification of Breast Cancer. *PET Clinics*, *18*(4), 441-458. https://doi.org/10.1016/j.cpet.2023.04.002

RStudio, BPC. (2020). RStudio : Integrated Development for R.

Ruff, R., Light, R., Parker, S., & Levin, H. (1997). The Psychological Construct of Word Fluency. *Brain and Language*, *57*(3), 394-405. https://doi.org/10.1006/brln.1997.1755

Ruiz-Casado, A., Álvarez-Bustos, A., De Pedro, C. G., Méndez-Otero, M., & Romero-Elías, M. (2021). Cancer-related Fatigue in Breast Cancer Survivors : A Review. *Clinical Breast Cancer*, *21*(1), 10-25. https://doi.org/10.1016/j.clbc.2020.07.011

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn Taking for Conversation. *Language*, *50*(4), 696-735. https://doi.org/10.1016/B978-0-12-623550-0.50008-2

Sammler, D., Grosbras, M.-H., Anwander, A., Bestelmeyer, P. E., & Belin, P. (2015). Dorsal and Ventral Pathways for Prosody. *Current Biology*, *25*(23), 3079-3085. https://doi.org/10.1016/j.cub.2015.10.009

Schilder, C. M., Seynaeve, C., Beex, L. V., Boogerd, W., Linn, S. C., Gundy, C. M., Huizenga, H. M., Nortier, J. W., Van De Velde, C. J., Van Dam, F. S., & Schagen, S. B. (2010). Effects of Tamoxifen and Exemestane on Cognitive Functioning of Postmenopausal Patients With Breast Cancer : Results From the Neuropsychological Side Study of the Tamoxifen and Exemestane Adjuvant Multinational Trial. *Journal of Clinical Oncology*, *28*(8), 1294-1300. https://doi.org/10.1200/JCO.2008.21.3553

Schmalz, X., Biurrun Manresa, J., & Zhang, L. (2023). What is a Bayes factor? *Psychological Methods*, *28*(3), 705-718. https://doi.org/10.1037/met0000421

Schmidt, C. S., Nitschke, K., Bormann, T., Römer, P., Kümmerer, D., Martin, M., Umarova, R. M., Leonhart, R., Egger, K., Dressing, A., Musso, M., Willmes, K., Weiller, C., & Kaller, C. P. (2019). Dissociating frontal and temporal correlates of phonological and semantic fluency in a large sample of left hemisphere stroke patients. *NeuroImage : Clinical*, *23*, 101840. https://doi.org/10.1016/j.nicl.2019.101840

Sekar, A., Panouillères, M., & Kaski, D. (2024). Detecting Abnormal Eye Movements in Patients with Neurodegenerative Diseases – Current Insights. *Eye and Brain*, *Volume 16*, 3-16. https://doi.org/10.2147/EB.S384769

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What Do Verbal Fluency Tasks Measure? Predictors of Verbal Fluency Performance in Older Adults. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00772

Shao, Z., Roelofs, A., & Meyer, A. S. (2012). Sources of individual differences in the speed of naming objects and actions : The contribution of executive control. *Quarterly Journal of Experimental Psychology*, *65*(10), 1927-1944. https://doi.org/10.1080/17470218.2012.670252

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3-4), 591-611. https://doi.org/10.1093/biomet/52.3-4.591

Shriberg, E. (2001). To 'errrr' is human : ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, *31*(1), 153-169. https://doi.org/10.1017/S0025100301001128

Silverman, D. H. S., Dy, C. J., Castellon, S. A., Lai, J., Pio, B. S., Abraham, L., Waddell, K., Petersen, L., Phelps, M. E., & Ganz, P. A. (2007). Altered frontocortical, cerebellar, and basal ganglia

activity in adjuvant-treated breast cancer survivors 5–10 years after chemotherapy. *Breast Cancer Research and Treatment*, *103*(3), 303-311. https://doi.org/10.1007/s10549-006-9380-z

Simon, A. C., & Christodoulides, G. (2016). Perception of prosodic boundaries by naïve listeners in French. *Speech Prosody 2016*, 1158-1162. https://doi.org/10.21437/SpeechProsody.2016-238

Skoogh, J., Steineck, G., Stierner, U., Cavallin-Ståhl, E., Wilderäng, U., Wallin, A., Gatz, M., Johansson, B., & on behalf of Swenoteca. (2012). Testicular-cancer survivors experience compromised language following chemotherapy : Findings in a Swedish population-based study 3–26 years after treatment. *Acta Oncologica*, *51*(2), 185-197. https://doi.org/10.3109/0284186X.2011.602113

Small, B. J., Rawson, K. S., Walsh, E., Jim, H. S. L., Hughes, T. F., Iser, L., Andrykowski, M. A., & Jacobsen, P. B. (2011). Catechol-O-methyltransferase genotype modulates cancer treatment-related cognitive deficits in breast cancer survivors. *Cancer*, *117*(7), 1369-1376. https://doi.org/10.1002/cncr.25685

Smith, A. (1973). Symbol Digit Modalities Test. https://doi.org/10.1037/t27513-000

Smith, K. M., Ash, S., Xie, S., & Grossman, M. (2018). Evaluation of Linguistic Markers of Word-Finding Difficulty and Cognition in Parkinson's Disease. *JOURNAL OF SPEECH LANGUAGE AND HEARING RESEARCH*, *61*(7), 1691-1699. https://doi.org/10.1044/2018_JSLHR-L-17-0304

Sousa, H., Almeida, S., Bessa, J., & Pereira, M. G. (2020). The Developmental Trajectory of Cancer-Related Cognitive Impairment in Breast Cancer Patients : A Systematic Review of Longitudinal Neuroimaging Studies. *Neuropsychology Review*, *30*(3), 287-309. https://doi.org/10.1007/s11065-020-09441-9

Spearman, C. (1904). The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, *15*(1), 72. https://doi.org/10.2307/1412159

Starkweather, A., Lyon, D., Elswick, R., Montpetit, A., Conley, Y., & McCain, N. (2013). A Conceptual Model of Psychoneurological Symptom Cluster Variation in Women with Breast Cancer : Bringing Nursing Research to Personalized Medicine. *Current Pharmacogenomics and Personalized Medicine*, *11*(3), 224-230. https://doi.org/10.2174/18756921113119990004

Stewart, A., Collins, B., Mackenzie, J., Tomiak, E., Verma, S., & Bielajew, C. (2008). The cognitive effects of adjuvant chemotherapy in early stage breast cancer : a prospective study. *Psycho-Oncology*, *17*(2), 122-130. https://doi.org/10.1002/pon.1210

Strober, L., Bruce, J., Arnett, P., Alschuler, K., Lebkuecher, A., Di Benedetto, M., Cozart, J., Thelen, J., Guty, E., & Roman, C. (2020). A new look at an old test : Normative data of the symbol digit modalities test –Oral version. *Multiple Sclerosis and Related Disorders*, *43*, 102154. https://doi.org/10.1016/j.msard.2020.102154

Sumowski, J. F., Benedict, R., Enzinger, C., Filippi, M., Geurts, J. J., Hamalainen, P., Hulst, H., Inglese, M., Leavitt, V. M., Rocca, M. A., Rosti-Otajarvi, E. M., & Rao, S. (2018). Cognition in multiple sclerosis : State of the field and priorities for the future. *Neurology*, *90*(6), 278-288. https://doi.org/10.1212/WNL.0000000000004977

Taillia, H., Léger, I., Moroni, C., Bonne, M., Berlin, C., Maillet, D., & Lange, M. (2015). Batterie Du Groupe de Réflexion Sur Les Évaluations Cognitives En Oncologie (GREC-ONCO). In *GREMOIRE 2 : Tests et Échelles Des Maladies Neurologiques Avec Symptomatologie Cognitive*. De Boeck-Solal.

Tangalos, E. G., Smith, G. E., Ivnik, R. J., Petersen, R. C., Kokmen, E., Kurland, L. T., Offord, K. P., & Parisi, J. E. (1996). The Mini-Mental State Examination in General Medical Practice : Clinical Utility and Acceptance. *Mayo Clinic Proceedings*, *71*(9), 829-837. https://doi.org/10.4065/71.9.829

Tate, M. C., Herbet, G., Moritz-Gasser, S., Tate, J. E., & Duffau, H. (2014). Probabilistic map of critical functional regions of the human cerebral cortex : Broca's area revisited. *Brain*, *137*(10), 2773-2782. https://doi.org/10.1093/brain/awu168

Team, A. (1999). Audacity.

Tenda, Y., Miyashita, M., Yamada, A., Shimizu, C., Nakayama, K., Honma, N., Taira, N., & Sawaki, M. (2022). Older patients' experience of living with cognitive impairment related to hormone therapy for breast cancer : A qualitative study. *European Journal of Oncology Nursing*, *57*, 102115. https://doi.org/10.1016/j.ejon.2022.102115

Themistocleous, C., Eckerstrom, M., & Kokkinakis, D. (2020). Voice Quality and Speech Fluency Distinguish Individuals with Mild Cognitive Impairment from Healthy Controls. *PLOS ONE*, *15*(7). https://doi.org/10.1371/journal.pone.0236009

Thompson, H. E., Almaghyuli, A., Noonan, K. A., Barak, O., Lambon Ralph, M. A., & Jefferies, E. (2018). The contribution of executive control to semantic cognition : Convergent evidence from semantic aphasia and executive dysfunction. *Journal of Neuropsychology*, *12*(2), 312-340. https://doi.org/10.1111/jnp.12142

Tilsen, S. (2019). Space and time in models of speech rhythm. *Annals of the New York Academy of Sciences*, *1453*(1), 47-66. https://doi.org/10.1111/nyas.14102

Tilsen, S., & Tiede, M. (2023). Parameters of unit-based measures of speech rate. *Speech Communication*, *150*, 73-97. https://doi.org/10.1016/j.specom.2023.05.006

Tobii AB. (2023). Tobii Pro Lab.

Tremblay, P., & Dick, A. S. (2016). Broca and Wernicke are dead, or moving past the classic model of language neurobiology. *Brain and Language*, *162*, 60-71. https://doi.org/10.1016/j.bandl.2016.08.004

Tremblay, P., & Gracco, V. L. (2010). On the selection of words and oral motor responses : Evidence of a response-independent fronto-parietal network. *Cortex*, *46*(1), 15-28. https://doi.org/10.1016/j.cortex.2009.03.003

Turk, A., & Shattuck-Hufnagel, S. (2013). What Is Speech Rhythm ? A Commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong. *Laboratory Phonology*, *4*(1). https://doi.org/10.1515/lp-2013-0005

Ullman, M. T. (2016). The Declarative/Procedural Model. In *Neurobiology of Language* (p. 953-968). Elsevier. https://doi.org/10.1016/B978-0-12-407794-2.00076-6

Umesh, P. (2012). Image Processing in Python. *CSI Communications*, 23.

Van Dam, F. S. A. M., Boogerd, W., Schagen, S. B., Muller, M. J., Droogleever Fortuyn, M. E., Wall, E. V., & Rodenhuis, S. (1998). Impairment of Cognitive Function in Women Receiving Adjuvant Treatment for High-Risk Breast Cancer : High-Dose Versus Standard-Dose Chemotherapy. *JNCI : Journal of the National Cancer Institute*, *90*(3), 210-218. https://doi.org/10.1093/jnci/90.3.210

Van Der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). Scikit-image : image processing in Python. *PeerJ*, *2*, e453. https://doi.org/10.7717/peerj.453

Van Dyk, K., Crespi, C. M., Bower, J. E., Castellon, S. A., Petersen, L., & Ganz, P. A. (2019). The cognitive effects of endocrine therapy in survivors of breast cancer : A prospective longitudinal study up to 6 years after treatment. *Cancer*, *125*(5), 681-689. https://doi.org/10.1002/cncr.31858

Van Dyk, K., Crespi, C. M., Petersen, L., & Ganz, P. A. (2020). Identifying Cancer-Related Cognitive Impairment Using the FACT-Cog Perceived Cognitive Impairment. *JNCI Cancer Spectrum*, *4*(1), pkz099. https://doi.org/10.1093/jncics/pkz099

Van Dyk, K., Ganz, P. A., Ercoli, L., Petersen, L., & Crespi, C. M. (2016). Measuring cognitive complaints in breast cancer survivors : psychometric properties of the patient's assessment of own functioning inventory. *Supportive Care in Cancer*, *24*(12), 4939-4949. https://doi.org/10.1007/s00520-016-3352-6

Van Kuppevelt, J. (1995). Main structure and side structure in discourse. *ling*, *33*(4), 809-833. https://doi.org/10.1515/ling.1995.33.4.809

Veal, B. M., Scott, S. B., Jim, H. S. L., & Small, B. J. (2023). Subjective cognition and memory lapses in the daily lives of breast cancer survivors : Examining associations with objective cognitive performance, fatigue, and depressed mood. *Psycho-Oncology*, *32*(8), 1298-1305. https://doi.org/10.1002/pon.6185

Verfaillie, S. C., Witteman, J., Slot, R. E., Pruis, I. J., Vermaat, L. E., Prins, N. D., Schiller, N. O., Van De Wiel, M., Scheltens, P., Van Berckel, B. N., Van Der Flier, W. M., & Sikkes, S. A. (2019). High amyloid burden is associated with fewer specific words during spontaneous speech in individuals with subjective cognitive decline. *Neuropsychologia*, *131*, 184-192. https://doi.org/10.1016/j.neuropsychologia.2019.05.006

Verstappen, C. C. P., Heimans, J. J., Hoekman, K., & Postma, T. J. (2003). Neurotoxic Complications of Chemotherapy in Patients with Cancer : Clinical Signs and Optimal Management. *Drugs*, *63*(15), 1549-1563. https://doi.org/10.2165/00003495-200363150-00003

Vigneau, M., Beaucousin, V., Hervé, P., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas : Phonology, semantics, and sentence processing. *NeuroImage*, *30*(4), 1414-1432. https://doi.org/10.1016/j.neuroimage.2005.11.002

Vincze, V., Szatloczki, G., Toth, L., Gosztolya, G., Pakaski, M., Hoffmann, I., & Kalman, J. (2021). Telltale silence : temporal speech parameters discriminate between prodromal dementia and mild Alzheimer's disease. *Clinical Linguistics & Phonetics*, *35*(8), 727-742. https://doi.org/

10.1080/02699206.2020.1827043
WOS :000573693900001.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van Der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Vázquez-Baeza, Y. (2020). SciPy 1.0 : fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261-272. https://doi.org/10.1038/s41592-019-0686-2

Voleti, R., Liss, J. M., & Berisha, V. (2020). A Review of Automated Speech and Language Features for Assessment of Cognitive and Thought Disorders. *IEEE Journal of Selected Topics in Signal Processing*, *14*(2), 282-298. https://doi.org/10.1109/JSTSP.2019.2952087

Von Ah, D., Crouch, A. D., Monahan, P. O., Stump, T. E., Unverzagt, F. W., Storey, S., Cohee, A. A., Cella, D., & Champion, V. L. (2022). Association of cognitive impairment and breast cancer survivorship on quality of life in younger breast cancer survivors. *Journal of Cancer Survivorship*, *16*(4), 812-822. https://doi.org/10.1007/s11764-021-01075-x

Von Ah, D., Habermann, B., Carpenter, J. S., & Schneider, B. L. (2013). Impact of perceived cognitive impairment in breast cancer survivors. *European Journal of Oncology Nursing*, *17*(2), 236-241. https://doi.org/10.1016/j.ejon.2012.06.002

Vos, L., Williams, M. W., Poritz, J. M. P., Ngan, E., Leon-Novelo, L., & Sherer, M. (2020). The Discrepancy Between Cognitive Complaints and Neuropsychological Test Findings in Persons With Traumatic Brain Injury. *Journal of Head Trauma Rehabilitation*, *35*(4), E382-E392. https://doi.org/10.1097/HTR.0000000000000557

Wagner, L., Sweet, J., Butt, Z., Lai, J., & Cella, D. (2009). Measuring Patient Self-Reported Cognitive Function : Development of the Functional Assessment of Cancer Therapy–Cognitive Function Instrument. *The Journal of Supportive Oncology*, *7*(6), W32-W39.

Waks, A. G., & Winer, E. P. (2019). Breast Cancer Treatment : A Review. *JAMA*, *321*(3), 288. https://doi.org/10.1001/jama.2018.19323

Wechsler, D. (1945). Wechsler Memory Scale.

Wechsler Memory Scale–Fourth Edition. (2012, novembre). https://doi.org/10.1037/t15175-000

Wefel, J., Vardy, J., Ahles, T., & Schagen, S. (2011). International Cognition and Cancer Task Force Recommendations to Harmonise Studies of Cognitive Function in Patients with Cancer. *Lancet Oncol*, *12*, 703-08. https://doi.org/10.1016/S1470-2045(10)70294-1

Whiteside, D. M., Kealey, T., Semla, M., Luu, H., Rice, L., Basso, M. R., & Roper, B. (2016). Verbal Fluency : Language or Executive Function Measure ? *Applied Neuropsychology : Adult*, *23*(1), 29-34. https://doi.org/10.1080/23279095.2015.1004574

Whittaker, A. L., George, R. P., & O'Malley, L. (2022). Prevalence of cognitive impairment following chemotherapy treatment for breast cancer : a systematic review and meta-analysis. *Scientific Reports*, *12*(1), 2135. https://doi.org/10.1038/s41598-022-05682-1

WHO Guidelines on Physical Activity and Sedentary Behaviour (1st ed). (2020). World Health Organization.

Williams, A. M., Shah, R., Shayne, M., Huston, A. J., Krebs, M., Murray, N., Thompson, B. D., Doyle, K., Korotkin, J., Van Wijngaarden, E., Hyland, S., Moynihan, J. A., Cory-Slechta, D. A., & Janelsins, M. C. (2018). Associations between inflammatory markers and cognitive function in breast cancer patients receiving chemotherapy. *Journal of Neuroimmunology*, *314*, 17-23. https://doi.org/10.1016/j.jneuroim.2017.10.005

Williams, K., Myers, J., Hu, J., Manson, A., & Maliski, S. (2021). Psycholinguistic Screening for Cognitive Decline in Cancer Survivors : A Feasibility Study. *Oncology Nursing Forum*, *48*(5), 474-480. https://doi.org/10.1188/21.ONF.474-480

Winocur, G., Vardy, J., Binns, M., Kerr, L., & Tannock, I. (2006). The effects of the anti-cancer drugs, methotrexate and 5-fluorouracil, on cognitive function in mice. *Pharmacology Biochemistry and Behavior*, *85*(1), 66-75. https://doi.org/10.1016/j.pbb.2006.07.010

Winocur, G., Berman, H., Nguyen, M., Binns, M. A., Henkelman, M., Van Eede, M., Piquette-Miller, M., Sekeres, M. J., Wojtowicz, J. M., Yu, J., Zhang, H., & Tannock, I. F. (2018). Neurobiological Mechanisms of Chemotherapy-induced Cognitive Impairment in a Transgenic Model of Breast Cancer. *Neuroscience*, *369*, 51-65. https://doi.org/10.1016/j.neuroscience.2017.10.048

Winter, B., & Grice, M. (2021). Independence and generalizability in linguistics. *Linguistics*, *59*(5), 1251-1277. https://doi.org/10.1515/ling-2019-0049

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining : Practical Machine Learning Tools and Techniques*. Elsevier. https://doi.org/10.1016/C2009-0-19715-5

Wright, H. H., Koutsoftas, A. D., Capilouto, G. J., & Fergadiotis, G. (2014). Global coherence in younger and older adults : Influence of cognitive processes and discourse type. *Aging, Neuropsychology, and Cognition*, *21*(2), 174-196. https://doi.org/10.1080/13825585.2013.794894

Yang, Y., & Ah, D. V. (2024). Cancer-Related Cognitive Impairment : Updates to Treatment, the Need for More Evidence, and Impact on Quality of Life—a Narrative Review. *Annals of Palliative Medicine*, *13*(5), 1265-1280. https://doi.org/10.21037/apm-24-70

Yao, C., Bernstein, L. J., & Rich, J. B. (2017). Executive functioning impairment in women treated with chemotherapy for breast cancer : a systematic review. *Breast Cancer Research and Treatment*, *166*(1), 15-28. https://doi.org/10.1007/s10549-017-4376-4

Yusuf, A. R. S., Heiling, H. M., Deal, A. M., Jensen, C. E., Mangieri, N. J., Nyrop, K. A., Lichtman, E. I., Rubinstein, S. M., Grant, S. J., Wood, W. A., Tuchman, S. A., & Nakamura, Z. M. (2022). Longitudinal Analysis of Patient-Reported Cognitive Function in Multiple Myeloma. *Clinical Lymphoma Myeloma and Leukemia*, *22*(12), 920-927. https://doi.org/10.1016/j.clml.2022.08.002

Zeng, Y., Dong, J., Huang, M., Zhang, J.-e., Zhang, X., Xie, M., & Wefel, J. S. (2020). Nonpharmacological interventions for cancer-related cognitive impairment in adult cancer patients : A network meta-analysis. *International Journal of Nursing Studies*, *104*, 103514. https://doi.org/10.1016/j.ijnurstu.2019.103514

Zhang, H. (2004). The Optimality of Naive Bayes. *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference*, *1*, 3.

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety And Depression Scale. *Acta Psychiatr Scand.*, *67*, 361-370.

Zimmerer, V. C., Hardy, C. J., Eastman, J., Dutta, S., Varnet, L., Bond, R. L., Russell, L., Rohrer, J. D., Warren, J. D., & Varley, R. A. (2020). Automated profiling of spontaneous speech in primary progressive aphasia and behavioral-variant frontotemporal dementia : An approach based on usage-frequency. *Cortex*, *133*, 103-119. https://doi.org/10.1016/j.cortex.2020.08.027

Annexes

Assessing Cancer-Related Cognitive Impairment for breast cancer survivors with speech analysis

Amélie B. Richard[a,b], Manon Lelandais[c], Sophie Jacquin-Courtois[b,d]

[a] Université de Montpellier Paul-Valéry, CNRS UMR5267; Praxiling, 34090 Montpellier, France [b] Université de Lyon, Université Lyon 1, INSERM U1028; CNRS UMR5292; Lyon Neuroscience Research Centre, F-69676 Lyon, France [c] Université Paris Cité, ALTAE URP3967, 75006 Paris, F-75013 France [d] Physical Medicine and Rehabilitation Department, Henry-Gabrielle Hospital, Hospices Civils de Lyon, 69610 Pierre-Bénite, France

Correspondence to Amélie B. Richard: amelie.richard@inserm.fr.

ABSTRACT

Background: This paper presents a first study on the assessment of Cancer-Related Cognitive Impairment (CRCI) with speech analysis. CRCI is a reported change in cognitive performance by cancer patients, which is too subtle to be detected by the current tests in use. Speech analysis might provide a solution to this issue. Speech analysis can be instrumental in the detection of subtle cognitive impairment, as speech contains fine-grained segmental and suprasegmental parameters that are sensitive to changes in cognition. Pauses have particularly been highlighted as potential behavioral markers of neurocognitive disorders. However, the absence of simple, detailed methods compromises the feasibility of speech analyses in clinical practice. **Objectives:** This study aims (i) to identify breast cancer survivors with CRCI using a new practical method centered on speech pauses, and (ii) to provide the detailed protocol that supports this study, which is intended to be applicable in clinical contexts. **Methods:** Thirty-three breast cancer survivors with a cognitive complaint, eleven breast cancer survivors without a cognitive complaint and thirteen controls were included in the study. Participants were instructed to tell a picture-based story. Their narratives were recorded, automatically transcribed with Whisper, and analyzed using the SPPAS and Praat software. Silent pauses, filled pauses, and sustained vowels were annotated, then processed in RStudio for statistical analysis. **Results:** Only silent pause duration was significantly longer for breast cancer survivors with a cognitive complaint than for controls. **Conclusions:** The results suggest that silent pause duration is a good marker for detecting CRCI. Automatizing the transcription and annotation of speech data improves the feasibility of speech analysis in clinical contexts, although a manual check is required.

Keywords: Cognitive impairment, cancer, speech analysis

Introduction

This paper describes a pilot study in which we use speech analysis to discriminate between patients with Cancer-Related Cognitive Impairment (henceforward CRCI) who have complained about cognitive change and healthy controls without any complaint. In the following subsections, we detail how speech analysis has been shown useful in the detection of cognitive impairment and why semi-automatic methods that are easy to implement are still needed in clinical contexts. We then describe the specific category of cognitive impairment our case population suffers from, along with the objectives and hypotheses of our pilot study.

Our pilot study is supported by a protocol for the analysis of prosody for patients with subtle cognitive impairment. The method it follows is intended to facilitate the diagnosis of subtle cognitive impairment. It is primarily addressed to certified speech and language therapists, students in speech and language pathology, and researchers who are interested in fluency-related phenomena.

Mild NeuroCognitive Disorders

Neurocognitive Disorders are considered mild (mild Neurocognitive Disorders, i.e., mNCDs) when the individual's autonomy is preserved American Psychiatric Association (2022). They cover a wide range of underlying physiopathologies, including neurodegenerative pathologies (e.g., early-stage Alzheimer's Disease, Lewy Body Dementia), vascular pathologies (e.g., stroke), traumatic pathologies (e.g., Traumatic Brain Injury), as well as infectious or oncological contexts (e.g., Covid-19 sequelae, HIV, Cancer-Related Cognitive Impairment). Mild NCDs also encompass cognitive-behavioral syndromes such as Mild Cognitive Impairment, for which the diagnosis of an etiology has not been established.

The prevalence of mNCDs is difficult to assess due to the diversity of possible underlying physiophathologies, and therefore might be underestimated. For traumatic brain injury, 10% to 15% of patients may develop a persistent post-concussion syndrome including cognitive impairment (Groupe de Travail Traumatisme Crânio-Cérébral Léger en Ile-de-France, 2021). In a study including seventy-four post minor stroke patients whose MMSE scores are in the range of norms (> 27/ 30), 70% of the patients (i.e., 52/74 patients) reported difficulty with social functioning, which was correlated with a working memory deficit (Hommel et al., 2009). This finding signals a high risk of developing cognitive impairment following a minor stroke. The authors also suggest that the cognitive screening tests in use might be not sensitive enough to detect this subtle form of cognitive impairment.

During clinical consultations, patients or their relatives complain of progressive cognitive difficulties, described as subtle but invalidating. However, their complaints are rarely correlated with scores on cognitive screening tests such as the MMSE or MoCA. As a result, mNCDs often remain underdiagnosed. The scientific public authority for health in France (i.e., Haute Autorité de Santé) stresses that any concerning cognitive complaint, particularly in the elderly, should be thoroughly investigated (Haute Autorité de Santé, 2018). The same level of caution has to apply to possibly younger populations in other contexts (i.e., minor stroke, mild Traumatic Brain Injury, Cancer-Related Cognitive Impairment, etc.) with high stakes in medical, economical, and public health aspects.

Cancer-related Cognitive Impairment

Among frequent mNCDs, Cancer-Related Cognitive Impairment (henceforward CRCI) is a form of subtle cognitive impairment induced by cancer (except central nervous system cancer) and/or its treatments, which may affect up to 75% of oncology patients (Janelsins et al., 2011). CRCI alters patients' quality of life. For some, the symptomatology is long-lasting, lasting more than two years after oncology treatments have been stopped. The persistence of the disorder limits participation in social activities and hinders the return to work for a potentially young population. The subtle nature of CRCI makes it an invisible disability that is often under-diagnosed, even though appropriate assessment and treatment options are available.

CRCI is characterized by a cognitive complaint affecting several aspects of cognition (memory, executive functions, attention, processing speed, language) that can be objectified by standardized questionnaires (e.g., FACT-Cog, Wagner et al. [2009]). One of the most frequent complaints is forgetting words. Despite the intensity of the complaint, the current speech and neuropsychological tests in use struggle to fully display the difficulties reported by patients, as they are in majority designed to assess more severe disorders. Scores obtained on verbal fluency and naming tests in particular are often within the norm. It is therefore necessary to develop new approaches and/or new tools to detect such subtle cognitive disorders.

Linguistic markers of cognitive impairment

The literature in aphasiology, whether vascular or neurodegenerative, has shown the benefit of analyzing patients' speech and voice to reveal cognitive disorders (see for instance Martínez-Nicolás et al. [2021]). Speech analysis has been found useful to assess the cognition of patients because of its non-invasive nature, its sensitivity to subtle cognitive changes and its high ecological validity (Lanzi et al., 2023). Speech can be studied through segmental and suprasegmental analyses, depending on whether the focus is on singular segments such as phonemes, or on attributes that can extend over phonemes such as fundamental frequency (Hz) and tones, intensity (dB) or rhythm, e.g., with duration (ms). Prosody, i.e., the study of speech elements including suprasegmental phenomena, is particularly

relevant to investigate mild cognitive impairment, as it contains fine-grained elements that can signal changes in a speaker's cognition (Ivanova et al., 2023). These prosodic elements include parameters related to voice quality, such as variations in intonational amplitude (Themistocleous et al., 2020), as well as broader speech units such as pauses (Vincze et al., 2021). For instance, Parjane et al. (2021) compared descriptive speech productions between patients with the nonfluent-agrammatic variant of Primary Progressive Aphasia, patients with Corticobasal Degeneration and healthy controls. Both patient groups showed longer, more frequent pauses and reduced fundamental frequency ranges (i.e., F0) compared to controls.

However, the variability of the methodologies in use (e.g., elicited speech type, linguistic variables, tools for analysis) makes it difficult to compare results across studies even for studies with similar objectives (Richard, Lelandais, et al., 2024). This is even more the case for cancer patients whose cognition has not been investigated through speech yet. As part of a harmonization effort, Angelopoulou et al. (2024) proposed a speech analysis method that was used to compare two types of elicited descriptive speech (i.e., a picture description and an autobiographical description) in a healthy population. The interest of their study mainly lies in the description of the measurement of nineteen speech variables, although some concepts such as "words" or "pauses" are not defined (e.g., the authors do not differentiate pauses due to physiological constraints from pauses having a demarcative or hesitation function).

However, no mention of semi-automatic tool is made in the methodology, which limits its feasibility in clinical practice. The authors for instance report an orthographic transcription stage during which they applied coding conventions, but they did not specify whether they used an automatic transcription tool or script. The transcription and annotation stages are time-consuming, and are therefore difficult to apply and replicate in clinical practice. A detailed method that is easy to implement in the clinic is still needed.

Semi-automatic speech analysis

In this section, we argue for the use of automatic, off-line and open source software.

Use of automatic tools

Speech analysis is a long and tedious task. After recording a patient's oral production, it is often necessary to transcribe it, either orthographically or phonetically. Manual transcription is especially time-consuming. Transcribing one minute of speech may take up to twelve minutes (Bryant et al., 2017). This step is also particularly costly in terms of concentration, since it requires several listening sessions of the recording in order to avoid omitting words or disfluencies produced by the speaker. For instance, determining the exact number of repetitions of a lexical or phonemic item will typically

require several takes. Meanwhile, the process of multiple listenings entails biases linked to auditory fatigability, subjectivity and habituation, and thus increases the risk of transcription errors (e.g., addition of words by priming effect).

The use of Artificial Intelligence tools (i.e., computer tools that automate a cognitive task) helps to reduce time and such biases by, for instance, providing a "neutral" first transcription. Artificial Intelligence is a broad field in computer science that spans automated cognitive tasks using simple algorithms, understandable machine learning classification, prediction and clustering models, as well as complex deep learning systems such as generative AI. These tools are based on technologies of varying complexity, from simple automated tasks (e.g., automatic detection of silent pauses) to neural network models (e.g., automatic transcription), and offer considerable time-saving advantages for speech analysis (see O'Shaughnessy [2024] for a review of Automatic Speech Recognition systems). However, AI tools can vary in performance depending on the audio quality of the recording. Manual verification is therefore essential to correct any errors made by these tools.

Use of offline tools

The recorded voice of patients and study participants represents identifying data that must be protected at a high level of security and confidentiality (see the Commission Nationale de l'Informatique et des Libertés [2016]). As a result, it is crucial to choose transparent data management tools. In addition, available resources on the internet sometimes have obscure privacy policies, which increases the risk of data leaks. The offline tools described in this study meet the security criteria required for speech analysis in the sense that data is processed from its storage place.

Use of open source tools

Open source software is free software whose programming codes are accessible to any user. This type of collaborative tool not only guarantees total transparency in data management, but also reinforces the scientific reproducibility and reliability of results.

Objective and hypotheses

Pilot studies have shown the benefit of analyzing disfluencies (Richard et al., 2022) particularly pauses, in the speech production of breast cancer survivors with cognitive complaints (Richard et al., 2023). A qualitative analysis of hesitation pauses produced during a picture-based narration task showed potential to differentiate breast cancer survivors with a cognitive complaint from breast cancer survivors without a cognitive complaint and control participants (Richard, Foncelle, et al., 2024).

The aim of the present study is to test whether three different types of pauses can distinguish cancer survivors with a cognitive complaint from those without and healthy controls. We apply a semi-automatized speech analysis method to this population, and provide the full protocol for both research and clinical uses.

Based on these previous studies, we hypothesize that silent pauses, filled pauses and sustained vowels will be longer (in ms) for breast cancer patients with CRCI (i.e., the presence of a significant complaint according to the FACT-Cog questionnaire) than for breast cancer patients without CRCI and healthy controls.

Methods

Population

Fifty-seven breast cancer survivors and fourteen age- gender- and education-matched healthy controls volunteered for the current study.

Inclusion criteria

Inclusion criteria for all participants were (i) to be a woman aged between 30 and 70, (ii) to be a native French speaker or be fluent in French (iii) to have normal or corrected sight and hearing. In addition, cancer survivors had to (iv) have completed curative treatments (i.e., surgery, chemotherapy, radiotherapy) more than four weeks and less than one year ago, excluding hormone therapy.

Exclusion criteria

Non-inclusion criteria for all participants were (i) a history of oral language disorders, (ii) a neurological or psychiatric history, (iii) the use of medications that are likely to cause cognitive changes. Additionally, participants who had breast cancer had to (iv) not present metastases. Ten breast cancer survivors were not eligible due to brain or tumoral comorbidities, and one healthy control was not eligible for dyslexia.

Groups

The FACT-Cog questionnaire is a standardized questionnaire for assessing cognitive complaints in oncology patients (Wagner et al., 2009), and for which a threshold of less than or equal to 55/72 on the Perceived Cognitive Impairment (PCI) subscale indicates the presence of significant cognitive

complaints (Van Dyk et al., 2020). Participants completed the French version of the FACT-Cog v3 questionnaire (Costa et al., 2018; Joly et al., 2012). We used this score to split our population into three groups: breast cancer survivors with a cognitive complaint (BC_{yes}) according to the PCI score of the FACT-Cog questionnaire, breast cancer survivors with no cognitive complaint (BC_{no}) according to the PCI score of the FACT-Cog questionnaire, and healthy controls (HC).

In total, thirty-three BC_{yes} participants (nineteen from the ACTICOG study, six from REMCOG studies and eight from the DisCCo study), eleven BC_{no} participants (ten from the ACTICOG study and one from the DisCCo study), and thirteen age- and education-matched HC participants (DisCCo study) were eligible.

Narrative task

The given task to the participants was a narrative speech task based on a sequence of five images, in which participants had to imagine the sixth (Bézy et al., 2016). The images tell the story of a family arriving at the beach. The father is windsurfing while his son and wife are on the shore. Suddenly the wind picks up, and the father falls into the water. The last image shows the alarmed mother and son, with the mother calling for help on the phone. Participants were instructed to tell the story based on the images they see, with no time limit.

Recording material and conditions

The narrative productions of participants in the ACTICOG and REMCOG studies were recorded in a quiet room at the Henry Gabrielle rehabilitation hospital (Saint-Genis-Laval, France) using a Rode Lavalier Go® omnidirectional lavalier microphone connected to a laptop.

The speech samples of DisCCo participants were recorded in an anechoic booth at the Lyon Neuroscience Research Centre (Bron, France) using a Shure® omnidirectional lapel microphone lined to a M-Audio®portable audio interface.

Linguistic parameters

Based on our previous pilot studies, we analyzed silent pauses, filled pauses and sustained vowels (also referred to as lengthenings in the literature). Silent pauses (SP) were segments characterized by an absence of speech, cough, or yawn during 200ms or more, from the first to the last speech phonemes. Silent pauses might include lip or tongue noise or breathing. Filled pauses (FP) were non-semantic voiced segments such as "uh" and "uhm" produced for 200ms or more with a flat F0 contour. Sustained vowels (SV) were sustained vowels produced for 200ms with a flat F0 contour.

We extracted mean durations for silent pauses, filled pauses and sustained vowels for each participant.

Prosodic analysis

The following section describes the different steps in the prosodic analysis, starting from the acoustic signal. Depending on the reader's needs, sections can be used independently of each other. A full tutorial of our method is available (see Supplementary material section).

Acoustic signal pre-processing

Before starting our prosodic analyses, we segmented the audio recording to only retain the narrative production. This step was carried out using Audacity free software (Team, 1999).

Transcription

We used automatic speech transcription with the Whisper large-v3 model (Radford et al., 2022). Whisper is an open source code dedicated to automatic speech recognition. As part of its functionalities, Whisper provides an orthographic transcription of speech from the acoustic signal. A detailed description of how to install and use Whisper is available (see the Supplementary material section).

We checked each text for transcription errors, and manually added silent pauses (#) and filled pauses (uh). It is important to note that Whisper removes all disfluencies (e.g., pauses, repetition, false starts, etc.) by default.

Finally, we cleaned up the data, for instance removing punctuation marks or replacing numbers with their equivalent in letters. The Python code developed for this study is available (see the Supplementary material section).

Pause annotation

We annotated pauses with the SPPAS software(Bigi, 2015). SPPAS is an open source software package that has specifically been developed for speech analysis. It provides a wide range of automatic analyses, such as silent pause detection, grapheme-phoneme conversion and syllabification of transcriptions aligned with the acoustic signal (i.e., sound). Its graphical interface and interoperability (i.e., its ability to work with other software in terms of compatibility formats) make it equally suitable to research and clinical applications. To refine our analyses, we also used the free Praat software (Boersma & Weenink, 2024). This software has been designed for acoustic signal

analysis and visualization, as well as the extraction of numerous prosodic variables (e.g., fundamental frequency).

The following paragraphs describe the steps we took for the data annotation phase. We first uploaded the (.wav format) recordings and (.txt format) transcriptions into SPPAS. From the "Annotation" tab, we searched for Inter-Pausal Units (i.e., IPUs), then ran the automatic transcription fill-in. The first function, "Search for IPUs", automatically detects silent pauses, while the second, "Fill-In IPUs", fills in empty spaces with the transcription of speech segments. These functions can be accessed via the "STANDALONE" button. We have chosen the `.TextGrid` output format, which is processed by Praat.

We then performed a verification step by opening the `.TextGrid` files in Praat. We readjusted the boundaries of silent pauses when they did not match the acoustic signal. This represented no readjustment in good quality recordings, and approximately 1 to 5 readjustments in poor quality recordings of 82 seconds.

In the SPPAS "Annotation" tab, we ran the "Text Normalization", "Phonetization" and "Alignment" functions simultaneously. This step segments the audio signal into phonemes from the text, and temporally aligns the phonemes with the acoustic signal. Vowel lengthenings and filled pauses can this way easily be identified based on their duration of 200 ms or more.

We ran another check in Praat, opening the `.TextGrid` files, now named -palign, to ensure that the phonemes were correctly aligned with the acoustic signal.

In the SPPAS "Analyze" tab, we opened the -palign `.TextGrid` files, and selected the annotation tier "TokenAlign". We applied a filter (i.e., single filter) to create a new tier named "pauses". This tier contains vowel phonemes and pauses (i.e., U /, A/, O/, @, i, E, e, j, H, a , O , y, 2, 9, u, fp, # for French according to SPPAS conventions) of 200 ms or more. In Praat, we only retained vowel lengthening annotations that had a flat intonation curve.

Finally, we harmonized all annotations by replacing the existing annotations in the "pauses" tier with the following labels: SP for Silent Pauses, FP for Filled Pauses and SV for Sustained Vowels. To this effect, we ran the text replacement script developed by DiCanio (2011) in Praat, available at the following link: https://www.acsu.buffalo.edu/~cdicanio/scripts.html

Inter-annotator agreement

Calculating inter-annotator agreement consists in giving a share of the data (usually about 20%) along with its annotation protocol (i.e., a description of the elements to annotate) to one or more individuals who are naïve to the dataset and the goals of the study to ensure the validity and reliability of the annotation process. It is measured with Cohen's κ, which score ranges from 0 to 1. A score equal to or above .81 confirms the reliability of the annotation.

We gave 20% of our dataset to two naïve annotators to verify that our definition of silent pauses was replicable. The inter-annotator agreement reached 98.71% of agreement, qualifying as a near prefect ($\kappa = .97$).

Extraction of prosodic data

We extracted mean pause duration as well as standard deviation for all three pause types in each speech sample with SPPAS. In the "Analyze" tab, we selected all "pauses" tiers and opened the "Statistics" option to access data spreadsheets.

Psychological and cognitive assessment

In addition to the FACT-Cog questionnaire as a subjective assessment of CRCI, participants underwent cognitive screening and psychological evaluations to control for confounding variables that might influence pause production.

Participants took the Montreal Cognitive Assessment test (Nasreddine et al., 2005). This rapid, self-administered test is designed to detect fine-grained cognitive impairment by assessing several cognitive domains (e.g., memory, visuo-spatial abilities, language). A score below 26/30 indicates the presence of cognitive decline.

Participants completed the Hospital Anxiety Depression (HAD) scale (Zigmond & Snaith, 1983). This questionnaire comprises two subscales designed to assess anxiety (HAD-A) and depression (HAD-D). A score above 11/21 on either subscale indicates the presence of anxiety or depression.

Finally, participants completed the Functional Assessment of Chronic Illness Therapy — Fatigue (FACIT-F) (Bonomi et al., 1996; Cella et al., 1993). This questionnaire was specially designed to assess overall cancer-related fatigue. The higher the score, the greater the fatigue. There is no standardized threshold for this questionnaire.

Statistical analyses

The statistical analyses in this study were run using R language (R Core Team, 2021) in RStudio (RStudio, BPC, 2020). This software is free, off-line and open source. With multiple online tutorials and its active community, R is easy to use for parametric and non-parametric frequentist and Bayesian statistical models, as well as data visualization. Our data is stored in a spreadsheet in .csv format. Each column corresponds to a variable, and each row to a participant. Figure 1 shows the ordering of the data after its importation in RStudio.

Figure 1. Data ordering from four participants in RStudio.

ID <chr>	Age <int>	Group <chr>	FACT.Cog_PCI <int>	FACIT <int>	SP_mean <dbl>
AL01	41	BC	60	46	0.7843750
AL02	39	BC	72	39	0.6536719
AL03	68	BC_cog	31	19	1.2844118
AL04	41	BC_cog	38	38	1.2593611

Before starting the analyses, we checked the assumption of normality of our data by running a Shapiro-Wilk test (Shapiro & Wilk, 1965). The result gave a significant p-value, meaning that our dataset was not normally distributed. We consequently chose non-parametric tests for our statistical analyses. We ran Kruskal-Wallis U tests (Kruskal & Wallis, 1952) to compare the variables of the BC_{yes} participants with those of the BC_{no} and the HC participants, and Dunn's post-hoc analyses for multiple comparisons (Dunn, 1964) to see which group differed from the others. Finally, we performed a Spearman correlation (Spearman, 1904) to test the possibility of an effect of demographic and psychological variables and (i.e., age, anxiety, depression, fatigue) on pauses.

Results

Demographic and psychological data

Fifty-seven participants took part in this study (BC_{yes} N = 33, BC_{no} N = 11, HC N = 13). There was no missing data except for one participant who did not fill in the questionnaires on fatigue (FACIT-F), anxiety (HAD-A), and depression (HAD-D). Anxiety and depression scores were below alert thresholds for all groups.

Table 1 summarizes the results of the Kruskal-Wallis test comparing the demographic and psychological variables between groups.

Age

A Kruskal-Wallis test indicated no significant age difference across the three groups (X^2 (2, N = 57) = 4.85, p = .088).

Screening of cognition

A Kruskal-Wallis test indicated no significant difference in the MoCA cognitive screening across the three groups (X^2 (2, N = 57) = 2.668, p = .263). This result means that no participant showed cognitive decline according to the screening test.

Fatigue

A Kruskal-Wallis test indicated a significant difference in reported fatigue as measured with the FACIT-F questionnaire across the three groups (X^2 (2, N = 56) = 26.422, p < .001). Post-hoc comparisons using Dunn's test showed that the median fatigue score for BC_{yes} was significantly higher than for BC_{no} and HC, p < .001. However, there was no significant difference was found in the median fatigue scores for BC_{no} and HC, p = .174. One participant (BC_{yes}) did not fill in the questionnaire.

Anxiety

A Kruskal-Wallis test showed a significant difference in reported anxiety as measured with the anxiety subscale of the HAD scale across the three groups (X^2 (2, N = 56) = 7.710, p = .021). Post-hoc comparisons using Dunn's test showed significantly higher median anxiety scores for BC_{yes} than for HC, p = .029. However, no significant difference was found in median anxiety scores between BC_{yes} and BC_{no}, p = .195, and between BC_{no} and HC p = .557. One participant (BC_{yes}) did not fill in the questionnaire. Anxiety scores remained below alert thresholds for all groups.

Depression

A Kruskal-Wallis test indicated a significant difference in reported depression as measured with the HAD depression subscale across the three groups (X^2 (2, N = 56) = 20.841, p < .001). Post-hoc comparisons using Dunn's test showed a higher median depression score for BC_{yes} than for HC, p < .001. However, there was no significant difference in median depression scores between BC_{yes} and BC_{no}, p = .063, and between BC_{no} and HC p = .105. One participant (BC_{yes}) did not fill in the questionnaire. Depression scores remained below alert thresholds for all groups.

Table 1. Characteristics of the population. Median, SD.

Variables	BC_{yes}	BC_{no}	HC	*p*-value
N =	33	11	13	—
Age	50 (8.7)	46 (8.4)	42 (10.2)	.088
FACT-Cog PCI (/72)	35 (12.2) ∗†	62 (5.2) †	68 (4.6) ∗	< .001
MoCA (/30)	27 (2.2)	29 (1.8)	28 (2.3)	.263
FACIT-F (/52)	28 (13.2) ∗†	44 (13.8) †	47 (6.9) ∗	< .001
HAD-A (/21)	10 (4.4) ∗	7 (2.8)	5 (4.0) ∗	.021
HAD-D (/21)	7 (3.7) ∗	4 (2.9)	2 (1.6) ∗	< .001

Note: BC_{yes} = breast cancer survivors with a significant cognitive complaint, BC_{no} = breast cancer survivors without any significant cognitive complaint, HC = healthy controls. ∗ = significant difference between BC_{yes} and HC. † = significant difference between BC_{yes} and BC_{no}.

Corpus characteristics

Our corpus is made of 57 speech samples, with a 79-minute total duration (mean per sample = 82 s ± 32s). A total of 2823 pauses were annotated, with 634 silent pauses, 1229 filled pauses, and 960 sustained vowels.

Pause analysis results

Given our small sample size and our objective to stay close to participants' production, we decided to keep outliers in.

Silent pauses

Figure 2 shows the distribution of mean silent pause duration overlapped across the three groups. While BC_{yes} and HC have distinct curves, BC_{no}'s distribution is in-between these two other groups with a mean closest to that of BC_{yes}. The three curves show a large first peak, followed by a smaller second peak standing for longer pause duration. This shows that in each group, a majority of participants produced a similar mean silent pause duration, while some participants produced a longer mean duration than the majority of the group, around .400 ms after the group mean. The range of each group is thus composed of two distribution peaks.

Figure 2. Overlapping distributions of mean silent pause duration (ms) for all three groups. Dotted lines show the mean duration of silent pauses for each group.

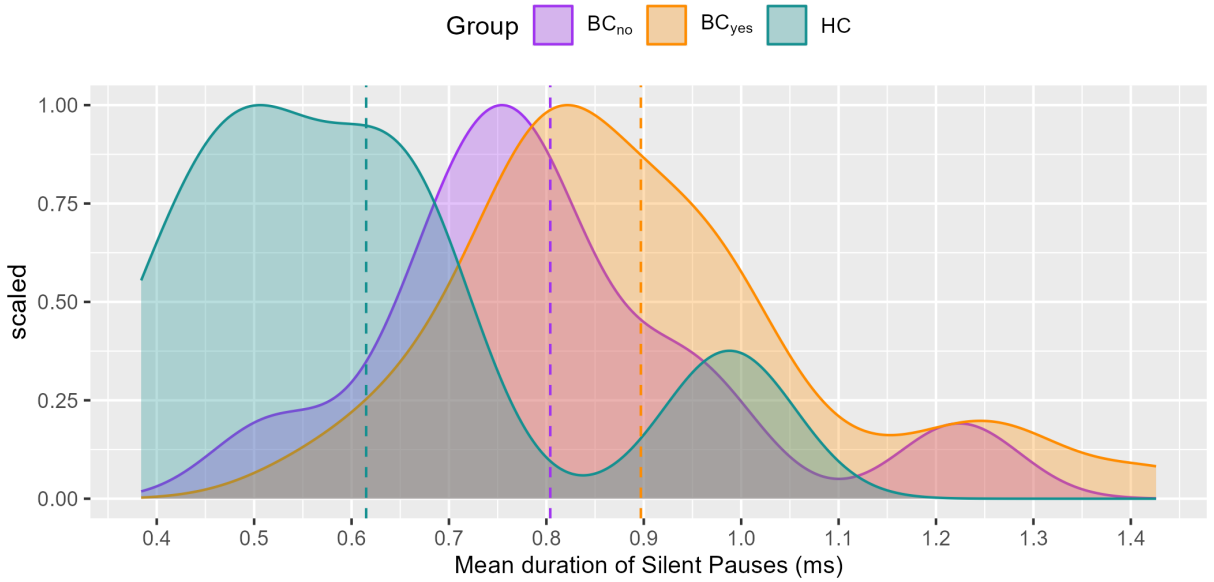
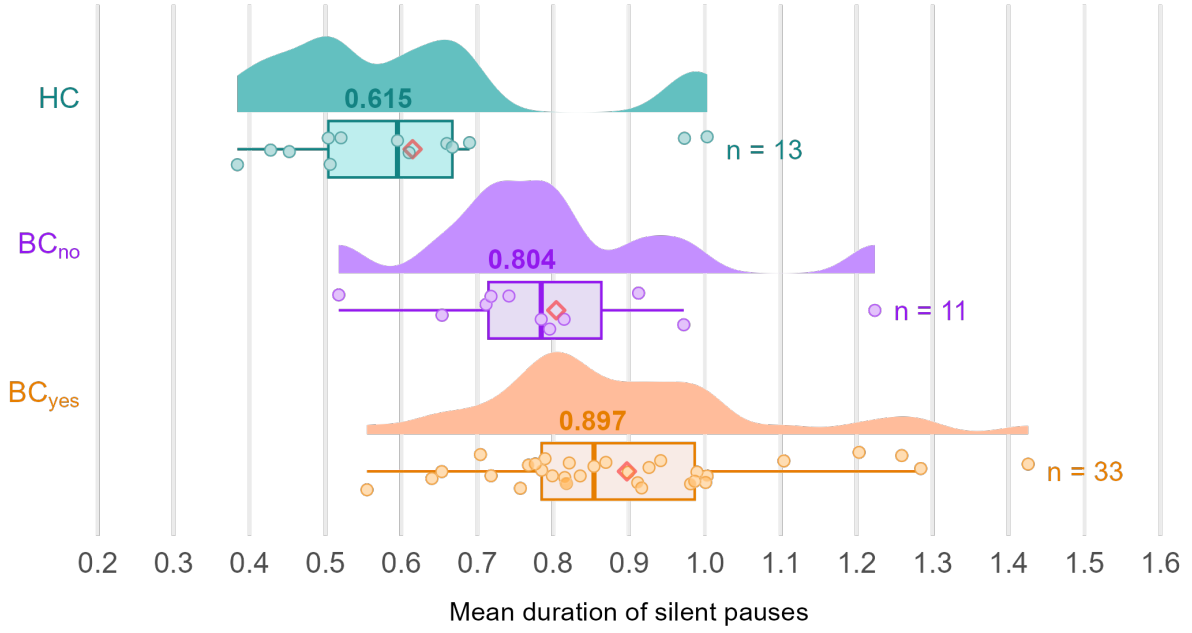


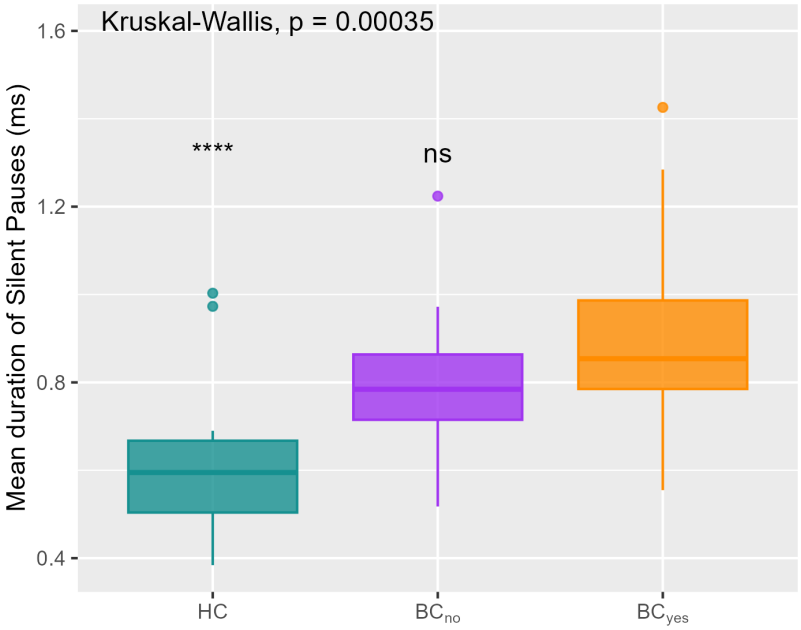
Figure 3 summarizes the distribution of mean silent pause duration. The box plots show the 1st and the 3rd interquartiles, as well as median pause duration. Red diamonds display the mean silent pause duration for each group. Each point represents the mean silent pause duration of a participant. As a group, HC shows little variability in results (apart from two points in the higher values), while BC_{yes} and BC_{no} show larger spreads.

Figure 3. Distribution of mean silent pause duration for each group. Red diamond = mean duration per group. Dots = mean duration per participant.



A Kruskal-Wallis test showed a significant difference in silent pauses across the three groups (X^2 (2, N = 57) = 15.894, p < .001). Post-hoc comparisons using Dunn's test indicated indicated a higher silent pause mean duration for the BC_{yes} compared to the HC group, p < .001. However, there was no significant difference in silent pause duration between the BC_{yes} and BC_{no} groups, p = .143, and between the BC_{no} and HC groups, p = .108. Figure 4 shows the Dunn's test results with BC_{yes} as the reference data, with a significant difference between HC vs. BC_{yes}, and a non-significant difference between BC_{no} and BC_{yes}.

Figure 4. Boxplots showing mean silent pause duration for each group. $****$ = significant difference between HC and BC_{yes}. ns = non-significant difference between BC_{no} and BC_{yes}.



The effect size for unbalanced data between the three groups was measured by Cohen's f and showed a large effect (f = 3.2). However, this result does not explain the overlap between the BC_{yes} and BC_{no} groups. We performed a power analysis to estimate the required number of BC_{yes}, BC_{no} and HC for balanced data to reach significance for mean silent pause duration results. The power analysis indicates that a sample size of 53 participants will provide enough data to expect a medium effect size with a 95% Confidence Interval ; pwr.anova.test(k=3, f=0.4, sig.level=.25, power=.8). A medium effect size will keep the inherent interspeaker variability of speech production.

Filled pauses

A Kruskal-Wallis test showed there was no significant difference in filled pause duration across the three groups (X^2 (2, N = 57) = 1.574, p = .455).

Sustained vowels

A Kruskal-Wallis test showed there was no significant difference in sustained vowel duration across the three groups (X^2 (2, N = 57) = 4.469, p = .107).

Table 2 summarizes the Kruskal-Wallis test results, comparing mean pause duration across groups.

Table 2. Mean (SD) pause duration per group.

Pause type	BC_{yes}	BC_{no}	HC	*p*-value
Silent Pauses	.897 (.192) ∗	.804 (.185)	.615 (.191) ∗	< .001
Filled Pauses	.649 (.197)	.599 (.187)	.520 (.186)	.455
Sustained Vowels	.415 (.078)	.356 (.067)	.381 (.057)	.107

Note: BC_{yes} = breast cancer survivors with a significant cognitive complaint, BC_{no} = breast cancer survivors without any significant cognitive complaint, HC = healthy controls. ∗ = significant difference between BC_{yes} and HC.

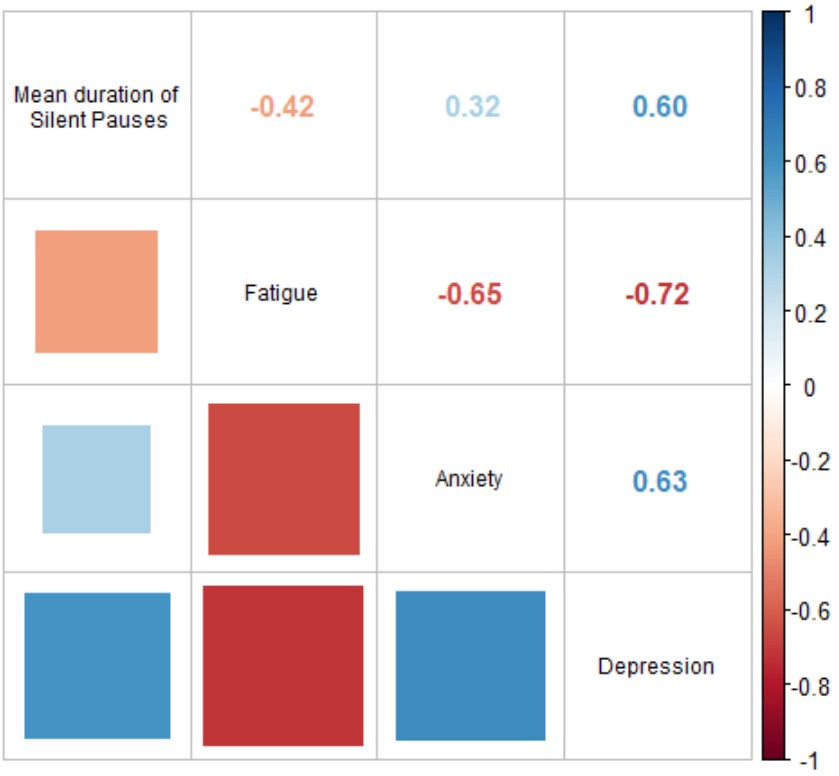
Co-factor control

Some psychological factors or humor states such as depression are known to influence speech production (Albuquerque et al., 2021). We performed a Spearman's rank correlation to examine the relationship between mean pause duration and reported fatigue, anxiety, and depression scores. Note that none of these scores are considered below pathological threshold. Therefore, they do not qualify as clinical observation. Correlations were only conducted for statistical purposes only. Additionally, lower scores on the fatigue scale show more reported fatigue, while higher scores on the anxiety and depression scales signal more reported anxiety and depression.

There was a negative correlation between mean silent pause duration and fatigue scores, r(2) = -.42, p < .001, and a positive correlation between mean silent pause duration and anxiety scores, r(2) = .32, p = .015. In addition, there was a positive correlation between mean silent pause duration and depression scores, r(2) = .60, p <.001. However, correlations coefficients were low, suggesting that mean silent pause duration is poorly influenced by co-factors.

Figure 5 displays the correlation matrix of the relationship between mean silent pause duration and the co-factors. Fatigue scores are negatively correlated with anxiety and depression scores, meaning that participants who reported fatigue also reported anxiety, r(2) = -.65, p < .001, and depression, r(2) = -.72, p < .001.

Figure 5. Correlation matrix of the relationship between the mean silent pause duration and co-factors. A red color indicates a negative correlation. A blue color indicates a positive correlation. Numbers = correlation coefficient. Note that the scores remain below alert thresholds.



There was no correlation between filled pause mean duration and fatigue scores, r(2) = -.15, p = .256, and between mean filled pause duration and anxiety scores, r(2) = .12, p = .382. There was no correlation either between mean filled pause duration and depression scores, r(2) = .13, p = .323.

There was no correlation between mean sustained vowel duration and fatigue scores, r(2) = -.14, p = .290. However, there was a positive correlation between mean sustained vowel duration and anxiety scores, r(2) = .02, p = .826, and between mean sustained vowel duration and depression scores, r(2) = .10, p = .435.

Discussion

The goal of our study was to test the potential use of pauses as a behavioral marker of Cancer-Related Cognitive Impairment. We measured silent pause, filled pause, and sustained vowel duration in a dataset of patients with Cancer-Related Cognitive Impairment and healthy controls. As a secondary goal, we aimed to showcase a detailed method using free off-line tools to deal with clinical data, in continuity with Angelopoulou et al. (2024) who provide a protocol for the analysis of connected speech. In the following subsections, we discuss the main result in our dataset analysis, i.e., the role

of silent pauses in the detection of Cancer-Related Cognitive Impairment. We then discuss the main advantages of our method and point out a number of starting points for further technical development.

Are silent pauses a speech marker of CRCI?

Based on our team's previous studies, we aimed to investigate silent pauses, filled pauses, and sustained vowels in cancer survivors with and without a cognitive complaint (i.e., CRCI) as well as healthy controls. The results show that in all three pause types, only mean silent pause duration can differentiate cancer survivors with CRCI from those without CRCI and healthy controls. Interestingly, this finding echoes other studies on pauses in individuals with different etiologies who suffer from subtle cognitive impairment. For instance, Bóna (2024) examined silent pauses, filled pauses and lengthenings among other types of disfluencies in individuals with Multiple Sclerosis. Multiple Sclerosis (MS) is a condition that impacts motor abilities, but can also cause cognitive impairment (Patti et al., 2015). In a cohort of twenty MS participants with no or mild dysarthria, Bóna (2024) found that only silent pauses were able to differentiate MS patients from healthy controls. This concerned both silent pause duration and frequency, as participants with MS produced more and longer silent pauses than controls.

Despite different etiologies and linguistic systems, silent pauses might better reflect a higher processing load than their voiced counterparts, i.e., filled pauses and sustained vowels. This idea has long been suggested by Ferreira (1993) and is supported by recent evidence in the field. Similarly, Betz et al. (2023) found a higher number of silent pauses compared to filled pauses and lengthenings in an experiment on the impact of cognitive load on speech in a card game with various levels of complexity. However, their results showed that lengthenings also increased with task difficulty. The authors related this phenomenon to the co-occurrence of lengthenings and silent pauses in their data, and to the nature of the speech task they used.

Our study presents a limitation regarding the number of participants, as the data is unbalanced. Our results need to be confirmed with further data since our sample size per group is relatively small. The power analysis we ran to estimate the ideal number of participants indicates that forty-eight participants per group are needed to increase the reliability of the results.

Reshaping speech analysis

We transcribed and annotated fifty-seven speech samples, with a total duration of seventy-nine minutes. Automatic Speech Recognition and Natural Language Processing tools considerably reduced the time spent on speech transcription and annotation.

Speeding up transcription while making it safer

Many private companies offer transcription solutions (e.g., Dragon, Nvivo). Most of them can be accessed either with a fee or for free online, with no guarantee to the use of the data. Such a lack in transparency represents a major ethical concern, especially when handling medical data. We used Whisper as a bypass to these issues. In addition to being time saving, it kept the data on our computer as a command. Whisper is a large language model that can transcribe more than twenty languages from different language families. However, the efficiency of such a model varies considerably across languages. It is measured via Word Error Rate (WER, i.e., the number of words that were poorly transcribed out of the total number of words in the data). Efficiency also depends on model size and audio quality for the recording. Table 3 shows examples of transcription errors from our corpus.

Table 3. Examples of transcription errors using the Whisper large v3 model.

Erroneous transcription	**Correction**	**Nature of the error**
Good audio recording quality		
[les vagues] qui **onduent**	[les vagues] qui **ondulent** *[the waves] ripple*	*Phonology*: The last phoneme is missing /ɔ̃dy/ and /ɔ̃dyl/
son scooter des **mères**	son scooter des **mers** *his water scooter*	*Semantics*: homophones « mère » (mother) and « mer » (sea)
il prend quelque chose dans la **glaciaire**	il prend quelque chose dans la **glacière** *he takes something in the cooler*	*Semantics*: homophones « glaciaire » (glacial) and « glacière » (cooler)
un monsieur **où** papa	un monsieur **ou** papa *a man or a dad*	*Syntax*: homophones « où » (where) et « ou » (or)
la planche à voile du papa qui du coup est tomb**ée** dans l'eau	la planche à voile du papa qui du coup est tomb**é** dans l'eau *the dad's windsurf who has fallen off*	*Morphosyntax*: the image depicts a man drowning (the father has fallen). Whisper makes the verb "tomber" (to fall) agree with the feminine noun "planche à voile" (windsurf) instead of the masculine noun "papa" (dad) because of the ambiguity of the pronoun "qui" (translated in this case "who", but also "which" in other contexts).
Poor audio recording quality		

(To be continued)

(Continued)

Erroneous transcription	Correction	Nature of the error
une planche de **sœur**	une planche de sur**f** *a surf board*	*Phonology*: the last phoneme is missing, /sœr/ (sister) and /sœrf/ (surf)
pas loin des coco**tis**	pas loin des coco**tiers** *not far from the coconut tree*	*Phonology*: the last phoneme is missing /kokoti/ and /kokotje/
sans que ni la maman ni le petit garçon ne remarqu**e**	sans que ni la maman ni le petit garçon ne remarqu**ent** *without either the mother or the little boy noticing*	*Morphosyntax*: absence of the plural mark to the verb "remarquer" (to notice) because Whisper does not take into account the link between "maman" (mom) and "garçon" (kid) represented by "ni … ni" (either)

Note. Good audio recording quality: speech samples recorded in a soundproof booth. Poor audio recording quality: speech samples recorded in a quiet room at the hospital, with echoes, background noise, mumblings from the corridors, and occasional outside noise (e.g., cars and helicopter).

Depending on the WER for a language, transcriptions need more or less inspection to fix potential errors. Larger models reduce WER but demand more computer power (CPU/GPU), and increase transcription time. In any case, we strongly recommend to manually checking transcripts to avoid errors. Model accuracy can be improved with Whisper's open source license. However, Whisper uses deep learning technology that is hard to fine-tune. In addition, Whisper automatically removes pauses and disfluencies, and is not yet able to recognize multiple speakers. There is thus a need for easy-to-use scripts that include pause and disfluency detection as well as multiple speaker recognition.

Improving precision in prosody analysis

Annotating speech is a repetitive task that requires high attentional resources. We used SPPAS to annotate the fifty-seven samples of our dataset in a few minutes. SPPAS is convenient in that it performs automatic annotation, allows the user to edit these annotations, and provides basic statistics in a single software. SPPAS handles French and English by default, and other languages can be implemented. Just like Whisper, the accuracy of the automatic annotation provided by SPPAS largely depends on audio quality. In the case of our dataset, this translated into segmentation errors during the silence detection and/or phonetization phases. We had a small number of misplaced boundaries for good-quality recordings, but this number increased with poor audio quality. We manually corrected the errors by aligning the boundaries to the audio signal. However, SPPAS lacks precision in acoustic signal visualization. We used Praat to overcome this limit.

Praat is the most used tool for speech analysis in research related to language and speech. It

benefits from a large, active community contributing many scripts for specific purposes. Praat is a powerful tool for visualizing, editing and analyzing various aspects of speech such as voice quality, rhythm, or intonational features. While Praat has become more user-friendly over time, a training phase is still required to exploit its ample possibilities.

Future prospects for speech analysis

Our secondary goal was to provide a ready-to-use method for speech analysis. The method we propose is still limited by the fact it uses three different software programs, and by the fact it requires some computer skills. Furthermore, while this method considerably reduces time and biases, speech analysis remains a long process that may not be applicable to all patients in clinical practice. Further research should especially develop software combining all three tools (i.e., Praat, Whisper and SPPAS) for clinical purposes, such as the Batchalign automated pipeline (Liu et al., 2023) which focuses on other aspects of language such as semantics and morphosyntax.

Research Highlights

- Cancer survivors are at risk to develop subtle cognitive impairment that is not detected by current neuropsychological tests.
- Speech analysis can be a powerful tool in the detection of subtle cognitive impairment.
- However, the absence of simple, detailed methods compromises the feasibility of such analyses in clinical practice.
- The main goal of this study is to demonstrate this method with a case in point application to breast cancer survivors with a cognitive complaint.
- In addition, we propose a semi-automatized method for speech analysis that provides guidance on free, secure tools for medical data.

Conclusion

We aimed to distinguish breast cancer survivors with a cognitive complaint from those without and healthy controls measuring pause duration. For this particular dataset, we had hypothesized we would find longer silent pauses, filled pauses and sustained vowels for breast cancer participants with a cognitive complaint than for breast cancer participants without any complaint and healthy controls, based on previous studies from our team. We did find a significant difference in silent pause duration only across the three participant groups, with post-hoc comparisons showing differences between cancer patients with a cognitive complaint and healthy controls, but no difference concerning

cancer patients without a cognitive complaint. No difference was found across groups for filled pause duration and sustained vowel duration. In agreement with other studies on subtle cognitive impairment, our results support the claim that silent pauses are more indicative of a higher processing load than voiced disfluencies.

We proposed a detailed semi-automatized method for speech analysis using open source, off-line tools to deal with clinical data. This methodology speeds up annotation while protecting the data. We applied this methodology to a dataset of patients with Cancer-Related Cognitive Impairment. This method is intended to help the diagnosis of subtle cognitive impairment via prosody. However, while our method and tutorials are ready to use, our protocol still relies on the use of several pieces of software. This calls for the development of automated pipelines combining different tools in order to make speech analysis even more applicable in clinical practice.

Acknowledgements

The authors would like to thank Farha Gheurbi and Marion Buchy for performing the inter-annotator agreement.

Authors' contributions

Amélie B. Richard: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration, Funding acquisition. **Manon Lelandais:** Conceptualization, Writing - Original Draft, Writing - Review & Editing. **Sophie Jacquin-Courtois:** Resources, Writing - Review & Editing, Supervision.

Funding

Amélie B. Richard is financed by a Ph.D. grant awarded by Paul-Valery Montpellier 3 University and funded by the French Ministry of Higher Education (Ministère de l'Enseignement Supérieur et de la Recherche - MESRI).

Supplementary material

Our speech dataset and participant demographics are not available due to medical rules. Our full protocol (Pause_analysis_protocol), our Whisper install tutorial (Whisper_install), and our code for cleaning transcripts (transcript_cleaner) are available in the Speech analysis repository, https://osf.io/q8vyb/?view_only=93d5c4b21abe424b8aed4803df5b87c1

Conflicts of interest/Competing interests

The authors declare no conflict of interests.

Ethics approval

The present study is part of three larger projects. The first project aims to evaluate the effectiveness of a physical rehabilitation program offered to women with breast cancer who have completed the main phase of their treatments (ACTICOG), and the second one evaluates the effectiveness of a cognitive rehabilitation program for individuals with subtle cognitive impairment due to stroke or cancer (REMCOG). They both have been approved by the ethics committee of the Hospices Civils de Lyon (respectively n°22-5009 and n°22-121). The third project focuses on cognitive impairment induced by cancer and its treatment in women who have had breast cancer (DisCCo), and was validated by the ethics committee Comité de Protection des Personnes Est-III in August 2023.

Consent to participate

Participants have provided written informed consent to participate.

Consent for publication

Participants have provided written informed consent to publish their data.

Availability of data and materials

Our speech dataset and participant demographics are not available due to medical rules. Our full protocol (Pause_analysis_protocol) and Whisper install tutorial (Whisper_install) are available in the Speech analysis repository, https://osf.io/q8vyb/?view_only=93d5c4b21abe424b8aed4803df5b87c1

Code availability

Our code for cleaning transcripts (transcript_cleaner) is available Speech analysis repository, https://osf.io/q8vyb/?view_only=93d5c4b21abe424b8aed4803df5b87c1

Vitae

Amélie B. Richard is a certified Speech-language Pathologist and a Ph.D. Candidate in Language Science (Université de Montpellier Paul-Valéry). Her research focuses on speech markers of subtle cognitive impairment, particularly in cancer patients, with a special interest in developing automatic speech-based diagnosis tools.

Manon Lelandais is an Associate Professor in Language Science (Université Paris Cité). She studies the interplay between the verbal, vocal, and gestural resources we use when we talk, to deepen our understanding of the real-time mechanisms of speech production.

Sophie Jacquin-Courtois is an M.D. in Physical Medicine and Rehabilitation and holds a Ph.D. in Neuropsychology. She currently works at the Neurology Rehabilitation unit of the Henry Gabrielle Hospital (Hospices Civils de Lyon) where she develops programs for women cancer survivors.

Méthode d’analyse des pauses pour l’évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment

---

Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment.

Auteurs :
Amélie Richard[1,2]
Fabrice Hirsch[1]
Sophie Jacquin-Courtois[2,3]

Affiliations :
[1]Université Montpellier 3, Praxiling, CNRS UMR 5267, Montpellier, France.
[2] Université de Lyon, Centre de Recherche en Neurosciences de Lyon, CNRS UMR 5292, INSERM U1028, Lyon, France.
[3] Hôpital Henry Gabrielle, Hospices Civils de Lyon, Lyon, France.

Autrice de correspondance :
Amélie Richard
amelie.richard@inserm.fr

Dates :
Soumission : 24/07/2024
Acceptation : 25/11/2024
Publication : 28/12/2024

Comment citer cet article :
Richard, A., Hirsch, F. & Jacquin-Courtois, S. (2024).Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment. *Glossa*, 142, 49-64. https://doi.org/10.61989/qp2vb532

e-ISSN :
2117-7155

Licence :
Ce travail est disponible sous licence Creative Commons Attribution 4.0 International.

Contexte. L'analyse de la parole permet de repérer des troubles cognitifs subtils, notamment car la prosodie contient des éléments fins, comme les pauses, qui peuvent être des marqueurs comportementaux de Troubles NeuroCognitifs. Cependant, l'absence de méthodes simples et détaillées compromet la faisabilité d'une telle analyse en clinique. Les Troubles NeuroCognitifs mineurs (TNCm) désignent un changement cognitif rapporté par un individu pour lequel l'autonomie est préservée. Cette appellation recouvre de nombreuses physiopathologies sous-jacentes, tel que le Cancer-Related Cognitive Impairment (CRCI). Ce trouble cognitif induit par un cancer et ses traitements, est un TNCm principalement caractérisé par une plainte mnésique (« j'oublie ce qu'on me dit ») et linguistique (« je cherche mes mots ») dont la fréquence et l'intensité ne sont pas corrélées aux scores des tests orthophoniques ou neuropsychologiques. Le manque d'outil d'évaluation suffisamment sensible pour déceler ce trouble subtil conduit souvent à un sous-diagnostic. De nouvelles méthodes, comme l'analyse de la parole, sont donc souhaitables pour évaluer le CRCI.

Objectifs. Les objectifs de cette étude sont (i) de proposer une méthode d'analyse des pauses applicable en clinique, (ii) d'identifier les personnes porteuses d'un CRCI grâce à cette méthode.

Méthode. Treize participantes post-cancer du sein et treize participantes contrôles saines ont été incluses. Toutes ont eu pour instruction de raconter une histoire à partir d'une séquence d'images qui leur était présentée. Leur production orale a été enregistrée, puis transcrite automatiquement avec Whisper et analysée sur les logiciels SPPAS et Praat. Les pauses silencieuses, pauses remplies (« euh ») et allongements vocaliques (« un petiiit garçon ») ont été annotés, puis ont fait l'objet d'une analyse statistique sur JASP.

Résultats. La durée des pauses silencieuses des participantes post-cancer du sein était significativement plus longue que celle des contrôles. En revanche, la durée des pauses remplies et des allongements vocaliques n'a pas montré de différence significative entre les deux groupes. De même, le taux de pauses (nombre d'occurrences / durée totale du discours) ne montre pas de différence significative entre les deux groupes.

Conclusions. Cette étude décrit en détails les étapes d'analyse des pauses en vue de son applicabilité clinique. Les résultats suggèrent que la durée des pauses silencieuses permet de différencier les participantes post-cancer ayant une plainte cognitive des participantes contrôles.

Mots-clefs : Troubles NeuroCognitifs mineurs ; Cancer-Related Cognitive Impairment ; pauses ; transcription automatique ; parole ; prosodie.

Method for the analysis of pauses for assessing subtle cognitive impairment applied to Cancer-Related Cognitive Impairment.

Background. Speech analysis can detect subtle cognitive impairment, particularly because prosody contains fine elements, such as pauses, which might be behavioral markers of NeuroCognitive Disorders. However, the absence of simple, detailed methods compromises the feasibility of such an analysis in clinics. Mild Neurocognitive Disorders (mNCD) refer to a cognitive change reported by an individual whose autonomy is preserved. This terminology covers many underlying pathophysiologies, such as Cancer-Related Cognitive Impairment (CRCI). CRCI is an mNCD characterized by a memory complaint (e.g., "I forget what I'm told") and a language complaint (e.g., "I search for my words") for which the frequency and intensity exceed the scores of speech-language tests or neuropsychological tests. The lack of sensitive tools to detect this subtle impairment often leads to its underdiagnosis. New methods such as speech analysis are needed to assess CRCI.

Objectives. The study aims are (i) to propose a method for analyzing pauses which is practical in clinical context, (ii) to identify participants with CRCI using this method.

Methods. Thirteen participants post breast cancer and thirteen healthy controls were included in the study. The participants were instructed to tell a picture-based story. Their narratives were recorded, then automatically transcribed with Whisper and analyzed using SPPAS and Praat software. Silent pauses, filled pauses (e.g., "uh"), and sustained vowels (e.g., "a smaall boy") were annotated, then examined on JASP for a statistical analysis.

Results. The duration of silent pauses in post-cancer participants was significantly longer than that of the controls. However, the duration of filled pauses and sustained vowels did not show a significant difference between the two groups. Similarly, the rate of pauses (i.e., number of occurrences/total speech duration) did not show a significant difference between the two groups.

Conclusions. This study provides a well-described method for pause analysis designed for the purpose of clinical context. The results suggest that the duration of silent pauses is a good marker for differentiating post-cancer participants from controls.

Key-words: Mild NeuroCognitive Disorders ; Cancer-Related Cognitive Impairment ; pauses ; automatic transcription ; speech ; prosody.

CONTEXTE

Cet article propose un protocole permettant de conduire une analyse de la prosodie chez des patients ayant des troubles cognitifs subtils. Cette méthode se veut une aide au diagnostic de troubles cognitifs subtils. Elle s'adresse aux orthophonistes diplômés, aux étudiants en orthophonie et aux chercheurs ou cliniciens s'intéressant aux questions relatives à la fluence.

Marqueurs linguistiques de troubles cognitifs

La littérature en aphasiologie, vasculaire ou neurodégénérative, a montré l'intérêt de l'analyse de la parole et de la voix des patients pour révéler les troubles cognitifs. En effet, son caractère non invasif, sa sensibilité aux changements cognitifs subtils et sa haute validité écologique en font un outil de choix pour évaluer la cognition des patients (Lanzi et al., 2023). La parole peut être étudiée à travers le prisme de la phonétique, c'est-à-dire des éléments segmentaux de la parole (les phonèmes), et celui de la prosodie (He et al., 2023), autrement dit des éléments supra-segmentaux de la parole (fréquence fondamentale (Hz), intensité (dB) et durée (ms)). La prosodie se révèle particulièrement pertinente pour investiguer les Troubles NeuroCognitifs mineurs (TNCm) car elle contient des éléments fins signalant une modification de la cognition d'un locuteur (Ivanova et al., 2023). Ces éléments prosodiques comprennent des paramètres liés à la qualité de la voix, tels que des variations d'amplitude intonative (Themistocleous et al., 2020), ainsi que des unités plus larges comme les pauses (Vincze et al., 2021). Parjane et al. (2021) ont comparé la production orale de discours descriptifs de patients ayant une Aphasie Progressive Primaire non-fluente agrammatique, à celle de patients ayant une Démence à Corps de Lewy et de participants contrôles sains. Les résultats montrent des pauses plus longues, plus fréquentes et une réduction de la fréquence fondamentale dans les deux groupes de patients par rapport aux locuteurs contrôles.

Une des problématiques soulevées par ces recherches est liée à la variabilité des méthodologies employées (type de discours, variables linguistiques, outils d'analyse), rendant les résultats entre études difficilement comparables. Dans un effort d'harmonisation, Angelopoulou et al. (2024) ont proposé une méthode d'analyse de la parole comparant deux types d'élicitation de discours descriptifs (une description d'images et une description autobiographique) chez une population saine. L'intérêt de leur étude réside principalement sur la description des mesures de dix-neuf variables de parole, bien que certains concepts comme « mot » ou « pause » ne soient pas explicités (les auteurs ne différencient pas les pauses dues à des contraintes physiologiques des pauses ayant une fonction démarcative ou d'hésitation).

Pour autant, la faisabilité de leur méthodologie en pratique clinique est questionnable. En effet, la méthode d'analyse proposée par cette équipe est entièrement manuelle. Or, les étapes de transcription et d'annotation sont chronophages et exigent une rigueur pour être réplicables, généralisables ou simplement interprétables, les rendant difficiles à appliquer en routine clinique. Une méthode détaillée et facile à exécuter dans l'exercice de la clinique serait donc souhaitable.

Analyse (semi-)automatique de la parole

Le choix des outils utilisés détermine la qualité d'une analyse de la prosodie. Cette section argumente en faveur de l'utilisation d'outils informatiques automatisés, hors-ligne et *open source*.

Importance des outils automatisés

L'analyse de la parole est une tâche longue et fastidieuse. Après avoir enregistré la production orale d'un patient, il est souvent utile de la transcrire soit de manière orthographique, soit de manière phonétique. La transcription manuelle est particulièrement chronophage. En effet, la transcription d'une minute de parole peut prendre jusqu'à douze minutes (Bryant et al., 2017). Cette étape est également coûteuse en termes de concentration puisqu'elle demande plusieurs écoutes de l'enregistrement afin de ne pas omettre de mots ou de disfluences produites par le locuteur. Par exemple, il est difficile de connaître le nombre exact de répétitions à la première écoute. Le processus d'écoutes multiples entraîne des biais liés à la fatigabilité auditive, des biais de subjectivité et d'habituation, augmentant ainsi les risques d'erreurs de transcription (comme un ajout de mots par effet d'amorçage).

L'utilisation d'outils dotés d'Intelligence Artificielle (outils informatisés qui automatisent une tâche cognitive) permet de réduire à la fois le temps consacré à la tâche de transcription et à ses

biais en apportant par exemple une première transcription « neutre ». Ces outils sont issus de technologies plus ou moins complexes, de la simple tâche automatisée (détection automatique des pauses silencieuses) aux modèles de réseaux neuronaux (transcription automatique) et apportent un gain de temps considérable pour les analyses de la parole. Pour autant, les outils d'IA peuvent montrer des performances variables, notamment selon la qualité audio de l'enregistrement. Il est donc indispensable d'effectuer une vérification manuelle afin de corriger les éventuelles erreurs commises par ces outils.

Importance des outils hors-ligne

L'enregistrement de la voix de patients ou de participants constitue une donnée identifiante qui doit être protégée à un haut niveau de sécurité (Commission Nationale de l'Informatique et des Libertés, 2016). C'est pourquoi le choix d'outils transparents en matière de gestion des données est crucial. Pour autant, les ressources disponibles sur internet ont parfois des politiques de confidentialité obscures, ce qui accroît le risque de fuite des données. Les outils hors-ligne décrits dans cette étude répondent aux critères de sécurité requis pour l'analyse de la parole au sens que les données sont traitées depuis leur lieu de stockage.

Importance des outils open source

Les logiciels *open source* sont des logiciels gratuits dont les codes de programmation sont accessibles à tout utilisateur. Ce type d'outil collaboratif non seulement garantit une totale transparence en matière de gestion des données, mais favorise également la reproductibilité scientifique et la fiabilité des résultats.

Troubles NeuroCognitifs mineurs

Les troubles neurocognitifs désignent une modification des performances cognitives constatée par l'individu ou ses proches, et sont dits mineurs lorsque l'autonomie de l'individu est préservée (American Psychiatric Association, 2022). Les TNCm couvrent un large éventail de physiopathologies sous-jacentes telles que les pathologies neurodégénératives (maladie d'Alzheimer à un stade précoce, Démence à Corps de Lewy, etc.), les pathologies vasculaires (comme les accidents vasculaires cérébraux) ou traumatiques (traumatisme crânien léger) et les contextes infectieux ou oncologiques (séquelles de Covid-19, VIH, troubles cognitifs liés au cancer, etc.). L'appellation TNCm recouvre également les syndromes comportementaux ou cognitifs tels que le Mild Cognitive Impairment, pour lesquels le diagnostic d'une étiologie n'est pas établi.

La prévalence des TNCm est difficile à estimer du fait de la diversité des possibles étiologies et physiopathologies sous-jacentes. Il est donc possible que ces chiffres soient sous-estimés. Dans le cadre du traumatisme crânien léger, 10 à 15% des patients peuvent développer un syndrome post-commotionnel persistant incluant des troubles cognitifs (Groupe de travail Traumatisme Crânio-Cérébral Léger -TCCL- en Ile-de-France, 2021). Dans une étude menée par Hommel et al. (2009) auprès de soixante-quatorze patients post-AVC mineur ayant un score au *Mini Mental State Examination* (MMSE, Folstein et al., 1975) dans la norme (> 27/30), 70% des patients (soit 52/74 patients) rapportaient des difficultés de fonctionnement social corrélées à un déficit en mémoire de travail. Ce résultat semble indiquer la présence élevée de possibles troubles cognitifs après un AVC mineur. Par ailleurs, les auteurs suggèrent que les tests de dépistage cognitif usuels manqueraient de sensibilité à repérer ces troubles fins.

Lors de consultations cliniques, les patients ou leur entourage peuvent rapporter l'apparition de difficultés cognitives discrètes mais invalidantes. Cependant, leur plainte est rarement corrélée aux scores des tests de repérage cognitif tels que le MMSE ou le *Montreal Cognitive Assessment* (MoCA, Nasreddine et al., 2005). De fait, les TNCm restent souvent sous-diagnostiqués et il est rappelé par la Haute Autorité de Santé que toute plainte cognitive inquiétante, en particulier chez le sujet âgé, doit faire l'objet d'une investigation approfondie (Haute Autorité de Santé, 2018). Néanmoins, cette même vigilance se doit d'être appliquée avec la même rigueur à une population possiblement plus jeune dans d'autres contextes (AVC mineur, traumatisme crânien léger, Cancer-Related Cognitive Impairment, etc.) dont l'enjeu sanitaire et médico-économique est important.

Cancer-Related Cognitive Impairment

Parmi les TNCm les plus fréquents, le Cancer-Related Cognitive Impairment (désormais CRCI) tient une place singulière qualitative et quantitative. En effet, il s'agit d'un trouble cognitif subtil induit

par un cancer (hors cancer du système nerveux central) et/ou ses traitements, qui peut concerner jusqu'à 75% des patients en oncologie (Janelsins et al., 2011). Outre son incidence élevée, il altère la qualité de vie des patients. Pour certains, la symptomatologie est durable, pouvant persister plus de deux ans après l'arrêt des traitements oncologiques. La permanence du trouble peut entraîner une limitation de la participation aux activités sociales et entraver la reprise professionnelle d'une population possiblement jeune. Le caractère subtil de l'expression de ses troubles en fait un handicap invisible souvent sous-diagnostiqué, alors même que des possibilités d'évaluation appropriée et de prises en charge existent.

Le CRCI est caractérisé par une plainte cognitive touchant plusieurs aspects de la cognition (mémoire, fonctions exécutives, attention, vitesse de traitement, langage) pouvant être objectivée par des questionnaires standardisés (comme le FACT-Cog de Wagner et al., 2009). Une des plaintes les plus fréquentes est le manque du mot. Malgré l'intensité de la plainte, les tests orthophoniques et neuropsychologiques actuels, généralement conçus pour évaluer des troubles plus sévères, peinent à révéler pleinement les difficultés rapportées par les patients. Des études comparant des patientes ayant reçu un traitement par chimiothérapie pour un cancer du sein et des participantes contrôles saines montrent une différence significative pour les compétences verbales (Amani et al., 2024 ; Jim et al., 2012). Toutefois, la taille d'effet de cette différence est relativement petite, ce qui signifie que les scores des patientes sont proches de ceux des contrôles. Dès lors, il est possible que les scores obtenus notamment aux épreuves de fluence verbale et de dénomination administrées en clinique, soient souvent dans les normes. De nouvelles approches et/ou de nouveaux outils sont donc à développer pour détecter ces troubles cognitifs subtils.

OBJECTIF ET HYPOTHÈSES

Concernant les marqueurs prosodiques du CRCI, les études pilotes menées par notre équipe ont montré l'intérêt de l'analyse des disfluences (Richard et al., 2022) et particulièrement des pauses dans la production discursive de patientes ayant eu un cancer du sein qui ont une plainte cognitive (Richard et al., 2023). L'analyse des pauses d'hésitation produites lors d'une tâche de narration élicitée par des images se révèle suffisamment sensible pour différencier des patientes post-cancer du sein ayant une plainte cognitive de participantes contrôles (Richard et al., 2024).

Les objectifs de la présente étude sont (i) de proposer une méthode d'analyse prosodique applicable en clinique, (ii) d'appliquer cette méthode à l'analyse des pauses chez des femmes ayant une plainte cognitive des suites d'un cancer du sein.

Nous émettons l'hypothèse que la méthode d'analyse que nous proposons permettrait de mettre en évidence un CRCI chez des patientes ayant eu un cancer du sein. Les pauses d'hésitation seraient plus fréquentes (en nombre/durée de discours) et plus longues (en ms) pour les participantes ayant une plainte cognitive liée au cancer que pour les participantes contrôles saines.

MÉTHODE

Démarches éthiques

La présente étude est issue de deux projets plus larges. Le premier a pour objectif d'évaluer l'efficacité d'un programme de réadaptation physique proposé à des femmes ayant eu un cancer du sein à l'issue de la phase principale de leurs traitements (ActiCog) et a été approuvé par le comité éthique des Hospices Civils de Lyon (n°22-5009). Le second projet porte sur le trouble cognitif induit par un cancer et ses traitements chez des femmes ayant eu un cancer du sein (DisCCo) et a été validé par le Comité de Protection des Personnes Est-III en août 2023. L'étude est enregistrée au numéro national 2023-A00267-38.

Population

Treize participantes ayant eu un cancer du sein (six participantes issues de l'étude ActiCog et sept participantes issues de l'étude DisCCo) et treize participantes contrôles saines (étude DisCCo) appariées en âge et en niveau de scolarisation ont été éligibles pour cette étude.

Critères d'inclusion

Les critères d'inclusion des participantes ayant eu un cancer (groupe Breast Cancer, désormais BC) étaient (i) d'avoir terminé les traitements curatifs (chirurgie, chimiothérapie, radiothérapie) depuis plus de quatre semaines et moins d'un an, hors hormonothérapie, (ii) d'avoir une vue et une

audition normales ou corrigées, (iii) d'avoir un score à la sous-échelle *Perceived Cognitive Impairment* (PCI) du questionnaire FACT-Cog inférieur ou égal à 55/72. Le questionnaire FACT-Cog (Wagner et al., 2009) est un questionnaire standardisé d'évaluation de la plainte cognitive des patients en oncologie pour lequel un seuil inférieur ou égal à 55/72 à la sous-échelle PCI indique la présence d'une plainte cognitive significative (Van Dyk et al., 2020).

Les critères d'inclusion des participantes contrôles (groupe Healthy Controls, désormais HC) étaient les mêmes que les participantes ayant eu un cancer du sein, excepté pour le score PCI du FACT-Cog (score > 55/72).

Critères d'exclusion

Les critères d'exclusion des participantes BC étaient (i) la présence de métastases, (ii) des antécédents de troubles du langage oral, (iii) des antécédents neurologiques ou psychiatriques, (iv) la prise de médicaments susceptibles d'entraîner une modification de la cognition.

En plus des critères d'exclusion énoncés ci-dessus, les participantes HC devaient ne jamais avoir eu de cancer.

Tâche de narration

Il était demandé aux participantes de produire une tâche de discours narratif à partir d'une séquence de cinq images (Bézy et al., 2016). Les images racontent l'histoire d'une famille qui arrive à la plage. Le père fait de la planche à voile pendant que son fils et sa femme sont sur le rivage. Soudain, le vent se lève et le père tombe à l'eau. La dernière image dépeint la mère et le fils alarmés, la mère appelant les secours au téléphone. Les participantes ont pour consigne de retracer l'histoire en images qu'elles voient, sans limite de temps.

Matériel et conditions d'enregistrement

Les productions discursives des participantes à l'étude ActiCog ont été enregistrées dans une pièce calme de l'hôpital de rééducation Henry Gabrielle (Hospices Civils de Lyon, Saint Genis Laval, France) avec un micro-cravate omnidirectionnel Rode Lavalier Go® relié à un ordinateur. Le micro était placé à quinze centimètres de la bouche des participantes.

Les productions discursives des participantes à l'étude DisCCo ont été enregistrées dans une cabine anéchoïque du Centre de Recherche en Neurosciences de Lyon (Bron, France) avec un micro-cravate omnidirectionnel Shure® relié à une carte son de marque M-Audio® et branché à un ordinateur. Le micro était placé à quinze centimètres de la bouche des participantes.

Paramètres linguistiques

Nous avons analysé les pauses silencieuses ainsi que les pauses remplies et allongements vocaliques liés à l'hésitation. Nous avons analysé les pauses égales ou supérieures à 200 ms. Ce seuil tient en effet compte de paramètres liés à la production de la parole, comme la durée des silences acoustiques des consonnes occlusives qui est inférieure à cette valeur. Il correspond également à une limite perceptive à partir de laquelle la pause est significativement perçue (Candea, 2000 ; Goldman et al., 2010). Toutes les pauses silencieuses au-dessus de ce seuil ont été incluses. Pour être considérés comme des pauses remplies et des allongements vocaliques, les phonèmes de plus de 200 ms devaient avoir une courbe intonative plate (Campione & Véronis, 2002 ; Duez, 2001). Suivant ces critères, nous avons identifié trois types de pauses :

- Les pauses silencieuses : segments caractérisés par l'absence de parole, de toux ou de bâillement d'une durée supérieure ou égale à 200 ms,

- Les pauses remplies : segments voisés de type « euh », « hum » d'une durée supérieure ou égale à 200 ms ayant une courbe intonative (F0) plate,

- Les allongements vocaliques : tenue d'un phonème vocalique d'une durée supérieure ou égale à 200 ms dont la courbe intonative (F0) est plate.

Pour chaque type de pauses, nous avons extrait la durée moyenne, médiane et totale, ainsi que la fréquence totale pour calculer le taux de pauses (nombre d'occurrence/durée totale du discours, pauses silencieuses incluses).

Analyse des pauses

La section suivante décrit les différentes étapes de l'analyse prosodique réalisée à partir du signal acoustique. Les différentes phases de la chaîne de traitement peuvent être conduites indépendamment pour adapter l'analyse aux besoins cliniques ou au focus d'une étude. Le temps de réalisation de cette procédure dépend largement de la durée et du nombre d'extraits

audio analysés, ainsi que du niveau de profondeur de l'analyse. Pour cette étude, l'analyse des vingt-six enregistrements représentant trente-cinq minutes de parole a duré environ deux heures dont quatre-vingts minutes de transcription automatique.

Nous avons rédigé un tutoriel complet détaillant chaque étape de notre analyse, disponible sur la plateforme OSF, fichier Protocole_annotation_pauses[1].

Les tutoriels que nous proposons se veulent accessibles à tous et ne requièrent pas de compétences particulières en informatique.

Pré-traitement du signal acoustique

Avant de commencer nos analyses prosodiques, nous avons segmenté l'enregistrement audio pour ne conserver que la production narrative en supprimant les parties qui ne seront pas étudiées (comme l'intervention de l'examinateur). Cette étape a été effectuée avec le logiciel libre Audacity (Audacity Team, 1999). Cette étape a duré environ cinq minutes.

Transcription

Nous avons réalisé les transcriptions automatiquement avec le modèle large V3 du code Whisper (Radford et al., 2022). Whisper est une commande *open source* dédiée à la reconnaissance automatique de la parole. Une de ses fonctions consiste à transcrire un discours oral en texte rédigé à partir du signal acoustique. Une description de l'installation et de l'utilisation de Whisper (testée avec deux volontaires naïfs de la programmation informatique) est disponible sur la plateforme OSF, fichier Installation_Whisper_SPPAS_Praat[2].

Nous avons vérifié chaque texte afin de corriger les erreurs de transcription. Les erreurs étaient peu nombreuses et concernaient surtout les noms propres et les troncatures de mots pour les enregistrements de qualité moyenne. Nous avons ensuite ajouté manuellement les pauses silencieuses (#) et les pauses remplies (euh). Il est important de noter que Whisper supprime par défaut toutes les disfluences (pauses, répétition, faux-départ, etc.).

Enfin, nous avons automatiquement nettoyé les données en supprimant par exemple les marques de ponctuation ou en remplaçant les chiffres par leur équivalent en lettres, grâce à un code Python développé pour cette étude. Ce code est disponible sur la plateforme OSF, fichier transcript_cleaner.py[3].

Cette étape a duré environ une heure et demie dont quatre-vingts minutes de transcription automatique.

Annotation des pauses

Nous avons annoté les pauses en utilisant le logiciel SPPAS (Bigi, 2015). SPPAS est un logiciel *open source* spécifiquement développé pour l'analyse de la parole. Il fournit de nombreuses analyses automatiques telles que la détection des pauses silencieuses, la conversion graphème-phonème ou encore la syllabation de la transcription alignée sur le signal acoustique. Son interface graphique et son caractère interopérable (possibilité de fonctionner avec d'autres logiciels) en font un outil autant applicable en recherche qu'en clinique. Pour affiner nos analyses, nous avons également utilisé le logiciel libre Praat (Boersma & Weenink, 2024). Celui-ci a été conçu pour l'analyse du signal acoustique et permet de visualiser avec précision le signal pour en extraire de nombreuses variables prosodiques (la fréquence fondamentale p. ex.).

Le paragraphe suivant décrit les étapes que nous avons réalisées pour la phase d'annotation du corpus. L'ensemble de ces étapes a duré environ trente minutes.

Nous avons d'abord chargé les enregistrements (format `.wav`) et les transcriptions (format `.txt`) dans SPPAS. Depuis l'onglet « Annotation », nous avons recherché les unités inter-pausales (*Inter Pausal Units*, IPUs), puis lancé la complétion automatique de la transcription. La première fonction, « Search for IPUs », détecte automatiquement les pauses silencieuses et la seconde, « Fill in IPUs », complète les espaces vides par la transcription des segments de parole. Ces fonctions sont accessibles via le bouton « STANDALONE ». Nous avons choisi le format de sortie « `.TextGrid` », format traité par Praat.

Nous avons ensuite procédé à une étape de vérification en ouvrant les fichiers `.TextGrid` dans Praat. Nous avons réajusté les frontières des pauses silencieuses quand celles-ci ne correspondaient pas au signal acoustique (cf. figure 1). Les erreurs d'alignement étaient nombreuses pour les enregistrements de qualité moyenne (avec bruit de fond) et rares pour les enregistrements de bonne qualité.

[1,2,3] https://osf.io/sqwnz/?view_only=8947f20e707f476c9cf44a35ca9d7c32

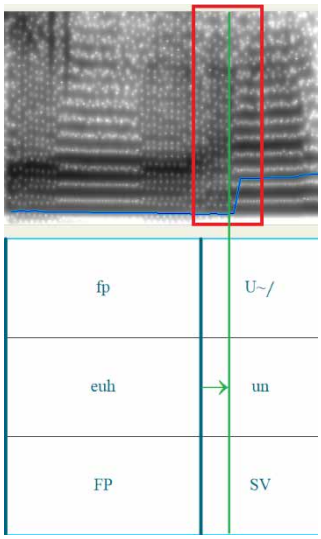


FIGURE 1 : Exemple d'erreur d'alignement. La pause « fp » (filled pause, pause remplie) est partiellement tronquée et la frontière (barre bleue) doit être tirée vers la droite jusqu'à la fin de l'émission du phonème (barre verte).

Dans l'onglet « Annotation » de SPPAS, nous avons lancé les fonctions « Text Normalization », « Phonetization » et « Alignment » simultanément. Cette étape permet de segmenter l'audio en phonèmes à partir du texte et d'aligner temporellement les phonèmes sur le signal acoustique. De cette manière, les allongements vocaliques et les pauses remplies sont facilement repérables par leur durée égale ou supérieure à 200 ms.

Nous avons effectué une nouvelle vérification dans Praat en ouvrant les fichiers `.TextGrid` nommés *-palign* pour nous assurer que les phonèmes étaient correctement alignés avec le signal acoustique.

Dans SPPAS, onglet « Analyze », nous avons ouvert les fichiers `.TextGrid` intitulés *-palign*, et nous avons sélectionné la ligne d'annotation (ou *tier*) « TokenAlign ». Nous avons appliqué un filtre (*single filter*) pour créer un nouveau *tier* nommé « pauses ». Ce *tier* contient les phonèmes vocaliques et les pauses `(U~/, A/, O/, @, i, E, e, j, H, a~, O~, y, 2, 9, u, fp, #` selon les conventions de SPPAS) d'une durée égale ou supérieure à 200 ms. Dans Praat, nous avons uniquement conservé les annotations des allongements vocaliques dont la courbe intonative est plate.

Enfin, nous avons harmonisé l'ensemble des annotations en remplaçant les annotations déjà existantes du *tier* « pauses » par les étiquettes suivantes : **SP** pour les pauses silencieuses (*Silent Pauses*), **FP** pour les pauses remplies (*Filled Pauses*) et **SV** pour les allongements vocaliques (*Sustained Vowels*). Pour cela, nous avons lancé le script de remplacement de texte développé par DiCanio (2011) dans Praat.

Accord inter-annotateur

L'accord inter-annotateur consiste à donner un certain pourcentage des données (environ 20%) ainsi que son protocole d'annotation (un descriptif des éléments à annoter) à une ou plusieurs personnes n'ayant pas de connaissances du corpus afin de vérifier la validité et la fiabilité du processus d'annotation. Il se calcule à l'aide du coefficient κ de Cohen, dont le score s'échelonne de 0 à 1. Un score égal ou supérieur à .81 atteste la validité et la fiabilité des annotations. En clinique, cette étape permet d'assurer une annotation comportant le moins de biais de subjectivité possible. En effet, la connaissance du patient et de ses troubles peut conduire à surestimer ou sous-estimer le nombre de pauses. Il est vrai cependant que l'accord inter-annotateur peut être difficile à réaliser en clinique (par manque de disponibilité de ses collègues p. ex.), aussi cette étape peut être optionnelle.

Nous avons confié 20% de nos données à deux annotatrices cliniciennes non-orthophonistes, naïves de l'analyse de la parole, pour comparer nos annotations des pauses. Nous leur avons fourni en amont un document détaillé de notre système d'annotation. Notre accord inter-annotateur a atteint 98,71% d'accord pour l'annotation des pauses silencieuses, soit un agrément presque parfait (κ = .97), 90,98% pour l'annotation des pauses remplies, soit un agrément presque

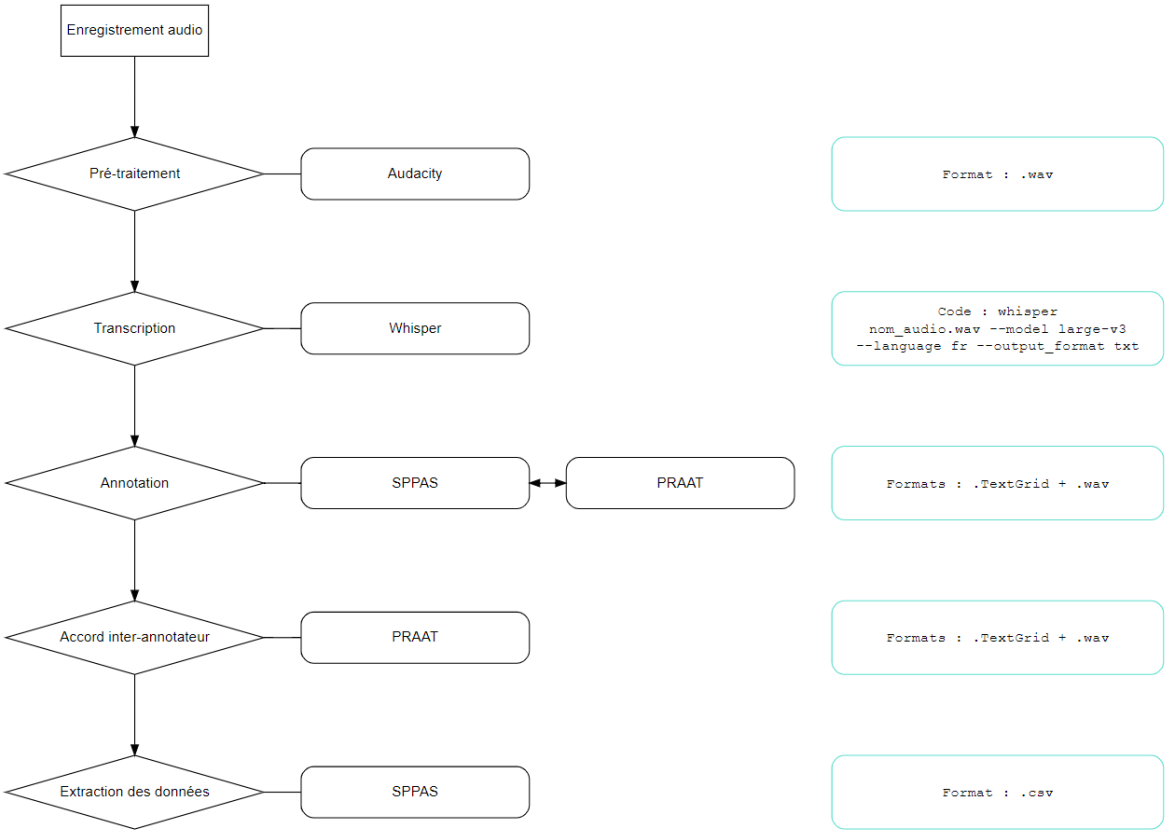


FIGURE 2 : Diagramme des étapes et des logiciels utilisés pour l'annotation de la parole.

parfait (κ = .81), et 86,88% pour l'annotation des allongements vocaliques, soit un agrément substantiel (κ = .74).

La figure 2 synthétise les différentes étapes que nous avons réalisées pour annoter la parole des participantes.

Extraction des données prosodiques

Nous avons extrait le nombre, la durée totale, moyenne, médiane et l'écart-type des trois types de pauses pour chaque échantillon audio à partir de SPPAS. Dans l'onglet « Analyze », nous avons sélectionné toutes les *tiers* « pauses », puis ouvert l'option « Statistics » pour accéder aux données sous forme de tableur. Cette étape a duré deux minutes.

Évaluation psychologique et cognitive

L'ensemble des participantes a bénéficié d'une évaluation psychologique et cognitive afin de contrôler les variables confondantes susceptibles d'influencer la production de pauses.

Elles ont également complété la version française du questionnaire FACT-Cog v3 (Costa et al., 2018 ; Joly et al., 2012) pour évaluer la présence d'une plainte cognitive significative (score inférieur ou égal à 55/72 à la sous-échelle *Perceived Cognitive Impairment*, PCI).

Les femmes ayant pris part à l'étude ont complété l'échelle *Hospital Anxiety Depression* (HAD, Zigmond & Snaith, 1983). Ce questionnaire comprend deux sous-échelles destinées à évaluer l'anxiété (HAD-A) et la dépression (HAD-D). Un score supérieur à 11/21 à l'une de ces sous-échelles atteste de la présence d'anxiété ou de dépression.

Elles ont ensuite complété le *Functional Assessment of Chronic Illness Therapy – Fatigue* (FACIT-F, Bonomi et al., 1996 ; Cella et al., 1993). Ce questionnaire est spécialement conçu pour évaluer la fatigue globale liée au cancer. Plus le score est élevé, moins la fatigue est importante. Il est à noter qu'il n'existe pas de seuil standardisé pour cette batterie de questions.

Enfin les participantes ont passé le test de repérage cognitif MoCA (Nasreddine et al., 2005). Ce test rapide à administrer est conçu pour détecter des troubles de la cognition en évaluant plusieurs

	id	site	group	age	complaint	anxiety	depression	cognition	FP_mean	SP_mean	SV_mean	FP_median
1	AL04	hospital	BC	41	38	0	2	26	0.464137214	1.259361109	0.464098824	0.25
2	AL08	hospital	BC	38	39	4	3	26	0.505	0.767647059	0.429166667	0.505
3	AL10	hospital	BC	38	31	11	4	30	0.568072969	0.756917516	0.318287043	0.48893414
4	AL13	hospital	BC	45	29	9	13	21	0.639166667	1.425927875	0.401111111	0.64
5	AL14	hospital	BC	39	49	6	11	29	0.626744848	0.799075957	0.393893653	0.5225

FIGURE 3 : Rangement des données de cinq participantes dans JASP.

domaines cognitifs (mémoire, capacités visuo-spatiales, langage, etc.). Un score supérieur ou égal à 26/30 est considéré comme normal.

Analyses statistiques

Les analyses statistiques de cette étude ont été réalisées avec le logiciel JASP (JASP Team, 2023). Ce logiciel présente l'avantage d'être gratuit, hors ligne, *open source* et simple d'utilisation. Son interface graphique rend accessibles plusieurs modèles de statistiques fréquentistes paramétriques et non paramétriques, ainsi que leurs équivalents Bayésiens (cf. https://jasp-stats.org/how-to-use-jasp/ pour une introduction à JASP). Nos données sont rangées dans un tableur au format .csv. Chaque colonne correspond à une variable et chaque ligne à une participante. La figure 3 montre le rangement des données après importation dans JASP.

L'effectif de notre étude est petit, raison pour laquelle nous avons choisi des tests non-paramétriques pour nos analyses statistiques. Nous avons effectué des tests U de Mann-Whitney (Mann & Whitney, 1947) pour comparer (1) l'âge, les scores au test de repérage cognitif et aux questionnaires des participantes BC par rapport aux participantes HC, ainsi que (2) les variables prosodiques des deux groupes. Enfin, nous avons effectué une corrélation de Spearman (Spearman, 1904) pour vérifier la possibilité d'un effet des variables psychologiques (anxiété, dépression, fatigue) sur les variables prosodiques.

L'ensemble des graphiques a été réalisé sur R studio (RStudio, BPC, 2020) avec les packages ggplot2 (Wickham et al., 2016) et GridExtra (Kassambara, 2016).

RÉSULTATS

Caractéristiques du corpus

Notre corpus est composé de 26 échantillons de parole pour une durée totale de 35 minutes (moyenne par échantillon = 82 s ± 41s). Au total, nous avons annoté 1222 pauses d'hésitation.

Données démographiques

Le tableau 1 montre les résultats du test U de Mann-Whitney réalisé pour la comparaison des variables démographiques et psychologiques entre les groupes. Comme attendu l'âge (p = .456) et le score MoCA ne révèlent pas de différence significative (p = .407), contrairement au score PCI du FACT-Cog (p < .001). Les participantes BC ont des scores significativement plus élevés que les contrôles aux questionnaires d'anxiété (p = .031) et de dépression (p = .009). Ces scores restent toutefois en-deçà des seuils pathologiques. Les participantes BC ont un score significativement

TABLEAU 1 : Comparaison des variables démographiques et des scores au test cognitif et aux questionnaires.

Variable	BC *moyenne (ET)*	HC *moyenne (ET)*	*p*-value
Âge	46 (9.9)	44 (10.6)	0.456
MoCA (/30)	27 (2.6)	27 (2.3)	0.407
FACT-Cog PCI (/72)	31 (10.7)	67 (4.8)	**< .001***
HAD-A (/24)	10 (5.0)	6 (4.2)	**0.031***
HAD-D (/24	7 (4.5)	3 (3.3)	**0.009***
FACIT-F (/52)	25 (8.6)	44 (7.1)	**< .001***

* = différence significative.

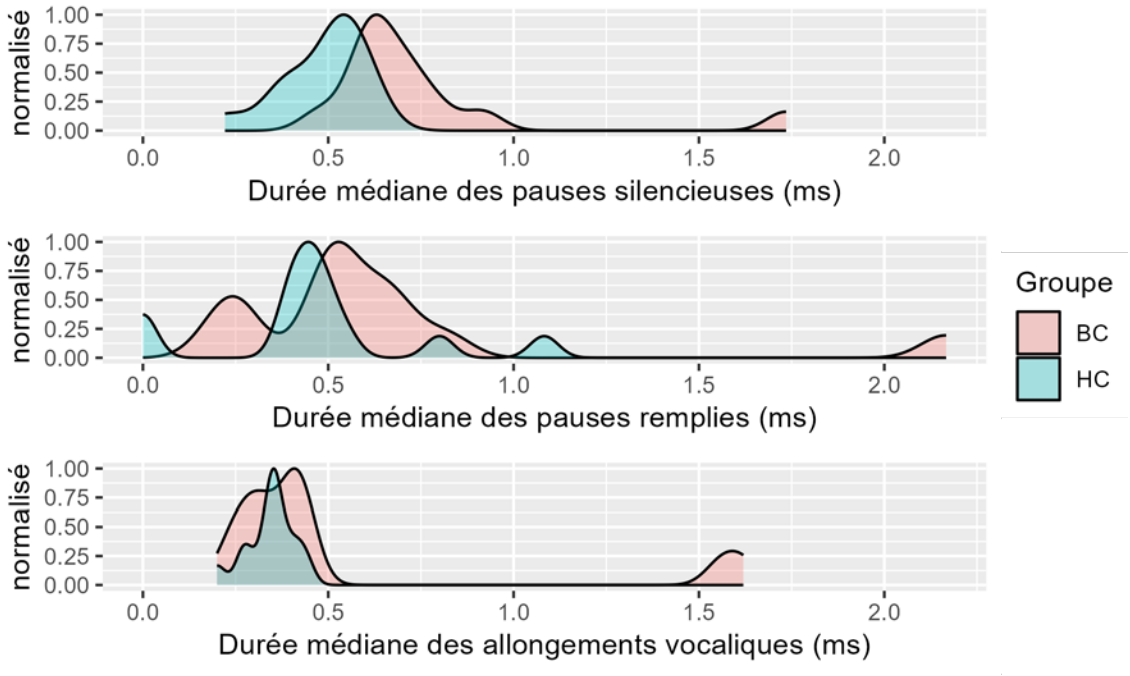


FIGURE 4 : Distribution de la durée médiane des pauses.

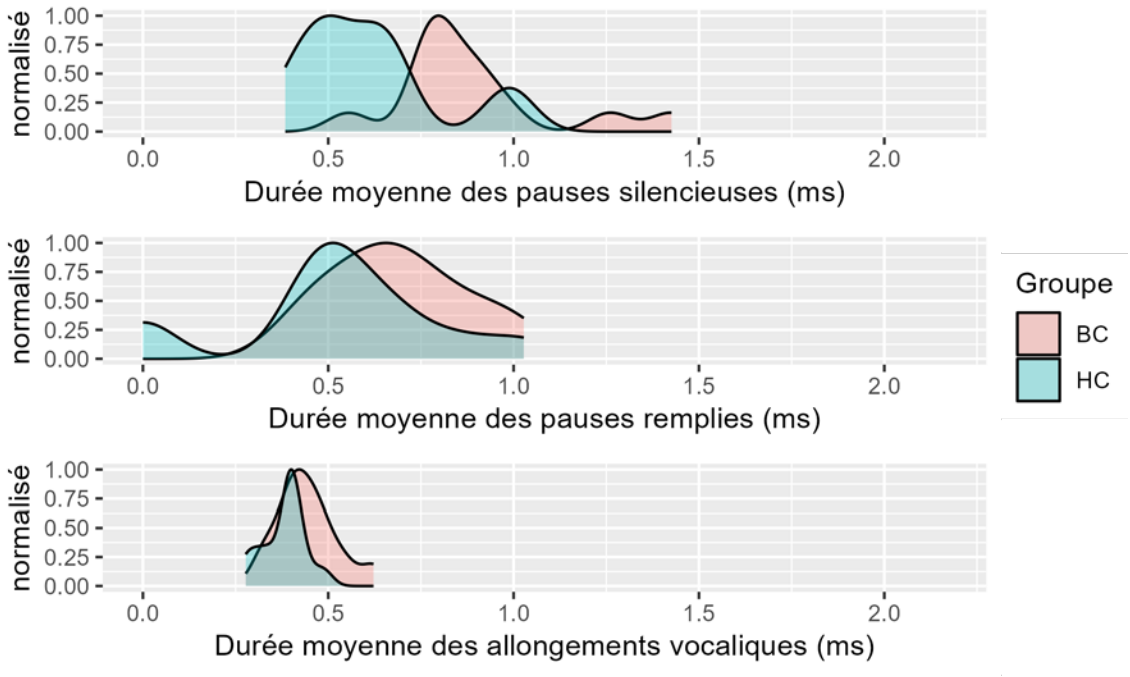


FIGURE 5 : Distribution de la durée moyenne des pauses.

plus faible au questionnaire de fatigue ($p < .001$), ce qui suggère la possible présence de fatigue chez cette population.

Résultats de l'analyse des pauses

La distribution des pauses est asymétrique avec une densité plus élevée chez les deux groupes pour les pauses de courte durée (~500 ms). La figure 4 montre la distribution de la durée médiane des pauses pour chaque type de pause.

La figure 5 montre la distribution de la durée moyenne des pauses pour chaque type de pause.

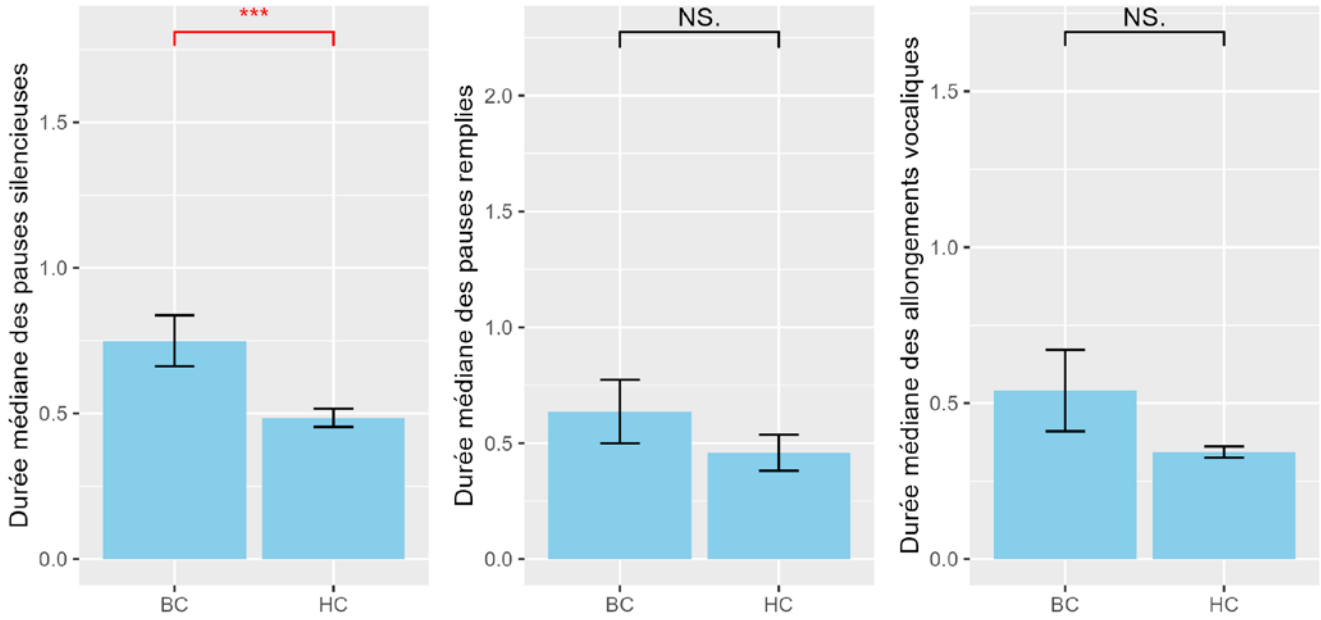


FIGURE 6 : Diagramme en barres de la durée médiane des pauses silencieuses, pauses remplies et allongements vocaliques pour les groupes Breast Cancer (BC) et Healthy Controls (HC). Les barres indiquent l'erreur-type. Les étoiles indiquent une différence significative entre les deux groupes.

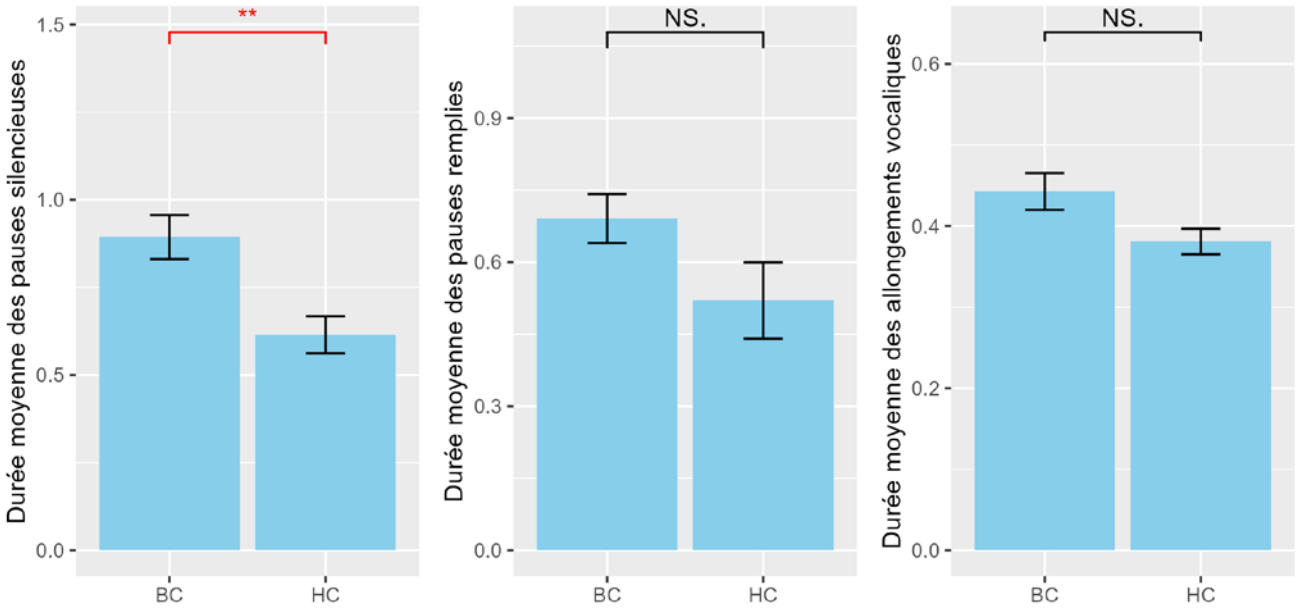


FIGURE 7 : Diagramme en barres de la durée moyenne des pauses silencieuses, pauses remplies et allongements vocaliques pour les groupes Breast Cancer (BC) et Healthy Controls (HC). Les barres indiquent l'erreur-type. Les étoiles indiquent une différence significative entre les deux groupes.

Durées médiane et moyenne des variables prosodiques

Nous avons réalisé un test U de Mann-Whitney pour comparer les durées médiane et moyenne des pauses silencieuses, des pauses remplies et des allongements vocaliques des participantes BC à celles des participantes HC. Les résultats indiquent des durées médiane et moyenne des pauses silencieuses significativement plus longues pour les participantes BC que pour les participantes HC (respectivement p = <.001 et p = .002). En revanche, les résultats ne montrent

pas de différence significative pour les durées médiane et moyenne des pauses remplies et des allongements vocaliques.

La figure 6 et la figure 7 montrent respectivement les durées médiane et moyenne des pauses silencieuses, pauses remplies et allongements vocaliques pour les groupes BC et HC.

Taux de pauses et d'allongements vocaliques

Les taux de pauses et d'allongements vocaliques par échantillon ont été calculés en divisant le nombre d'occurrences par la durée totale des échantillons.

Le tableau 2 résume les moyennes des taux de chaque variable prosodique. Nous avons réalisé un test U de Mann-Whitney pour comparer les taux des pauses silencieuses, des pauses remplies et des allongements vocaliques des participantes BC à ceux des participantes HC. Les résultats ne montrent pas de différence significative entre les groupes pour ces variables prosodiques.

Contrôle des co-facteurs

Certains facteurs psychologiques ou états comme la fatigue peuvent influencer la production d'un discours. Nous avons réalisé une corrélation de Spearman pour examiner la relation entre les scores d'anxiété, de dépression et de fatigue et la durée des pauses silencieuses.

Les résultats ne montrent pas de corrélation significative entre les durées médiane et moyenne des pauses silencieuses et le score d'anxiété ($r(24) = .317$, $p = .114$, et $r(24) = .264$, $p = .192$), ni entre la durée médiane des pauses silencieuses et le score de dépression ($r(24) = .346$, $p = .084$). En revanche, les résultats suggèrent une corrélation significative entre les durées médiane et moyenne des pauses silencieuses et le score de fatigue ($r(24) = -.446$, $p = .022$ et $r(24) = - 473$, $p = .015$), ainsi que la durée moyenne des pauses silencieuses et le score de dépression ($r(24) = .458$, $p = .019$). Toutefois, aucun score n'ayant été considéré pathologique, ces résultats doivent être interprétés avec précaution.

DISCUSSION ET CONCLUSION

Ce travail décrit les différentes étapes nécessaires à la réalisation d'une analyse de la prosodie à travers une application concrète chez des participantes ayant eu un cancer du sein. La production de parole de treize participantes post-cancer du sein ayant un trouble cognitif induit par un cancer et ses traitements et celle de treize participantes contrôles saines appariées a été obtenue au moyen d'une tâche de narration élicitée par des images. Le corpus final, composé de vingt-six échantillons de parole pour une durée totale de trente-cinq minutes, a ensuite été analysé grâce à une méthode spécifiquement développée pour les besoins de l'étude.

La méthode employée utilise des outils informatiques automatisés, hors ligne et *open source*. Outre le gain de temps procuré par ces outils, leur utilisation protège le contrôle des données et de leur traitement. L'alternance entre SPPAS et Praat constitue toutefois une limite de cette méthode. En effet, elle nécessite la maîtrise de deux logiciels et peut allonger le temps d'analyse. Il est possible de réaliser l'ensemble de l'analyse prosodique sur SPPAS. Cependant, la visualisation du spectrogramme fourni par Praat rend l'ajustement des frontières plus aisé. De futures études veilleront à développer un outil combinant les fonctionnalités de SPPAS et la précision de Praat afin d'alléger davantage la procédure.

Cette méthode peut s'adapter à d'autres pathologies ou syndromes. Son application clinique peut être multiple. Elle permet par exemple d'établir une ligne de base et d'évaluer l'évolution des patients, de confirmer ou d'infirmer une intuition clinique en comparant la production d'un patient à des données normatives issues de la littérature, et/ou de proposer une analyse détaillée

TABLEAU 2 : Taux des pauses silencieuses, des pauses remplies et des allongements vocaliques.

Taux des variables prosodiques	BC *moyenne (ET)*	HC *moyenne (ET)*	*p*-value
Pauses silencieuses	0.270 (0.056)	0.224 (0.072)	0.101 ns
Pauses remplies	0.149 (0.062)	0.130 (0.072)	0.573 ns
Allongements vocaliques	0.172 (0.073)	0.161 (0.079)	0.880 ns

* = différence significative, ns = différence non-significative.

des difficultés linguistiques (recherche lexicale, programmation syntaxique, erreurs phonétiques, etc.).

Concernant l'application clinique que nous avons réalisée pour illustrer cette méthode, nos principaux résultats suggèrent que la durée des pauses d'hésitation silencieuses permet de différencier le groupe de participantes ayant eu un cancer du sein avec plainte cognitive des participantes contrôles, contrairement aux pauses remplies et aux allongements vocaliques. Ce résultat est similaire à ceux obtenus par Bóna (2024) qui a comparé les pauses dans le discours de personnes ayant des troubles cognitifs liés à une sclérose en plaques (SEP) à celles de participants sains lors de tâches de narration. Les pauses silencieuses des patients avec SEP sont plus fréquentes et plus longues que chez les contrôles. En revanche, les pauses remplies ne présentent pas de différence entre les deux groupes. Il est intéressant de noter que, bien que les pathologies sous-jacentes des deux populations (cancer et sclérose en plaques) diffèrent, les pauses silencieuses permettent d'identifier un TNCm chez des populations jeunes. Une explication possible serait que les pauses silencieuses, plus que les pauses remplies, soient le reflet d'une charge cognitive élevée.

Angelopoulou et al. (2024) soulignent la nécessité d'examiner séparément les types de pauses. En effet, chaque type de pauses répond à un besoin différent tel que des besoins linguistiques, interactionnels ou cognitifs. Par conséquent, elles peuvent refléter divers processus cognitifs. Il est donc nécessaire d'analyser isolément les pauses jouant un rôle de marqueur linguistique des TNCm.

Une des limites de notre étude est le petit effectif de participantes. Nos résultats sont donc à confirmer par une cohorte plus importante. Afin d'enrichir les connaissances sur les marqueurs de troubles cognitifs subtils en contexte de TNCm, de futures études pourront inclure d'autres paramètres prosodiques tels que la variabilité de F0 ou le débit de parole.

Pour conclure, notre étude montre l'intérêt des outils informatisés semi-automatiques pour l'évaluation de patients ayant un TNCm grâce au gain de temps et d'objectivité qu'ils procurent. Si l'utilisation de ces outils peut sembler à première vue fastidieuse, notre méthode détaillée assure la faisabilité technique d'une telle analyse prosodique en clinique.

Enfin, notre étude participe de l'évolution de la profession d'orthophoniste grâce à l'utilisation d'outils d'analyse automatisés ou d'intelligence artificielle qu'elle propose.

REMERCIEMENTS

Les auteurs remercient Farha Gheurbi et Marion Buchy d'avoir réalisé l'accord inter-annotateur de cette étude.

DÉCLARATION D'INTÉRETS

Les auteurs ont déclaré n'avoir aucun lien d'intérêt en relation avec cet article.

FINANCEMENTS

Amélie B. Richard est financée par une bourse doctorale attribuée par l'Université Montpellier Paul-Valéry et allouée par le ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI).

RÉFÉRENCES

Amani, O., Mazaheri, M. A., Moghani, M. M., Zarani, F., & Choolabi, R. H. (2024). Chemotherapy-induced cognitive impairment in breast cancer survivors: A systematic review of studies from 2000 to 2021. *Cancer Reports, 7*(2), e1989. https://doi.org/10.1002/cnr2.1989

American Psychiatric Association. (2022). *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5-TR). https://doi.org/10.1176/appi.books.9780890425787

Angelopoulou, G., Kasselimis, D., Goutsos, D., & Potagas, C. (2024). A methodological approach to quantifying silent pauses, speech rate, and articulation rate across distinct narrative tasks: Introducing the Connected Speech Analysis Protocol (CSAP). *Brain Sciences, 14*(5), 466. https://doi.org/10.3390/brainsci14050466

Audacity Team (1999). *Audacity* (version 3.0) [logiciel informatique]. https://www.audacityteam.org/

Bézy, C., Renard, A., & Pariente, J. (dir.) (2016). *GRÉMOTS : batterie d'évaluation des troubles du langage dans les maladies neurodégénératives du GRECO*. De Boeck Supérieur.

Bigi, B. (2015). SPPAS - Multi-lingual approaches to the automatic annotation of speech. T*he Phonetician. Journal of the International Society of Phonetic Sciences*, 111–112, 54–69. https://hal.science/hal-01417876

Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* (Version 6.1.55) [logiciel informatique]. https://www.fon.hum.uva.nl/praat/

Bóna, J. (2024). Pausing and fluency in speech of patients with relapsing-remitting multiple sclerosis. *Clinical Linguistics & Phonetics, 38*(4), 332–344. https://doi.org/10.1080/02699206.2023.2223347

Bonomi, A. E., Cella, D. F., Hahn, E. A., Bjordal, K., Sperner-Unterweger, B., Gangeri, L., Bergman, B., Willems-Groot, J., Hanquet, P., & Zittoun, R. (1996). Multilingual translation of the Functional Assessment of Cancer Therapy (FACT) quality of life measurement system. *Quality of Life Research, 5*(3), 309–320. https://doi.org/10.1007/BF00433915

Bryant, L., Spencer, E., & Ferguson, A. (2017). Clinical use of linguistic discourse analysis for the assessment of language in aphasia. *Aphasiology, 31*(10), 1105–1126. https://doi.org/10.1080/02687038.2016.1239013

Campione, E., & Véronis, J. (2002). *A large-scale multilingual study of silent pause duration.* Speech Prosody 2002, 11-13 avril. Aix-En-Provence. https://doi.org/10.21437/SpeechProsody.2002-35

Candea, M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané. Etude sur un corpus de récits en classe de français.* [thèse de doctorat, Université de la Sorbonne nouvelle - Paris III]. HAL. https://theses.hal.science/tel-00290143v1

Cella, D. F., Tulsky, D. S., Gray, G., Sarafian, B., Linn, E., Bonomi, A., Silberman, M., Yellen, S. B., Winicour, P., & Brannon, J. (1993). The Functional Assessment of Cancer Therapy scale: Development and validation of the general measure. *Journal of Clinical Oncology, 11*(3), 570–579. https://doi.org/10.1200/JCO.1993.11.3.570

Commission Nationale de l'Informatique et des Libertés - CNIL. (2016). *Le règlement général sur la protection des données—RGPD*. https://www.cnil.fr/fr/reglement-europeen-protection-donnees

Costa, D. S. J., Loh, V., Birney, D. P., Dhillon, H. M., Fardell, J. E., Gessler, D., & Vardy, J. L. (2018). The Structure of the FACT-Cog v3 in cancer patients, students, and older adults. *Journal of Pain and Symptom Management, 55*(4), 1173–1178. https://doi.org/10.1016/j.jpainsymman.2017.12.486

DiCanio, C. (2011). *Text replacement script for Praat* [logiciel informatique]. https://www.acsu.buffalo.edu/~cdicanio/scripts.html

Duez, D. (2001). Caractéristiques acoustiques et phonétiques des pauses remplies dans la conversation en français. *Travaux Interdisciplinaires du Laboratoire Parole et Langage (TIPA)*, 20, 31-48. https://hal.science/hal-00285404v1

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*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Goldman, J.-P., Thomas, F., Roekhaut, S., & Simon, A.-C. (2010). Étude statistique de la durée pausale dans différents styles de parole. *Actes des 28e Journées d'Étude sur la Parole (JEP)*, 161–164. http:// hdl.handle.net/2078.1/81909

Groupe de travail Traumatisme Crânio-Cérébral Léger (TCCL) en Île-de-France (2021). *Rapport du groupe de travail TCCL*. Agence Régionale de Santé Île-de-France. https://www.iledefrance.ars.sante.fr/traumatisme-cranio-cerebral-leger-tccl-en-ile-de-france-rapport-du-groupe-de-travail

Haute Autorité de Santé. (2018). *Patients présentant un trouble neurocognitif associé à la maladie d'Alzheimer ou à une maladie apparentée. Guide parcours de soins.* https://www.has-sante.fr/jcms/c_2906096/en/patients-presentant-un-trouble-neurocognitif-associe-a-la-maladie-d-alzheimer-ou-a-une-maladie-apparentee

He, R., Chapin, K., Al-Tamimi, J., Bel, N., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J. P., Alegret, M., Sanabria, A., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (2023). Automated classification of cognitive decline and probable Alzheimer's dementia across multiple speech and language domains. *American Journal of Speech-Language Pathology, 32*(5), 2075–2086. https://doi.org/10.1044/2023_AJSLP-22-00403

Hommel, M., Miguel, S. T., Naegele, B., Gonnet, N., & Jaillard, A. (2009). Cognitive determinants of social functioning after a first ever mild to moderate stroke at vocational age. *Journal of Neurology, Neurosurgery & Psychiatry, 80*(8), 876–880. https://doi.org/10.1136/jnnp.2008.169672

Ivanova, O., Martínez-Nicolás, I., & Meilán, J. J. G. (2023). Speech changes in old age: Methodological considerations for speech-based discrimination of healthy ageing and Alzheimer's disease. *International Journal of Language & Communication Disorders, 59*(1), 13–37. https://doi.org/10.1111/1460-6984.12888

Janelsins, M. C., Kohli, S., Mohile, S. G., Usuki, K., Ahles, T. A., & Morrow, G. R. (2011). An update on cancer- and chemotherapy-related cognitive dysfunction: Current status. *Seminars in Oncology, 38*(3), 431–438. https://doi.org/doi:10.1053/j.seminoncol.2011.03.014

JASP Team. (2023). *JASP* (version 0.17.3) [logiciel informatique]. https://jasp-stats.org/team/

Jim, H. S. L., Phillips, K. M., Chait, S., Faul, L. A., Popa, M. A., Lee, Y.-H., Hussin, M. G., Jacobsen, P. B., & Small, B. J. (2012). Meta-analysis of cognitive functioning in breast cancer survivors previously treated with standard-dose chemotherapy. *Journal of Clinical Oncology, 30*(29), 3578–3587. https://doi.org/10.1200/JCO.2011.39.5640

Joly, F., Lange, M., Rigal, O., Correia, H., Giffard, B., Beaumont, J. L., Clisant, S., & Wagner, L. (2012). French version of the Functional Assessment of Cancer Therapy–Cognitive function (FACT-Cog) version 3. *Supportive Care in Cancer, 20*, 3297–3305. https://doi.org/10.1007/s00520-012-1439-2

Kassambara, A. (2016). *ggpubr: "ggplot2" Based publication ready plots* (version 0.6.0) [fichier de données]. https://doi.org/10.32614/CRAN.package.ggpubr

Lanzi, A. M., Saylor, A. K., Fromm, D., Liu, H., MacWhinney, B., & Cohen, M. L. (2023). DementiaBank: Theoretical rationale, protocol, and illustrative analyses. American *Journal of Speech-Language Pathology, 32*(2), 426–438. https://doi.org/10.1044/2022_AJSLP-22-00281

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics, 18*(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

OSF (2011). Plateforme open source de soutien à la recherche. Center for Open Science. https://osf.io/sqwnz/?view_only=8947f20e707f476c9cf44a35ca9d7c32

Parjane, N., Cho, S., Ash, S., Cousins, K. A. Q., Shellikeri, S., Liberman, M., Shaw, L. M., Irwin, D. J., Grossman, M., & Nevler, N. (2021). Digital speech analysis in progressive supranuclear palsy and corticobasal syndromes. *Journal of Alzheimer's Disease, 82*(1), 33–45. Embase. https://doi.org/10.3233/jad-201132

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust speech recognition via large-scale weak supervision. *arXiv: 2212.04356[*eess.AS*]*. https://doi.org/10.48550/ARXIV.2212.04356

Richard, A. B., Reilly, K. T., & Jacquin-Courtois, S. (2022, 17 juin). Que révèlent les disfluences sur le manque du mot rapporté par les patientes ayant un cancer du sein. *Actes des 34e Journées d'Etudes sur la Parole (JEP)*, 357-365. https://doi.org/10.21437/JEP.2022-38

Richard, A. B., Hirsch, F., Jacquin-Courtois, S., & Reilly, K. T. (2023). Self-interruptions in breast cancer patients who complain of anomia. Dans R. Skarnitzl & J. Volín (dir.), *Proceedings of the 20th International Congress of Phonetic Sciences* (p. 4012–4016). https://hal.science/hal-04195993v1

Richard, A. B., Foncelle, A., Hirsch, F., Jacquin-Courtois, S., Reilly, K., & Lelandais, M. (2024). *Speech markers of Cancer-Related Cognitive Impairment: A pilot study.* Speech Prosody 2024, 2-5 juillet. Leiden. https://doi.org/10.31219/osf.io/xz46h

RStudio, BPC. (2020). *RStudio: Integrated development for R* [logiciel informatique]. https://support.posit.co/hc/en-us/articles/206212048-Citing-RStudio

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology, 15*(1), 72-101. https://doi.org/10.2307/1412159

Themistocleous, C., Eckerström, M., & Kokkinakis, D. (2020). Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from healthy controls. *PLOS ONE, 15*(7), e0236009. https://doi.org/10.1371/journal.pone.0236009

Van Dyk, K., Crespi, C. M., Petersen, L., & Ganz, P. A. (2020). Identifying Cancer-Related Cognitive Impairment using the FACT-Cog perceived cognitive impairment. *JNCI Cancer Spectrum, 4*(1), pkz099. https://doi.org/10.1093/jncics/pkz099

Vincze, V., Szatlóczki, G., Tóth, L., Gosztolya, G., Pákáski, M., Hoffmann, I., & Kálmán, J. (2021). Telltale silence: Temporal speech parameters discriminate between prodromal dementia and mild Alzheimer's disease. *Clinical Linguistics & Phonetics, 35*(8), 727–742. https://doi.org/10.1080/02699206.2020.1827043

Wagner, L. I., Sweet, J., Butt, Z., Lai, J. S., & Cella, D. (2009). Measuring patient self-reported cognitive function: Development of the Functional Assessment of Cancer Therapy–cognitive function instrument. *The Journal of Supportive Oncology, 7*(6), W32–W39.

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & van den Brand, T. (2016). *ggplot2: Create elegant data visusalisations using the grammar of graphics*. https://ggplot2.tidyverse.org

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression scale. *Acta Psychiatrica Scandinavica, 67*(6), 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Speech markers of Cancer-Related Cognitive Impairment : A pilot study

---

Speech Prosody 2024
2-5 July 2024, Leiden, The Netherlands

Speech markers of Cancer-Related Cognitive Impairment: A pilot study

Amélie B. Richard[1,2], Alexandre Foncelle[2], Fabrice Hirsch[1], Sophie Jacquin-Courtois[2,3], Karen T. Reilly[2], Manon Lelandais[4]

[1]Université Paul-Valéry Montpellier 3, UMR5267 Praxiling, France
[2]Université de Lyon, U1028; UMR5292 Lyon Neuroscience Research Centre, France
[3]Henry-Gabrielle Hospital, Hospices Civils de Lyon, France
[4]Université Paris Cité, UR3967 CLILLAC-ARP, France
amelie.richard@inserm.fr, manon.lelandais@u-paris.fr

Abstract

Speech is sensitive to mild cognitive changes due to age-related diseases, and prosodic features can identify patients with early-stage dementia from controls. Few studies have investigated speech markers of subtle cognitive impairment in non-neurodegenerative pathologies in younger populations, such as Cancer-Related Cognitive Impairment (CRCI). Little is known about the cognitive mechanisms underlying CRCI, but it is frequently encountered by cancer patients who mainly report memory-related concerns (i.e., forgetting words). Despite its substantial impact on patient quality of life, CRCI is difficult to detect with neuropsychological tools and often remains underdiagnosed. Our aim is to test whether previously documented speech markers are likely to detect CRCI in patients with breast cancer. We compared speech rate, F0 variability and pause duration in 11 breast cancer survivors with a cognitive complaint, 11 breast cancer survivors without any cognitive complaint and 10 controls in two narrative tasks (memory-based; picture-based). A Bayesian analysis showed no significant effects of group or task, but a qualitative analysis of pauses allowed us to generate hypotheses about the cognitive mechanisms underlying the patients' reported memory concerns. Even though speech markers specific to CRCI have yet to be defined, prosodic analysis is a promising approach for detecting subtle cognitive impairment.

Index Terms: Speech markers, cognitive impairment, pauses, breast cancer.

1. Introduction

Prosody is known to be sensitive to changes in cognition. Language is a complex cognitive function which relies on phonological and semantical mechanisms, as well as memory, executive and attentional processes. Speech production, as its output, reflects everyday cognitive ability.

Some prosodic features reach satisfactory discrimination power distinguishing patients with cognitive decline from controls, and identifying variants of a same pathology with either manual or machine learning methods [1].Speech rate, F0 variability and pause duration in particular have been shown to be good candidates for assessing cognition. For instance, individuals with amnestic Mild Cognitive Impairment have more F0 variability than individuals with non-amnestic Mild Cognitive Impairment [2]. Research in psycholinguistics suggests that these speech markers might be associated with different brain processes. For instance, a lower speech rate might reveal deficits in episodic memory, executive functions, and lexical retrieval difficulties in individuals with early-stage dementia [3] while a longer pause duration might reflect lexical-semantic difficulties [4]. These speech markers are even more manifest when the cognitive load is increased with a demanding task [5].

However, the large majority of studies using prosodic analysis to detect cognitive impairment focus on geriatric populations with dementia. Given that dementia is caused by clear degenerative physiological alterations of brain structures, its speech characteristics may differ from those of cognitive change not related to dementia. Little is known about speech markers in younger individuals with acquired cognitive impairment in non-neurodegenerative contexts. A new challenge in clinical prosody analysis is to test the efficiency of speech markers in differentiating between individuals with subtle cognitive impairment and healthy controls.

Subtle non-neurodegenerative cognitive impairment is increasingly frequent, and covers a wide range of pathologies (e.g., mild traumatic brain injury, long-Covid-19, minor stroke sequalae) in a population that is often young. It refers to self-reported cognitive difficulties that are hard to assess using current diagnostic tools. One of these impairments is Cancer-Related Cognitive Impairment (CRCI). CRCI is a functional change in cognition encountered by up to 70% of patients with cancer in the years following the end of curative treatments. In addition to decreasing patients' quality of life, CRCI hinders their return to work and limits their participation in social activities.

Cancer patients mostly report difficulties with memory and language (i.e., forgetting words and names, forgetting what they wanted to say) [6], [7]. Subjective difficulties in CRCI are confirmed by validated self-report questionnaires which signal the presence of a general cognitive complaint. Despite this, most studies note no or little association between the intensity of cognitive complaints and scores on neuropsychological tests [8]. This discrepancy might be explained by the fact that such tests were designed for more severe disorders and evaluate a single cognitive function in optimal conditions. The lack of appropriate objective tools hinders our understanding of CRCI, and prevents cancer patients from accessing appropriate care. There is thus a need for new approaches to explore CRCI and its underlying mechanisms.

Based on the results of previous studies using speech markers for detecting subtle cognitive impairment [9], we assume that prosodic analysis is sensitive and ecological enough to discriminate cancer patients with CRCI from controls.

10.21437/SpeechProsody.2024-16

1.1. Aim & hypotheses

Our aim is to test whether previously documented speech markers are likely to be specific to breast cancer patients with a cognitive complaint. We hypothesized that breast cancer survivors with a cognitive complaint would produce slower speech, with a longer pause duration and more F0 variability than those without a cognitive complaint and healthy controls. Based on the patients' memory complaints we think that they will perform worse in a memory-based narrative task than in a picture-based narrative task.

2. Method

The present study is part of two larger projects. The first one aims at assessing the effectiveness of a rehabilitation program offered to women with breast cancer who have finished the main phase of their treatment (ActiCog), and has been approved by the ethical committee of the Hospices Civils de Lyon (n°22-5009). The second project focuses on Cancer-Related Cognitive Impairment in women with breast cancer (DisCCo), and was approved by the Ethics Committee Est-III in August 2023.

2.1. Population

Thirty-seven survivors enrolled in a physical rehabilitation program, Alizés, had a 30-minute one-to-one interview with a speech therapist. The Alizés program is offered by the Physical Medicine and Rehabilitation department at the Henry Gabrielle hospital (Hospices Civils de Lyon), and is open to women who are less than one year after completion of an adjuvant treatment for breast cancer (surgery, chemotherapy and radiation therapy). Women taking hormone therapy were included.

The women interviewed for the current study were required to be enrolled in the Alizés program and to have a good level of French (i.e., native speaker or fluent). Exclusion criteria included: neurological (i.e., stroke) or psychiatric (i.e., schizophrenia) history, currently taking medication known to modify cognitive ability (e.g., antidepressants), and the presence of a speech or language disorder. Four survivors were not eligible because of epilepsy, depression, or stroke history, and one was not fluent in French. The FACT-Cog questionnaire is a validated measure of cognitive complaint in patients with cancer. A score below 55/72 on the Perceived Cognitive Impairment subscale signals the presence of a cognitive complaint [10]. Among the thirty-three eligible participants, we randomly chose eleven who had a significant complaint revealed by the FACT-Cog, and selected eleven others with scores above this cutoff. In total, twenty-two participants were included in the current study. In addition, we selected ten healthy control participants from the DisCCo study without any cancer history or cognitive complaint. Inclusion and exclusion criteria for controls were the same as described above.

To sum up, participants were divided into three groups: 11 cancer survivors with a cognitive complaint (BCcog), 11 cancer survivors without any cognitive complaint (BC), and 10 healthy controls (HC).

2.2. Procedure

2.2.1. Tasks and speech variables

Participants were asked to complete two narrative tasks: a **picture-based** task for which they were instructed to tell a story from a 5-picture sequence [11], and a **memory-based task** for which they were instructed to listen to and memorize a 1-minute short story [12], then to immediately recall it with a maximum of detail. Healthy controls were instructed to read and memorize a short text displayed on a screen for forty seconds, then to immediately recall it with a maximum of detail.

The samples for survivors were recorded with a Rode Lavalier Go in a quiet room at the hospital (Saint-Genis-Laval, France). The samples for healthy controls were recorded with a Shure lavalier in a soundproof cabin at the Lyon Neuroscience Research Centre (Bron, France). Three speech variables were selected according to the following characteristics:

- **Pause duration (ratio):** was measured by dividing total pause time by total speech time. Pauses included silent pauses (speechless segments with a duration above 250 ms), filled pauses (non-lexical fillers perceived as hesitation markers such as "euh" and "hum" (/ʌ/, /ø/ and /m/) with a flat F0 contour and a duration above 250 ms), and lengthenings (voiced phonemes with flat F0 contour and a duration above 250 ms). The literature on pause length agrees on a threshold equal to or above 200 ms for pauses related to cognitive processes such as hesitation or demarcation [13]. However, we recorded survivors' samples in a hospital resulting in medium audio quality. We thus chose a threshold of 250 ms to avoid annotating unrelated acoustic phenomena such as echoes or background noise.
- **Speech rate (wpm):** was measured by dividing the number of words per minute. This included lexical and grammatical items and word fragments, and excluded filled pauses and noises.
- **F0 variability (Hz):** was automatically measured with Praat [14] as the standard deviation of the total speech sample.

2.2.2. Subjective evaluation

In addition to the PCI subscale of the FACT-Cog questionnaire, participants completed the Perceived Cognitive Ability subscale. Despite having no standardized cutoff, the subscale is used in clinical contexts as a useful tool to measure self-reported cognitive ability. Participants also filled the standardized Hospital Anxiety and Depression (HAD) scale for anxiety and depression.

2.3. Data processing

We recorded 45 minutes of picture-based task and 31 minutes of memory-based task. All samples were automatically transcribed with Whisper [15] and manually checked by the first author. Pause annotation was performed with the SPPAS software [16] and errors were manually corrected by the first author in Praat. In total, we annotated 874 pauses for the picture task and 1355 pauses for the memory task.

2.4. Statistical analysis

2.4.1. Linear mixed-effects model

The null hypothesis (H0) postulates that there is no difference in F0 variability, pause ratio or speech rate, depending on groups and conditions (i.e. tasks). To test H0 against H1 (i.e., the presence of a significant difference), linear mixed models were designed and compared for each of these three dependent variables. A random intercept was attributed to the *subject* variable, as well as fixed effects to *group* (HC, BC and BCcog

levels), *condition* (picture and memory levels), *age*, *anxiety*, *depression*, and the various interaction terms were combined with all these fixed effects. Models were conducted in R [14] using the lme4 package [17].

2.4.2. Bayesian statistics

To confirm the results of the mixed models, we conducted a Bayesian analysis. There was no strong *a priori* on the parameter values of the models, so non-informative priors were set for all parameters in all models with a Normal prior distribution N(0,3). The Bayes factor was calculated between each model comparison for hypothesis testing and, for the most likely model, the summary of the parameter estimate with a 95% credibility interval was considered. The Bayesian analysis was conducted in R using the brms package [18].

3. Results

3.1. Population description

Table 1 shows the population's mean age and scores in questionnaires on cognitive complaint and anxiety/depression, along with standard deviations. Controls were slightly younger than cancer outpatients, although a Kruskal-Wallis test showed no significant difference between groups (p = .31). A lower score on the PCI subscale signals a higher cognitive complaint. As mentioned in the Method section, only BCcog had a significant cognitive complaint. A lower score on the PCA subscale signals lower perceived cognitive abilities. Cognitive abilities were significantly lower in the BCcog group. A score above 11/21 on the HAD subscales signals the presence of anxiety or depression. None of the participants reported anxiety or depression.

Table 1: *Population age and scores in questionnaires on cognitive complaint, and anxiety/depression; mean (sd)*

Characteristics	**BC**	**BCcog**	**HC**
Age	47.18 (7.15)	49.36 (9.11)	44.10 (9.96)
FACT-Cog PCI score /72	62.73 (5.10)	40.55 (9.32)	66.70 (5.12)
FACT-Cog PCA score /28	20.27 (6.02)	13.27 (4.55)	23.70 (2.87)
HAD – Anxiety score /21	6.36 (2.99)	5.11 (2.77)	8.20 (3.34)
HAD – Depression score /21	4.18 (3.59)	4.44 (3.74)	6.30 (3.38)

3.2. Prosodic variables

Table 2 shows median pause duration ratio, speech rate and F0 variability for each group in the memory and picture conditions. Pause duration ratio is higher for the BCcog group in the picture condition (0.42) than for BC (0.23) and controls (0.28).

Table 2: *Median pause duration ratio, speech rate and F0 variability for the memory and picture conditions*

	BC	**BCcog**	**HC**
Memory			
Pause duration ratio	0.30	0.31	0.23
Speech rate (wpm)	165	173	150
F0 variability (sd)	35.66	38.26	31.91
Picture			
Pause duration ratio	0.23	**0.42**	0.28
Speech rate (wpm)	180	152	162
F0 variability (sd)	34.81	37.72	33.45

Figure 1 shows the violin box distributions of pause duration (a), speech rate (b), and F0 variability (c) between groups for the memory (blue) and picture (orange) conditions. The white line represents the median. A trend can be observed in 1a with a higher pause duration ratio in BCcog than in controls, especially for the picture condition (see Table 2).

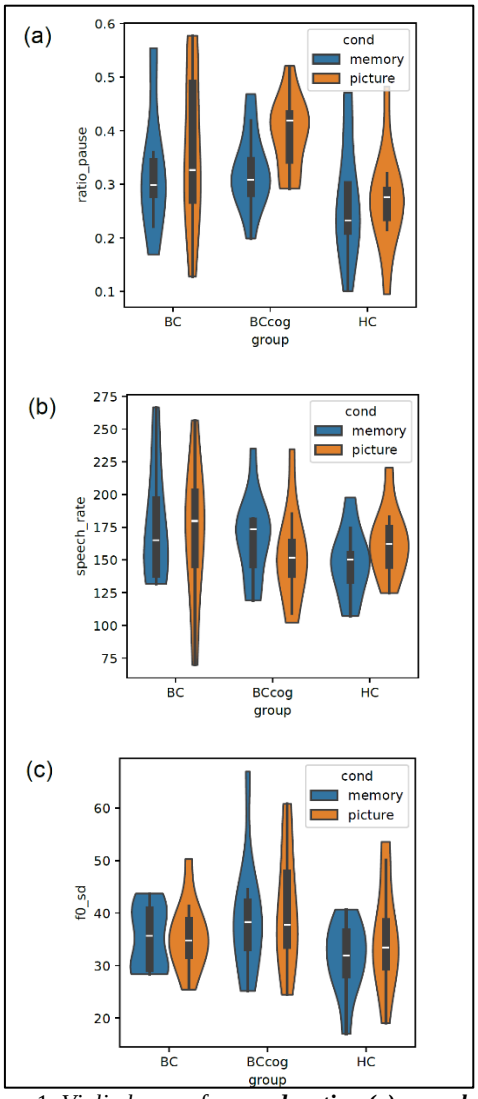


Figure 1: *Violin boxes of **pause duration (a), speech rate (b)**, and **F0 variability (c)** for each group in the memory and picture conditions.*

3.3. Linear and Bayesian analyses

The linear mixed models showed no effect of group, condition, age, anxiety or depression on the three speech markers. The Bayes factor indicates evidence for the null hypothesis (i.e., no difference between conditions and groups), $BF_{H0/H1\ pause\ ratio}$ = $1.63e^7$, $BF_{H0/H1\ speech\ rate}$ = 1.11, $BF_{H0/H1\ F0sd}$ = 3.41, meaning that the data are respectively approximately $1.63e^7$, 1.11, and 3.41 times more likely to occur under H0 than under H1. These results indicate very strong evidence in favor of the null

hypothesis for pause ratio, no evidence for speech rate, and moderate evidence for F0 variability. In other words, the Bayesian analysis suggests that there is no difference between condition and groups for pause ratio, that no conclusion can be made for speech rate, and that there might be no difference for F0.

3.4. Qualitative analysis of pause duration

In order to interpret the trend concerning pause duration ratio alongside the absence of quantitative results, we carried out a finer-grained analysis of pauses. We specifically examined pause duration between two functional groups of pauses, e.g. demarcation and hesitation.

In French, rhythmic groups are marked by final accents [19] with a clear rising or falling F0, coinciding with syntactic boundaries. Pauses that occur after these prosodic boundary cues have a demarcative function [19] and were labeled as such, as shown in Figure 4. Pauses that follow a flat F0 and break syntactic linearity signal hesitation and were labeled as such, as shown in Figure 5. Filled pauses and lengthenings following a silent pause and beginning a clause were also considered hesitation marks. We measured the median duration of both demarcation and hesitation pauses.

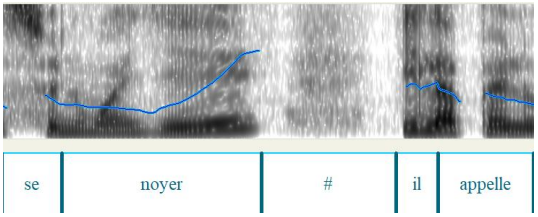


Figure 2: *Silent pause with a demarcative function between two prosodic and syntactic domains ("he's starting to drown # he's calling for help").*

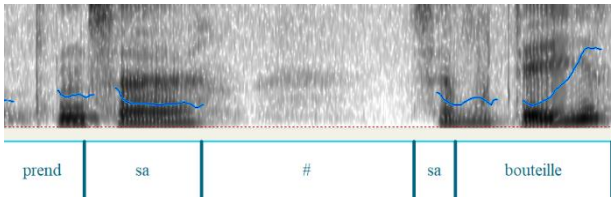


Figure 3: *Silent pause with a hesitation function within a syntactic unit ("he's taking his # his bottle").*

Table 3 shows median pause duration for each pause function between groups in the memory and picture conditions. Demarcation pauses were overall longer than hesitation pauses, except hesitation pauses produced by the BCcog group in the picture condition (1.120 ms).

Table 3: *Median pause duration by function*

Function	**BC**	**BCcog**	**HC**
Memory			
Demarcative pauses	1.065	1.280	0.865
Hesitation pauses	0.460	0.510	0.390
Picture			
Demarcative pauses	0.845	0.851	0.570
Hesitation pauses	0.458	**1.120**	0.410

4. Discussion

The statistical analyses yielded no conclusive results for pause duration, speech rate, and F0 variability. However, a finer-grained investigation of pause functions showed differences between groups. Cancer survivors with a cognitive complaint produced longer hesitation pauses than those without any cognitive complaint and controls. This finding suggests that targeting only hesitation pauses might be relevant for identifying subtle cognitive impairment with a prosodic analysis.

Furthermore, in relation with task type, the descriptive results showed that survivors with a cognitive complaint produced longer pauses in the picture task compared to the memory task. This contradicts our hypothesis stating that survivors would produce longer pauses in the memory-based narrative task due to their reported memory concerns. However, this result is in line with previous research on pauses in individuals with early Alzheimer's Disease, which showed that patients paused more during a picture narrative than in a memory-based narrative [20]. According to the authors, memory-based narratives might be more ecological than picture-based narratives, and might provide an easier context for participants to complete the task. This raises the question of cooccurring cognitive deficits underlying CRCI for cancer survivors. Indeed, a longer hesitation pause duration might reflect the increase of cognitive load due to the nature of the task, which requires visual exploration as well as lexical retrieval and speech planification. Although memory concerns are frequently reported among this population, they might overshadow other deficits than memory such as attentional and executive processes.

Finally, our results may be limited by the small number of participants. To estimate the number of cancer participants with a cognitive complaint and healthy controls required to reach significance for pause duration ratio (with a level of α below 0.05 and a power level above 80%), we performed a power analysis. Twenty participants per group are needed to show a significant difference in the picture condition between BCcog and HC. Otherwise, thirty participants per group would be needed to conclude that there is no difference in memory and picture conditions between BCcog and HC.

5. Conclusions

We conducted a speech analysis using three prosodic markers (pause duration ratio, speech rate and F0 variability) in two narrative tasks (memory-based and picture-based) in an attempt to discriminate cancer survivors with a cognitive complaint from those without any complaint and healthy controls. Although the mixed-effects model and Bayesian analyses showed no significant results, survivors with a cognitive complaint had a higher pause duration ratio than the other groups. A finer-grained analysis of pause function revealed that survivors with a cognitive complaint paused longer especially in the picture-based narrative task. This suggests that identifying hesitation pauses might help detect subtle cognitive impairment in cancer survivors. Furthermore, a longer pause duration in the picture condition might reflect an increase of cognitive load due to underlying deficits such as visual attention and executive processes. Further investigation will thus be needed to confirm that hesitation pauses in picture condition is a good prosodic marker of Cancer-Related Cognitive Impairment.

6. References

[1] O. Ivanova, I. Martínez-Nicolás, and J. J. G. Meilán, 'Speech changes in old age: Methodological considerations for speech-based discrimination of healthy ageing and Alzheimer's disease', *Intl J Lang & Comm Disor*, pp. 1460-6984.12888, May 2023, doi: 10.1111/1460-6984.12888.

[2] J. A. Mefford *et al.*, 'Varied performance of picture description task as a screening tool across MCI subtypes', *PLOS Digit Health*, vol. 2, no. 3, p. e0000197, Mar. 2023, doi: 10.1371/journal.pdig.0000197.

[3] C. De Looze *et al.*, 'Cognitive and Structural Correlates of Conversational Speech Timing in Mild Cognitive Impairment and Mild-to-Moderate Alzheimer's Disease: Relevance for Early Detection Approaches', *Front. Aging Neurosci.*, vol. 13, p. 637404, Apr. 2021, doi: 10.3389/fnagi.2021.637404.

[4] J. E. Mack *et al.*, 'What do pauses in narrative production reveal about the nature of word retrieval deficits in PPA?', *Neuropsychologia*, vol. 77, pp. 211–222, Oct. 2015, doi: 10.1016/j.neuropsychologia.2015.08.019.

[5] M. Belz, 'Defining Filler Particles: A Phonetic Account of the Terminology, Form, and Grammatical Classification of "Filled Pauses"', *Languages*, vol. 8, no. 1, p. 57, Feb. 2023, doi: 10.3390/languages8010057.

[6] G. Bolton and A. Isaacs, 'Women's experiences of cancer-related cognitive impairment, its impact on daily life and care received for it following treatment for breast cancer', *Psychology, Health & Medicine*, vol. 23, no. 10, pp. 1261–1274, 2018, doi: https://doi.org/10.1080/13548506.2018.1500023.

[7] F. M. Henderson, A. J. Cross, and A. R. Baraniak, '"A new normal with chemobrain": Experiences of the impact of chemotherapy-related cognitive deficits in long-term breast cancer survivors', *Health Psychology Open*, vol. 6, no. 1, p. 205510291983223, Jan. 2019, doi: 10.1177/2055102919832234.

[8] V. J. Bray, H. M. Dhillon, and J. L. Vardy, 'Systematic review of self-reported cognitive function in cancer patients following chemotherapy treatment', *J Cancer Surviv*, vol. 12, no. 4, pp. 537–559, Aug. 2018, doi: 10.1007/s11764-018-0692-x.

[9] I. Martínez-Nicolás, T. E. Llorente, F. Martínez-Sánchez, and J. J. G. Meilán, 'Ten Years of Research on Automatic Voice and Speech Analysis of People With Alzheimer's Disease and Mild Cognitive Impairment: A Systematic Review Article', *Front. Psychol.*, vol. 12, p. 620251, Mar. 2021, doi: 10.3389/fpsyg.2021.620251.

[10] J. E. Fardell, V. Bray, M. L. Bell, B. Rabe, H. Dhillon, and J. L. Vardy, 'Screening for cognitive symptoms among cancer patients during chemotherapy: Sensitivity and specificity of a single item self-report cognitive change score', *Psycho-Oncology*, vol. 31, no. 8, pp. 1294–1301, Aug. 2022, doi: 10.1002/pon.5928.

[11] C. Bézy, A. Renard, and J. Pariente, *GREMOTS Batterie d'évaluation des troubles du langage dans les maladies neurodégénératives*. Louvain-la-Neuve, Belgique: De Boeck Supérieur, 2016.

[12] Y. Joanette, B. Ska, and H. Côté, *Protocole Montréal d'Evaluation de la Communication*. Isbergues, France : Ortho Edition, 2004.

[13] E. Campione and J. Véronis, 'A Large-Scale Multilingual Study of Silent Pause Duration', presented at the SPeech Prosody, Aix-En-Provence, 2002.

[14] P. Boersma and D. Weenink, 'Praat: Doing Phonetics by Computer'. 2024. Accessed: Oct. 26, 2021. [Online]. Available: www.praat.org

[15] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, 'Robust Speech Recognition via Large-Scale Weak Supervision', 2022, doi: 10.48550/ARXIV.2212.04356.

[16] B. Bigi, 'SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech', *the Phonetician" - International Society of Phonetic Science*, vol. 111–112, no. 2015-1–11, pp. 5–69, 2015.

[17] D. Bates, M. Mächler, B. Bolker, and S. Walker, 'Fitting Linear Mixed-Effects Models Using **lme4**', *J. Stat. Soft.*, vol. 67, no. 1, 2015, doi: 10.18637/jss.v067.i01.

[18] P.-C. Bürkner, 'Bayesian Item Response Modeling in *R* with **brms** and *Stan*', *J. Stat. Soft.*, vol. 100, no. 5, 2021, doi: 10.18637/jss.v100.i05.

[19] C. Astésano, 'Prosodic characteristics of Reference French', in *Varieties of Spoken French*, 1st ed., S. Detey, J. Durand, B. Laks, and C. Lyche, Eds., Oxford University PressOxford, 2016, pp. 68–85. doi: 10.1093/acprof:oso/9780199573714.003.0006.

[20] A. Pistono, J. Pariente, C. Bézy, B. Lemesle, J. Le Men, and M. Jucla, 'What happens when nothing happens? An investigation of pauses as a compensatory mechanism in early Alzheimer's disease', *Neuropsychologia*, vol. 124, pp. 133–143, Feb. 2019, doi: 10.1016/j.neuropsychologia.2018.12.018.

Self-interruptions in Breast Cancer Patients Who Complain of Anomia

---

Self-interruptions in Breast Cancer Patients Who Complain of Anomia

Amélie B. RICHARD[1,2], Fabrice HIRSCH[1], Sophie JACQUIN-COURTOIS[2,3,] Karen T. REILLY[2]

(1) Paul-Valery Montpellier 3 University, CNRS UMR5267, Praxiling, 34090 Montpellier, France
(2) Université de Lyon, Université Lyon 1, INSERM U1028; CNRS UMR5292; Lyon Neuroscience Research Center, Trajectories Team, F-69676 Lyon, France
(3) Physical Medicine and Rehabilitation Department, Henry-Gabrielle Hospital, Hospices Civils de Lyon, 69610 Pierre-Bénite, France

ABSTRACT

Patients with cancer and cancer survivors sometimes complain of cognitive difficulties related to the cancer and its treatments. Even though one of their most frequent complaints is anomia, currently used neuropsychological tests are unable to detect word-finding difficulties.
The aim of this study was to explore the relationship between pauses at minor boundaries and their right contexts to better describe pauses associated with subtle anomia.
The spontaneous speech of seventeen patients with breast cancer was recorded. Pauses in suspensive interruptions (i.e., the phrase continues after the pause) and disfluent interruptions (i.e., the phrase is modified after the pause) were analysed as well as their right contexts (i.e., lexical-semantic content, independent phrases, or tool words).
The main result showed that suspensive pauses are more likely to be followed by lexical content. This finding suggests that the analysis of pauses in syntactic disruptions could provide evidence about subtle anomia.

Keywords: Breast cancer, pauses, anomia, disfluencies.

1. INTRODUCTION

Patients with cancer and cancer survivors sometimes complain of cognitive changes, called Cancer-Related Cognitive Impairment (CRCI). CRCI negatively affects patients' quality of life by limiting their social participation, and in some cases compromising their return to work. Patients complain of difficulties in the domains of memory (forgetting names and dates) [1], executive functioning (trouble multitasking) [2], attention (distractibility), and language [3]. One of the most frequent language complaints is anomia. According to self-report studies, difficulties with word-finding or word-retrieval in everyday speech could concern between one third and up to 100 percent of patients with CRCI [4], [5].

While the cognitive complaint can be reliably identified using validated questionnaires, neuropsychological tests struggle to detect many difficulties [6]. This discrepancy between subjective and objective assessments may be due to a lack of sensitivity and/or ecological validity of the latter [7]. Indeed, the tasks currently in use were originally designed for use with more severe disorders like aphasia, and do not correspond to the subtle nature of most cancer patients' difficulties [6]. Hence, there is a need for new diagnostic methods.
One solution, with high sensitivity and ecological validity, could be phonetics and discourse analysis. This approach has already been used in research on Mild Cognitive Impairment (MCI) [8], [9], and although MCI patients have a very different aetiology and evolution, their language complaint resembles that of patients with CRCI. These studies on MCI have shown differences between speech in patients and healthy controls at a prosodic level (e.g., a higher number of pauses) [10].
Everyday language contains speech errors called disfluencies. These include pauses, breaks, or repetitions, and one study with Alzheimer's Disease (AD) patients found an association between the number of disfluencies and word finding difficulties. These patients also produced more pauses at minor syntactic boundaries than healthy controls [11]. The authors interpreted these results as evidence of difficulties with lexical retrieval and planning. Based on these results, the current study investigated whether similar phenomena exist in breast cancer patients who complain of anomia. The goal of this research is to explore the relationship between pauses at minor syntactic boundaries and their right context with the aim of better describing pauses associated with subtle anomia. This exploratory study is part of a larger project aimed at developing a language-based diagnostic tool for CRCI.

2. METHOD

Twenty-three women participating in a physical reconditioning program for patients with breast cancer took part in the study. This rehabilitation program is offered by the Henry Gabrielle hospital

(Saint-Genis-Laval, France) to patients who are at least 6 weeks post-treatment (chemotherapy, radiotherapy, and surgery).
All patients were French native speakers or fluent in French. Patients with a history of neurological or psychiatric disorders, those taking medications that can affect consciousness, and/or having a diagnosed speech language impairment were not included. Patients were included if they reported language difficulties when questioned by a speech and language therapist. In total, seventeen patients were eligible for the study. Six patients were not included due to psychiatric conditions (2), vascular disease (1), language impairment history (2), or no language complaint (2). Since the study design was exploratory, we did not include a control group.
This research is part of the ActiCog study at the Hospices Civils de Lyon.

2.1. Speech samples

Patients were interviewed individually by a speech and language therapist. Each interview lasted approximately 30 minutes and consisted of a brief language examination (to verify the absence of any language impairment prior to the cancer diagnosis) followed by three tasks. In the first task, patients were asked to narrate a story based on five images. In the second task, they were asked to recall as many details as possible from a 5-paragraph text read by the speech and language therapist. Finally, they conversed with the investigator about the destination for their next holidays. The investigator intervened rarely, providing backchannels and questions when necessary, in order for the conversation to continue. Data presented in this study come from the spontaneous discourse related to their next holiday destination.
Interviews were recorded using a Rhode Lavalier Go™ omnidirectional microphone linked to a computer.
Patients completed a brief cognition function screening test (Montreal Cognitive Assessment, MoCA) and a speed of processing test (Symbol Digit Modalities Test, SDMT). They also completed auto-questionnaires about their perceived cognitive function (Functional Assessment of Cancer Therapy-Cognitive Function – Cognitive Function, FACT-Cog), their fatigue (FACIT-F, Functional Assessment of Chronic Illness Therapy – Fatigue), and their anxiety/depression (Hospital Anxiety and Depression Scale, HADS).

2.2. Linguistic features

All interviews were manually transcribed and annotated using PRAAT [12]. Silent pauses over 200 msec, filled pauses such as “euh” (“uh”), and lengthening over 200 msec [13] with a flat F0 contour were all annotated.
Interruptions were labelled using the annotation system proposed by Pallaud and colleagues [14] based upon their analysis of auto-interrupted utterances in spontaneous conversations of French speakers. In their study, Pallaud and colleagues identified two kinds of self-interruptions: suspensive and disfluent. In suspensive interruptions, the utterance continues after the pause without any change in the phrase. In contrast, disfluent interruptions are characterised by a morpho-syntactic disruption whereby the interrupted utterance is modified or left unfinished.
The right context of the pauses was analysed to investigate the type of content following the pause. Three content types were defined: Lexical Phrases (LP) such as nouns, noun phrases, or verbs, Independent Phrases (IP) such as verb phrases, and Tool Phrases (TP) such as grammatical words or locutions.

The extracted variables were Pause type (Suspensive Pause and Disfluent Pause) and right context type (Lexical Phrase, Independent Phrase, and Tool Phrase) were extracted. For each variable, we measured the rate (i.e., number of occurrences / 100 words).

2.3. Statistical analyses

A D’Agostino & Pearson’s test was run to test whether the data follow a normal distribution.
Pearson’s correlations were conducted to investigate the relationship between average pause rate and questionnaire and test scores.
Pause rate data were entered into a two-way ANOVA with factors pause type (Suspensive vs Disfluent) and right context type (LP, IP, and TP) with Bonferroni post hoc tests. Finally, for each pause type and for each patient we plotted pause rate as a function of right context type.

3. RESULTS

3.1 Sample characteristic

The mean age of the 17 patients included in this study was 49.7 (± 9.5).
The neuropsychological assessments did not detect any cognitive problems, whereas the FACT-Cog questionnaire revealed a cognitive complaint (score of < 60 on the perceived cognitive impairment subscale [15]) in 13 /17 patients.
One patient did not complete the anxiety, depression, and fatigue questionnaires. Anxiety and depression scores were low for 13 of the remaining 16 patients.

There was no correlation between pause rate and scores on the HADS anxiety subscale (ρ = - .39, p = .12) and depression subscale (ρ = 0.03, p = .91). Patients reported a moderate level of fatigue but there was no correlation between pause rate and scores on the FACIT-F self-report fatigue questionnaire (ρ = - 11, p = .68).

To ensure that our data set did not include data from patients who made a small number of very long pauses, we calculated the correlation between the number of the pauses and average pause duration. The relationship between these two variables was highly correlated (ρ = 0.72, p <.0001), which attests to the quality of our dataset which did not include data from a small number of very long pauses.

3.2 Relationship between Pause Type and Right Contexts

Figure 1 shows the average pause rate separately for the two pause types and the three right contexts. For suspensive pauses, mean pause rate was highest for the LP right context and lowest for the TP right context. Whereas, for Disfluent pauses it was highest for the IP right context and lowest for the TP right context. The two-way ANOVA on these data showed a significant interaction (F(2) = 4.87, p = .009). Bonferroni post-hoc tests revealed a significant difference between LP and TP for both Suspensive (p < .001, 95% C.I. = [-2.24, -0.73]) and Disfluent Pauses (p < .05, 95% C.I. = [-1.43, 0.07]), a significant difference between IP and TP in both Suspensive (p < .01, 95% C.I. = [-1.6, -0.09]), and Disfluent Pauses (p < .001, 95% C.I. = [-1.99, -0.48]). LP and IP pause rates were also significantly different, but only for Suspensive Pauses (p < 0.05, 95% C.I. = []). There was a significant main effect of Right Context (p < .0001), but no effect of Pause Type was found (p =.52).

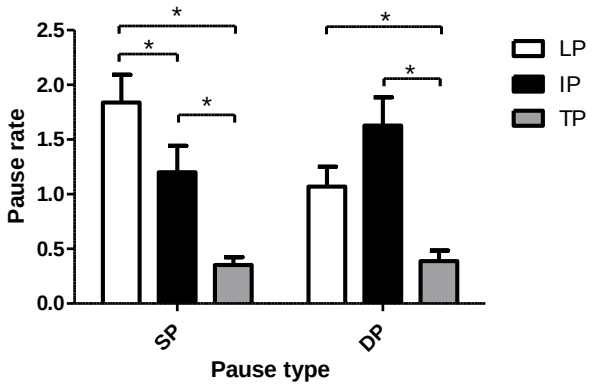


Figure 1: Comparison of mean pause rate for Suspensive and Disfluent Pauses as a function of right context: Lexical Phrase (LP), Independent Phrase (IP), and Tool Phrase (TP).

The mean pause rate data shown in Figure 1 reveal different data patterns for Suspensive and Disfluent pauses. To explore whether data from individual patients followed these same patterns, we plotted each patient's pause rate as a function of right context for Suspensive (Figure 2) and Disfluent (Figure 3) pauses. These figures show that nearly 75% of patients showed the same pattern as the mean pause rate data shown in Figure 1 (13/17 for suspensive pauses and 12/17 for disfluent pauses).

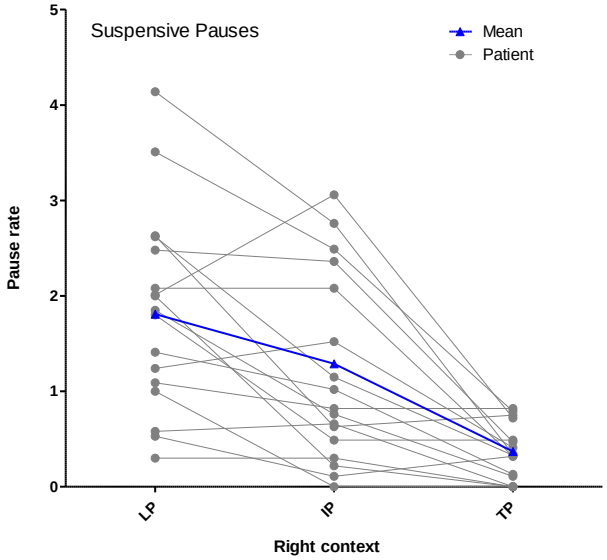


Figure 2: Right context pattern for each patient for Suspensive Pauses. Each line represents a patient, and the blue line represents the mean for all patients.

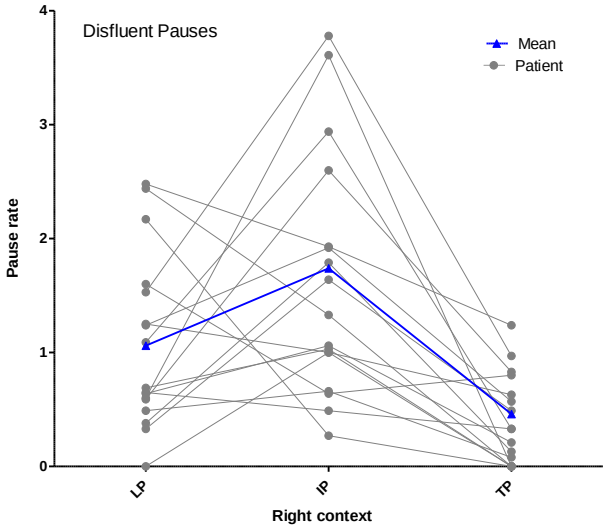


Figure 3: Right context pattern for each patient for Disfluent Pauses. Each line represents a patient, and the blue line represents the mean for all patients.

4. DISCUSSION

The spontaneous speech of seventeen women who had completed treatment for breast cancer and who complained of anomia was analysed.

4.1 Pause type

In the present corpus, pause rates showed a different pattern across right contexts for suspensive and disfluent pauses. Suspensive pauses occurred more often before lexical-semantic content than before independent phrases or tool words. In contrast, disfluent pauses occurred more often before independent phrases than before lexical-semantic content or tool words. This pattern of results suggests that suspensive pauses could be more likely to be associated with lexical retrieval processes.

Tool phrases are generally less frequent than other content type. The grammatical content analysed in the current study consisted of tool words such as determiners, or pronouns, as well as discourse markers (e.g., “voilà”, “quoi”, “en fait”). In the current study, tool words were often found before another break between two weak boundaries, suggesting that they might be part of lexical or syntactic difficulties. Discourse markers were less frequent than tool words, and often occurred before a strong boundary, suggesting that they might help the patient end her unfinished utterance.

While we deliberately separated our pauses into Suspensive and Disfluent, when the data from the two pause types are combined (the main effect of Right Context in our two-way ANOVA), lexical-semantic content and independent phrases follow pauses just as often as each other. While the interaction we observed means that these data should be interpreted with caution, they nevertheless suggest two possible hypotheses. First, anomia could be hidden in rephrasing, whereby patients who have trouble with word-finding use a new phrase to express their idea without the sought word. Second, the anomia reported by patients could arise from more global cognitive difficulties. For example, if anomia is caused by CRCI-related mechanisms then it could be a symptom of changes in cognitive functions including attention, memory, and executive functioning. Indeed, altered function in any one of these cognitive domains could have a direct impact on language. If this is the case, then patients who complain of anomia are likely to encounter trouble with phrasing or difficulties in syntactic planning.

4.2 Pause rate patterns

Nearly 75% of patients had pause rate patterns similar to the mean rate of the current sample. Research on the onset of CRCI is scarce. In contrast, we know that once cognitive difficulties have been identified they can last anywhere between two and ten years after the end of treatment [16]. Since these cognitive difficulties can compromise return to work [17] and limit social activities, there is an urgent need for early identification of patients who are potentially at risk of developing them. Comparing an individual’s pause rate pattern to patterns derived from a large population of cancer patients as well as a control population could be a reliable method for identifying those patients most at risk of developing long-lasting cognitive impairment.

4.3 Relationship between cofactors and pause rates

Consistent with the literature [6], scores on the cognitive complaint scale revealed more cognitive difficulties than the neuropsychological test scores. However, no correlation was found between pause rates and either cognitive complaint scores or neuropsychological test scores. Similarly, pause rates were not correlated with depression and anxiety or fatigue. Hence, pause rates could be independent of these psychological factors.

4.4 Strengths and limitations

This study is the first to explore pauses at minor syntactic boundaries with the goal of investigating phonetic markers of subtle anomia. Its major strength relies on the discourse type it analysed. Indeed, spontaneous speech provides the best ecological validity because of its proximity to everyday discourse.

Due to the exploratory nature of this study, no control group was included. Further investigations with healthy individuals and cancer patients with no complaint of anomia are necessary to achieve the long-term goal of developing a new method for the diagnosis of subtle anomia.

5. REFERENCES

[1] M. L. Kanaskie et S. J. Loeb, « The experience of cognitive change in women with breast cancer following chemotherapy », *J Cancer Surviv*, vol. 9, n° 3, p. 375-387, sept. 2015, doi: 10.1007/s11764-014-0387-x.

[2] F. M. Henderson, A. J. Cross, et A. R. Baraniak, « ‘A new normal with chemobrain’: Experiences of the impact of chemotherapy-related cognitive deficits in long-term breast cancer survivors », *Health Psychology Open*, vol. 6, n° 1, p. 205510291983223, janv. 2019, doi: 10.1177/2055102919832234.

[3] G. Bolton et A. Isaacs, « Women’s experiences of cancer-related cognitive impairment, its impact on daily life and care received for it following treatment

for breast cancer », *Psychology, Health & Medicine*, vol. 23, nº 10, p. 1261-1274, 2018, doi: https://doi.org/10.1080/13548506.2018.1500023.

[4] J. S. Myers, « Chemotherapy-Related Cognitive Impairment: The Breast Cancer Experience », *Oncology Nursing Forum*, vol. 39, nº 1, p. E31-E40, janv. 2012, doi: 10.1188/12.ONF.E31-E40.

[5] D. Von Ah, B. Habermann, J. S. Carpenter, et B. L. Schneider, « Impact of perceived cognitive impairment in breast cancer survivors », *European Journal of Oncology Nursing*, vol. 17, p. 236-241, 2013, doi: http://dx.doi.org/10.1016/j.ejon.2012.06.002.

[6] D. S. J. Costa et J. E. Fardell, « Why Are Objective and Perceived Cognitive Function Weakly Correlated in Patients With Cancer? », *Journal of Clinical Oncology*, vol. 37, nº 14, p. 1154-1158, 2019, doi: https://doi.org/10. 1200/JCO.18.02363.

[7] A. Hutchinson, J. R. Hosking, G. Kichenadasse, J. Mattiske, et C. Wilson, « Objective and subjective cognitive impairment following chemotherapy for cancer: A systematic review », *Cancer Treatment Reviews*, vol. 38, p. 926-934, 2012, doi: http://dx.doi.org/10.1016/j.ctrv.2012.05.002.

[8] L. Toth *et al.*, « A Speech Recognition-based Solution for the Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech », *CAR*, vol. 15, nº 2, p. 130-138, doi: 10.2174/1567205014666171121114930.

[9] C. De Looze *et al.*, « Cognitive and Structural Correlates of Conversational Speech Timing in Mild Cognitive Impairment and Mild-to-Moderate Alzheimer's Disease: Relevance for Early Detection Approaches », *Front. Aging Neurosci.*, vol. 13, p. 637404, avr. 2021, doi: 10.3389/fnagi.2021.637404.

[10] K. W. Kim, S.-H. Na, Y.-C. Chung, et B.-S. Shin, « A Comparison of Speech Features between Mild Cognitive Impairment and Healthy Aging Groups », *Dement Neurocogn Disord*, vol. 20, nº 4, p. 52, 2021, doi: 10.12779/dnd.2021.20.4.52.

[11] F. Gayraud, H.-R. Lee, et M. Barkat-Defradas, « Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects », *Clinical Linguistics & Phonetics*, vol. 25, nº 3, p. 198-209, mars 2011, doi: 10.3109/02699206.2010.521612.

[12] P. Boersma et D. Weenink, « Praat: Doing Phonetics by Computer ». 2021. Consulté le: 26 octobre 2021. [En ligne]. Disponible sur: www.praat.org

[13] J.-P. Goldman, S. Roekhaut, et A.-C. Simon, « Étude statistique de la durée pausale dans différents styles de parole », in *Actes des 28èmes journées d'étude sur la parole (JEP)*, 2010, p. 161-164. [En ligne]. Disponible sur: http:// hdl.handle.net/2078.1/81909

[14] B. Pallaud, R. Bertrand, P. Blache, L. Prévot, et S. Rauzy, « Suspensive and disfluent self interruptions in French language interactions », in *Fluency and Disfluency across Languages and Language Varieties*, Corpora and Language in Use – Proceedings 4., L. Degand, G. Gilquin, L. Meurant, et A.-C. Simon, Éd. Louvain-la-Neuve: Presses universitaires de Louvain, 2019, p. 109-138.

[15] K. Van Dyk, C. M. Crespi, L. Petersen, et P. A. Ganz, « Identifying Cancer-Related Cognitive Impairment Using the FACT-Cog Perceived Cognitive Impairment », *JNCI Cancer Spectrum*, vol. 4, nº 1, p. pkz099, févr. 2020, doi: 10.1093/jncics/pkz099.

[16] M. B. de Ruiter *et al.*, « Cerebral hyporesponsiveness and cognitive impairment 10 years after chemotherapy for breast cancer », *Hum. Brain Mapp.*, vol. 32, nº 8, p. 1206-1219, août 2011, doi: 10.1002/hbm.21102.

[17] R. Bijker, S. F. A. Duijts, S. N. Smith, R. de Wildt-Liesveld, J. R. Anema, et B. J. Regeer, « Functional Impairments and Work-Related Outcomes in Breast Cancer Survivors: A Systematic Review », *J Occup Rehabil*, vol. 28, nº 3, p. 429-451, sept. 2018, doi: 10.1007/s10926-017-9736-8.

Speech and Oculomotor Features of Cancer-Related Cognitive Impairment: A Comprehensive English Summary

Authors' Contribution

This section states the contributions of the doctoral student and her collaborators to the development of this dissertation. The CRediT taxonomy (*Contributor Role Taxonomy*) is used to identify the contribution of each co-author (see https://credit.niso.org/)

PhD Candidate's Contribution

Amélie Richard, throughout the dissertation, articles and proceedings: Conceptualization, Methodology, Software, Formal analysis, Resources, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project administration.

Co-authors' Contribution

Maude Beaudoin for the article Oculomotor behavior of breast cancer survivors reporting cognitive changes: Supervision.

Anne Cheylus, for the article Oculomotor behavior of breast cancer survivors reporting cognitive changes: Software.

Alexandre Foncelle, for the article Using machine learning to identify speech markers of reported cognitive impairment in breast cancer survivors, as well as for the proceedings Speech markers of Cancer-Related Cognitive Impairment: A pilot study and Self-interruptions in Breast Cancer Patients Who Complain of Anomia: Formal analysis, Data Curation, Writing – Review & Editing.

Fabrice Hirsch, for the article Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment, as well as for the proceedings Speech markers of Cancer-Related Cognitive Impairment: A pilot study and Self-interruptions in Breast Cancer Patients Who Complain of Anomia: Writing – Review & Editing, Supervision.

Sophie Jacquin-Courtois, for the articles Using machine learning to identify speech markers of reported cognitive impairment in breast cancer survivors, Assessing Cancer-Related Cognitive Impairment for breast cancer survivors with speech analysis and Méthode d'analyse des pauses pour

l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment, as well as for the proceedings Speech markers of Cancer-Related Cognitive Impairment: A pilot study and Self-interruptions in Breast Cancer Patients Who Complain of Anomia: Resources, Supervision. For the article Assessing Cancer-Related Cognitive Impairment for breast cancer survivors with speech analysis: Conceptualization, Writing – Review & Editing, Supervision. For the article Oculomotor behavior of breast cancer survivors reporting cognitive changes: Supervision.

Manon Lelandais, for the article Linguistic markers of subtle cognitive impairment in connected speech: A Systematic Review: Conceptualization, Methodology, Data Curation, Writing – Original Draft, Review & Editing, Supervision. For the article Assessing Cancer-Related Cognitive Impairment for breast cancer survivors with speech analysis , as well as the proceedings Speech markers of Cancer-Related Cognitive Impairment: A pilot study: Writing – Original Draft, Writing – Review & Editing.

Pauline Mouchès, for the article Using machine learning to identify speech markers of reported cognitive impairment in breast cancer survivors: Methodology, Formal analysis, Data Curation, Writing – Review & Editing, Supervision.

Laure Pisella, for the article Oculomotor behavior of breast cancer survivors reporting cognitive changes: Conceptualization, Writing – Review & Editing.

Karen T. Reilly for the article Linguistic markers of subtle cognitive impairment in connected speech: A Systematic Review as well as for the proceedings Speech markers of Cancer-Related Cognitive Impairment: A pilot study and Self-interruptions in Breast Cancer Patients Who Complain of Anomia': Conceptualization, Writing – Review & Editing, Supervision.

CRediT's 14 Contributor Roles

- **Conceptualization**: Ideas; formulation or evolution of overarching research goals and aims
- **Methodology**: Development or design of methodology; creation of models
- **Software**: Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components
- **Validation**: Verification, whether as a part of the activity or separate, of the overall replication/ reproducibility of results/experiments and other research outputs
- **Formal analysis**: Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data
- **Investigation**: Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection
- **Resources**: Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools
- **Data Curation**: Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later reuse
- **Writing – Original Draft**: Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation)
- **Writing – Review & Editing**: Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or postpublication stages
- **Visualization**: Preparation, creation and/or presentation of the published work, specifically visualization/data presentation
- **Supervision**: Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team
- **Project administration**: Management and coordination responsibility for the research activity planning and execution
- **Funding acquisition**: Acquisition of the financial support for the project leading to this publication

Introduction

Speech contains fine elements whose granularity allows us to understand the functioning of human cognition. Accordingly, speech analysis stands as a promising method for detecting subtle cognitive disorders. Specifically, identifying prosodic variables could help the diagnosis of cognitive disorders that are difficult to reveal by current speech-language pathology or neuropsychological tests, such as Cancer-Related Cognitive Impairment (CRCI). This impairment manifests itself through mild symptoms, although its impact on the patients' quality of life is substantial. Among other problems, patients describe difficulties with memory, attention maintenance, performing routine tasks, and speaking fluently. Word-finding difficulties are one of the most commonly reported concerns, regardless of the type of cancer or its treatments. These functional alterations significantly interfere with the patients' daily lives. Therefore, an early detection seems essential in order to limit the long-term effects of CRCI.

CRCI can be revealed through standardized questionnaires that collect the symptoms experienced by patients in their daily lives. However, current neuropsychological tests often struggle to detect this disorder. This limit results in a discrepancy between perceived difficulties and the actual scores, which remain within normal range. The lack of sensitivity in diagnosis tools raises several issues. On the one hand, CRCI is an invisible disability, as patients tend to compensate for or minimize its manifestations (Lewis et al., 2016). The subtlety of these symptoms often causes a gap between the patients' perception of their difficulties and their recognition by the medical community. In the absence of a clear diagnosis, patients cannot receive the necessary care. On the other hand, the variety of possible cognitive deficits associated with this disorder (whether related to memory, attention, language, etc.) complicates the understanding of the processes that lead to CRCI. In other words, although patients describe a range of cognitive problems, from memory to language, including executive functions and attention, it remains challenging to precisely pinpoint the nature and extent of these alterations. The gap in knowledge slows down the development of tailored therapeutic interventions.

The interaction between language and other cognitive functions offers the opportunity to investigate CRCI in a way that is closely tied to the patients' everyday environment. Speaking is an ordinary act, and even the slightest change in speech can provide insights into the cognitive status of the speaker. However, speech is subject to considerable inter-individual variability, and its analysis presents a complexity inherent to the properties of language. On its own, speech does not always allow for a clear interpretation of the cognitive processes underlying language production. Therefore, the addition of a complementary approach such as eye-tracking can strengthen hypotheses regarding the mechanisms recruited in speech production.

Relevance of the Study

The perspectives offered by speech analysis to reveal the presence of CRCI are encouraging. However, its application raises several methodological questions.

First, only two psycholinguistic studies have been made within the context of CRCI (Myers et al., 2022; K. Williams et al., 2021). These studies focused on syntactic complexity in the speech of patients who had recovered from prostate cancer. While they highlight the value of analyzing linguistic parameters for the detection of CRCI, using syntax as an entry point, seems to introduce some inter-rater biases. These factors compromise both the feasibility and replicability of such studies.

In contrast, prosodic analysis is based on the measurement of acoustic signal. The numerous computerized and automated tools dedicated to acoustic signal processing now available make the analysis of prosody a promising method for detecting CRCI. However, the best parameters for detecting cognitive disorders is inconsistent. Yet, the variety of documented pathologies and the diversity in the examined variables prevent the identification of potential CRCI markers.

Research Question

Qualitative studies on CRCI are valuable resources regarding the difficulties experienced by patients. Interviews conducted with cancer survivors show that word-finding difficulties (or anomia) are a common source of concern within this population. Moreover, these difficulties are often associated with the report of memory problems. To date, no specific study has been conducted on this phenomenon in the context of CRCI. Consequently, little is known about its underlying cognitive mechanisms. Since patients report memory problems, it may be tempting to assume that their anomia results from a degradation of semantic knowledge. However, clinical observations and studies using neuroimaging data do not support this hypothesis. On the contrary, these studies suggest the presence of attentional or executive deficits. Since the management of word-finding difficulties depends on its cognitive substrates, further insight into this widespread symptom is necessary. Moreover, the effects of word-finding difficulties on patients' discourse production is an interesting entry point to disentangle the underlying mechanisms of CRCI.

The exploratory research reported in this dissertation aims to achieve two objectives: (i) Identify prosodic markers that may indicate the presence of CRCI, (ii) Examine the cognitive processes underlying this disorder.

Hypotheses

First, the speech of breast cancer survivors may contain specific prosodic markers to the cognitive changes they face. To test this hypothesis, a pre-selection of parameters will be incorporated into machine learning models. These models will be able to classify survivors reporting cognitive difficulties, survivors without complaints, and control subjects.

Second, the anomia reported by breast cancer survivors would result from an alteration of attentional or executive dysfunction, rather than a memory deficit. This hypothesis will be investigated through the analysis of oculomotor behavior during speech production. This approach will aim to determine the extent to which the aforementioned cognitive mechanisms influence the lexical retrieval difficulties observed in these patients.

Dissertation Structure

The **Theoretical Framework** first presents a state of the art on Cancer-Related Cognitive Impairment, then offers an overview of the relationship between language and cognition. This chapter continues with the investigation of speech processes through an eye-tracking technique. The **methodology article, Assessing Cancer-Related Cognitive Impairment for breast cancer survivors with speech analysis**, details the experimental framework of this dissertation and justifies the choice of statistical models and machine learning techniques used to examine prosodic variables. The **first article, Linguistic markers of subtle cognitive impairment in connected speech: A Systematic Review**, reviews the commonly used linguistic parameters to evaluate subtle cognitive disorders. The **second article, Using machine learning to identify speech markers of reported cognitive impairment in breast cancer survivors**, proposes a set of prosodic markers that can detect the presence of CRCI using machine learning models. The **third article, Oculomotor behavior of breast cancer survivors reporting cognitive changes**, sheds light on the phenomenon of word-finding difficulties in the context of CRCI by exploring eye-movement behavior during speech production. Finally, the results and findings from these experimental chapters will be discussed as part as a wider context in the **General Discussion**.

Theoretical Framework

1 Introduction

Cancers are widespread heterogeneous diseases around the world. According to the International Agency for Research on Cancer, about twenty million new cases were diagnosed in 2022 (Ferlay et al., 2024). In France, the National Cancer Institute estimates the incidence of diagnosed cancers at 443,136 cases per year, with a prevalence of 3.8 million people living after cancer (Institut National du Cancer, 2024). Thanks to advances in knowledge and technology in cancer research, some malignant tumors such as breast cancer are better diagnosed and treated. As a result, the number of survivors is increasing, and the quality of life after cancer has now become one of the major public health challenges (Institut National du Cancer, 2021). Alongside alopecia, fatigue, and other well-documented side effects, cognitive impairment is a common and disabling aftereffect of cancer. Cancer-Related Cognitive Impairment (CRCI) — sometimes called Cancer-Related Cognitive Changes — refers to a range of cognitive difficulties that affect the daily functioning of survivors. CRCI is a multifactorial syndrome whose underlying cognitive mechanisms are still not well understood.

This theoretical framework is organized in four sections: Section Breast Cancer presents current knowledge on breast cancer. Section Cancer-Related Cognitive Impairment is a state of the art about the cognitive sequelae of cancer. Section Speech: A Gateway into Cognition highlights the benefits of speech analysis to investigate cognition and its dysfunction. Eventually, Section Investigating Linguistic Processes Through Eye Behavior shows the interest of using an eye-tracking technique as a complementary behavioral measure to speech analysis.

2 Breast Cancer

General information

According to data from the International Agency for Research on Cancer, breast cancer was the most diagnosed cancer in the world in 2020, with more than 2.26 million new cases per year (Ferlay et al., 2024). On a national scale, it remains the most common cancer among women and is the second

leading cause of death due to morbidity in this population (Institut National du Cancer, 2024), just behind cardiovascular diseases. However, the improvement in screening, diagnosis, and treatments has made this cancer one of the best-treated in France. The five-year survival rate for breast cancer after diagnosis is 88% (Institut National du Cancer, 2018, 2024).

A cancer cell is characterized by its independence from the signals that control cell growth and division, its ability to resist the process of apoptosis (i.e., cell death), its capacity to migrate (i.e., to move from one environment to another), and to invade adjacent tissues (i.e., to colonize the cells of other environments) (Brown et al., 2023). The reference molecular classification of breast cancers identifies at least four main subtypes of breast cancer (Roy et al. [2023], see Organisation mondiale de la santé [2019] for a complete classification): luminal A and B cancers, HER2-overexpressed cancers, and triple-negative cancers. Luminal A and B breast cancers are hormone-sensitive (HR+). They express estrogen (ER+) and progesterone (PR+) receptors and account for about 70% of breast cancers. HER2-overexpressed (HER2+) or HER2-like cancers, on the other hand, have a receptor that promotes the multiplication of cancer cells. They account for about 15% of breast cancers. Finally, the Basal-Like and triple-negative cancer category refers to the absence of hormonal receptors (ER- and PR-) or the HER2 receptor (HER2-). They represent about 15% of breast cancers. The choice of surgical and pharmacological treatments is guided by this classification, as well as the tumor's location, its severity stage (i.e., its size and the potential presence of metastases), and the patient's health status.

Depending on the size of the tumor, the surgical intervention may involve a lumpectomy or, in some cases, a total mastectomy, including the removal of the nipple. The technique known as "sentinel lymph node biopsy" involves analyzing the first lymph node of the affected breast (Giammarile et al., 2022). If it contains cancer cells, a lymph node dissection will then be considered. Surgery is generally proposed as the first line of treatment, followed by adjuvant pharmacological treatments. Among these treatments, chemotherapy, radiotherapy, hormone therapy, targeted therapies, and immunotherapy are the most commonly used in the management of breast cancer (Burguin et al., 2021; Kerr et al., 2022).

Chemotherapy delivered as a neoadjuvant therapy (before surgery) helps reduce the size of the tumor prior to the procedure. As an adjuvant treatment, it prevents the proliferation of cancer cells. It is most often administered intravenously over a period of three to six months. Various molecules and regimens, such as anthracyclines (e.g., doxorubicin, epirubicin), taxanes (e.g., paclitaxel [Taxol]), and 5-fluorouracil (or 5-FU), are prescribed based on the cancer subtype and its severity (Waks & Winer, 2019).

Targeted therapies aim to inhibit or slow down the action of proteins responsible for the growth of cancer cells. They are administered intravenously, usually over a period of six to twelve months. Immunotherapy, a form of targeted therapy, strengthens the immune system to enhance its ability to fight cancer cells (Debien et al., 2023). The treatments are delivered intravenously over a period of three to six weeks. Monoclonal antibodies, such as trastuzumab, are among the available treatments

(Kerr et al., 2022).

Life After Cancer

The side effects of cancer treatments can have both short-term and long-term impacts on the quality of life of survivors. The VICAN-5 report, which assesses patients' lives in France five years after the end of cancer treatment, highlights a number of sequelae commonly encountered by women who are cured of breast cancer (Institut National du Cancer, 2018). Among these are aesthetic aftereffects, persistent fatigue, chronic pain, psychological disorders, as well as gynecological and functional disturbances (Delrieu et al., 2021). In addition to these effects, cognitive difficulties may arise during the oncology care process and persist after cancer recovery.

3 Cancer-Related Cognitive Impairment

The term CRCI is generally attributed to solid and hematopoietic tumors (i.e., blood cancers) excluding cancers of the Central Nervous System. Brain tumors (e.g., glioblastomas, oligodendrogliomas) are highly likely to cause cognitive dysfunctions due to their location (Meskal et al., 2016; Patchell, 2003). Moreover, CRCI is primarily documented in women with breast cancer due to its high incidence and significant prevalence, meaning better chances of recovery. However, this cognitive impairment is not specific to this population. Patients with other types of cancer, such as testicular, prostate, lung, colon cancer, etc., may also face cognitive difficulties (Demos-Davies et al., 2024).

CRCI refers to a set of cognitive difficulties described by oncology patients, whose symptoms decrease their quality of life. The autonomy of survivors is generally preserved. The clinical expression of CRCI may be subtle, while its functional repercussions on their daily life are significant. Cancer survivors often report, for instance, complaints about short-term memory, concentration problems, and word-finding difficulties. These problems, which disrupt their social interactions and interfere with the performance of routine tasks, may affect up to 75% of women with breast cancer (Lange et al., 2019). However, these figures may be underestimated due to the lack of sufficiently sensitive diagnosis tools to detect CRCI.

CRCI varies in time according to patients, their treatment path, their type of cancer, and the administered treatments (see, for instance, the longitudinal study by Yusuf et al. [2022] on CRCI in patients with multiple myeloma [blood cancer]). Regarding the prevalence of CRCI in women with breast cancer, a systematic review shows that a quarter of patients experience cognitive impairment before receiving adjuvant treatments, 24% immediately after the end of treatments, and 21% six months afterwards (Dijkshoorn et al., 2021). While an increasing number of studies highlight the presence of cognitive changes before starting adjuvant treatments (Ahles et al., 2008), the expression

of the disorder is, in most cases, the highest six months after the end of anticancer medications (Janelsins et al., 2018). This time window corresponds to a period when patients resume their activities and are more tangibly confronted with their cognitive difficulties. For most patients, cognitive complaints tend to decrease within the two years following the end of adjuvant treatments, either due to the spontaneous regression of CRCI or because of coping strategies, i.e., adaptation to cognitive changes, adopted by the patients. However, for some of them, CRCI is a symptom that persists beyond two years after the end of treatments (Whittaker et al., 2022), and even up to ten years later (De Ruiter et al., 2011). These patients describe the feeling of never having fully regained their pre-cancer cognitive abilities. Currently, there is no clinical tool exists for early detection of patients at risk of developing persistent symptoms, or for determining vulnerability factors.

Risk Factors of CRCI

To date, the origins of CRCI remain poorly understood. This section describes the main factors involved in the onset and/or worsening of CRCI.

Cancer Biology

CRCI has been previously labeled as chemofog or chemobrain due to the association of its symptoms with chemotherapy (Hurria et al., 2007). However, advances in research have shown that CRCI could affect nearly 40% of patients even before treatment begins (Demos-Davies et al., 2024). Other factors, such as other oncological treatments or the cancer itself, may play a role in its onset (Olson & Marks, 2019). For instance, Wefel et al. (2011) highlighted the existence of cognitive complaints and dysfunction in men diagnosed with testicular cancer before the administration of treatments. Nearly half of their sixty-nine patients had cognitive deficits in areas such as motor functions, verbal learning, and executive functions (i.e., inhibition, updating, and mental flexibility [Miyake & Friedman, 2012]) at the baseline assessment (Wefel et al., 2011). Similar results were found by Quesnel et al. (2009) in women treated for breast cancer. In this longitudinal study, the women reported lower cognitive performance compared to controls prior to chemotherapy and radiotherapy (Quesnel et al., 2009). These observations were confirmed by a later study, which incorporated brain imaging techniques in addition to neuropsychological assessments (Kesler et al., 2017). The dynamics of activated brain networks were significantly lower in patients, especially during tasks involving memory, executive functions, and verbal fluency (Kesler et al., 2017). According to the authors, these results highlight the impact of cancers (excluding brain tumors) on brain function, for instance, via inflammatory mechanisms. Cancer may provoke an inflammatory response that activates neurotoxic cytokines, such as interleukins and Tumor Necrosis Factor (TNF), which are often associated with CRCI (Lyon et al., 2016). Cytokines are proteins involved in the functioning of the immune system. Among them, interleukins modulate cell growth, differentiation, and activation during inflammatory and immune responses, while TNF regulates inflammation and sends signals to immune cells to destroy

tumor cells. In the case of malignant tumors, these cytokines weaken and cross the blood-brain barrier, which is the cellular tissue that separates blood from cerebrospinal fluid and protects the brain from pathogens and toxins in the blood. Cytokines thus stimulate brain inflammation by causing a dysregulation of neurotransmitter activity involved in proper cognitive functioning (Mayo et al., 2021). Moreover, a high concentration of pro-inflammatory cytokines would impair cognitive performance (Henneghan et al., 2020), such as short-term memory abilities (A. M. Williams et al., 2018).

Cancer Treatments

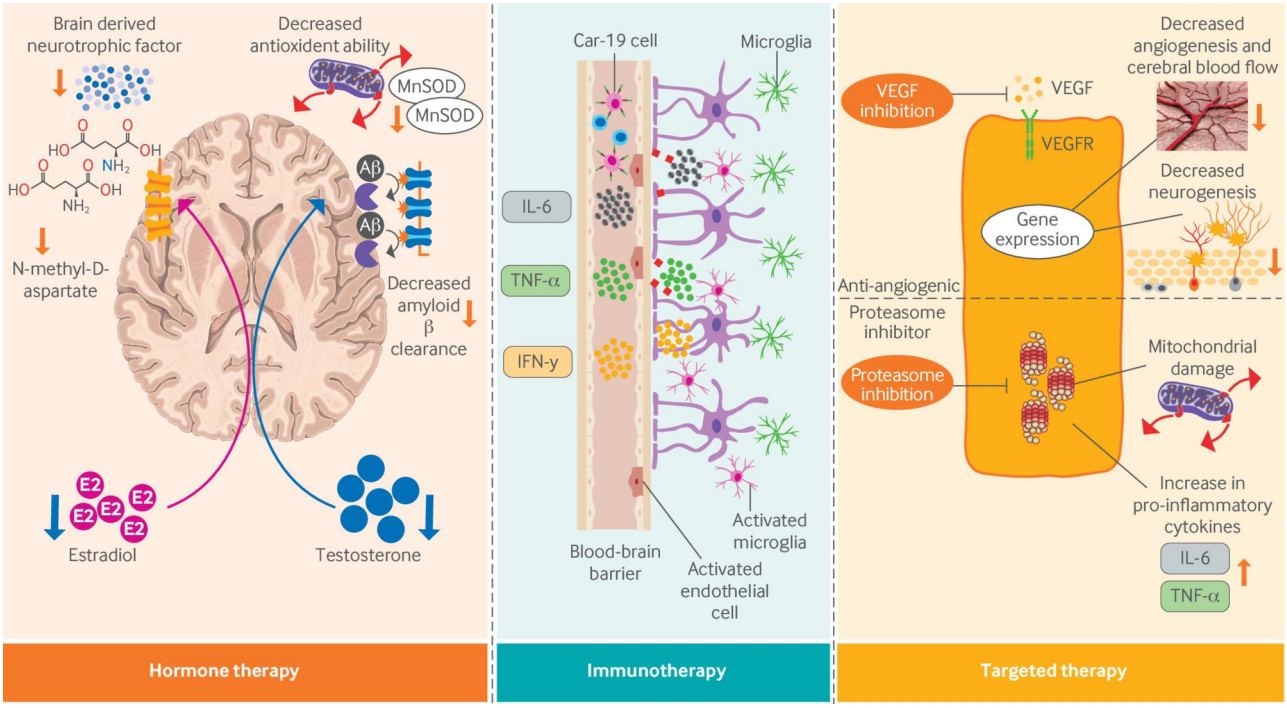
Cognitive difficulties are particularly pronounced during the treatment period. Some medications cross the blood-brain barrier (Deeken & Löscher, 2007), and their toxicity could lead to cognitive dysfunction (Hurria & Lachs, 2007; Winocur et al., 2006; Winocur et al., 2018). Although chemotherapy is frequently identified as the main cause of CRCI, it is possible that all anticancer treatments influence not only the onset but also the severity of CRCI (Demos-Davies et al., 2024; Fleming et al., 2023).

Chemotherapy induces oxidative stress (known to weaken the blood-brain barrier) and could increase the level of pro-inflammatory cytokines. During chemotherapy cycles, higher levels of cytokine have been associated with deficits in various cognitive domains, such as psychomotor speed, complex attention, executive functions, verbal and visual memory, and verbal flexibility, although the affected domains vary across studies (Jim et al., 2012; Li & Caeyenberghs, 2018). Furthermore, the neurotoxic effects of chemotherapeutic molecules on the peripheral nervous system are widely recognized (Verstappen et al., 2003). These substances could also have detrimental effects on the brain itself. For instance, traces of 5-fluorouracil have been detected in small quantities in the brains of non-human primate models (Bourke et al., 1973). These brain toxins trigger a cascade of processes that disrupts the repair of damaged DNA, which lead to neuronal death (Ren et al., 2019). Moreover, the intensity of chemotherapy doses appears to exacerbate the severity of CRCI (Van Dam et al., 1998).

Hormone therapy has also been identified as a potential factor in the development of CRCI. The impact of this treatment on cognition has been highlighted through questionnaires given to patients with prostate cancer (Boué et al., 2024). Another study emphasizes the effects of tamoxifen on the decrease in verbal fluency and verbal memory performance in patients with breast cancer (Boele et al., 2015). However, these results are not confirmed by Van Dyk et al. (2019). This gap in results could be related to the methodology used to assess cognitive abilities (Van Dyk et al., 2019).

Figure 1 illustrates the neurological mechanisms associated with hormone, immune, and targeted therapy that might play a part in the development of CRCI, from Fleming et al. (2023).

Figure 1. Possible mechanisms underlying cognitive impairment associated with hormone, immune, and targeted therapy.



Note: Reprinted with permission from Fleming et al. (2023). 2023 BMJ Publishing Group Ltd.

Genetic Factors

The role of three genetic predispositions has now been clearly established in the development and severity of CRCI. The presence of the ε4 allele of Apolipoprotein E (APOE4), a gene also involved in Alzheimer's Disease, would be associated with a decrease in cognitive performance in patients treated with chemotherapy, especially in the areas of visual memory and visuospatial abilities (Ahles et al., 2003). On the other hand, the nucleotide polymorphism (i.e., a genetic variation) of the enzyme catechol-O-methyltransferase (COMT), which is involved in the functioning of the prefrontal cortex, would lead to impaired cognitive performance in breast cancer patients treated with chemotherapy (Small et al., 2011). Finally, the brain-derived neurotrophic factor (BDNF) is a particularly active protein in the hippocampus and playing a key role in long-term memory. It is coded by the BDNF gene, of which the polymorphism seems to reduce the risk of developing CRCI. A study shows that patients with breast cancer carrying a BDNF gene polymorphism report fewer deficits in verbal abilities and dual-task management after chemotherapy, compared to other patients (Ng et al., 2016).

Psychological, Physiological and Socio-Demographic Factors

In addition to biological factors, psychological, physiological, and socio-demographic factors intersect and influence the expression of CRCI. Although these confounding factors do not directly

explain the onset of CRCI, their evaluation provides valuable insight into the cognitive trajectory of patients.

For some, the announcement of a cancer diagnosis is a traumatic event that can generate intense psychological stress, or even post-traumatic stress disorder (PTSD) in the most severe cases. Stress could thus be both a triggering and an aggravating factor for CRCI (Hermelink et al., 2015; Hermelink et al., 2017). However, the relationship between stress and the development of CRCI remains uncertain. Hermelink et al. (2017) studied the impact of post-traumatic stress on cognitive abilities in women with breast cancer. The study compared a group of patients before any treatment, a group of patients who had received chemotherapy, and a group of healthy controls, all of whom underwent a clinical assessment of post-traumatic stress and a battery of neuropsychological tests. The results showed that about half of the participants did not show signs of post-traumatic stress regardless of their group. Additionally, the overlap of scores between the three groups suggests that stress does not have a direct impact on cognitive performance, even though there was a tendency for the chemotherapy-treated group to show poorer cognitive performance. The authors concluded that although women with breast cancer independently of treatment experience subtle cognitive difficulties which may be partially modulated by post-traumatic stress. However, the latter would not explain the onset of CRCI. These results are corroborated by a more recent study indicating that patients who have not yet received treatment report more cognitive difficulties than controls, although there are no notable differences in terms of overall cognitive performance (Aspelund et al., 2024).

Moreover, the prevalence of post-traumatic stress in patients with cancer remains relatively low, around 10% (Arnaboldi et al., 2017). It should also be noted that stress symptoms decrease over time, and there is little evidence supporting the existence of long-term post-traumatic stress disorder in cancer survivors. Another study reveals that, thirty days after the diagnosis, around 19% of patients exhibit post-traumatic stress symptoms, while 70% show anxiety symptoms. However, the intensity of these symptoms significantly decreases within two years of the diagnosis (Arnaboldi et al., 2014). Similarly, psychological evaluations of oncology patients often highlight the presence of depressive and anxious symptoms. A systematic review has shown that between 9.4% and 66.1% of cancer survivors suffer from depression, while between 17.9% and 33.3% exhibit anxiety symptoms. However, these figures are not significantly higher than those observed in the general female population (Maass et al., 2015).

Cancer-related fatigue is another factor that is often associated with cancer and its treatments. This symptom often persists long after treatment has ended and is characterized by a generalized feeling of fatigue that is not relieved by sleep (Hofman et al., 2007; Kang et al., 2023; Rosas et al., 2023; Ruiz-Casado et al., 2021). It is therefore not surprising that this kind of fatigue can lead to a decrease in attentional capacities (Kohler et al., 2020). As a result, some patients may attribute their cognitive complaints to fatigue rather than a true cognitive deficit.

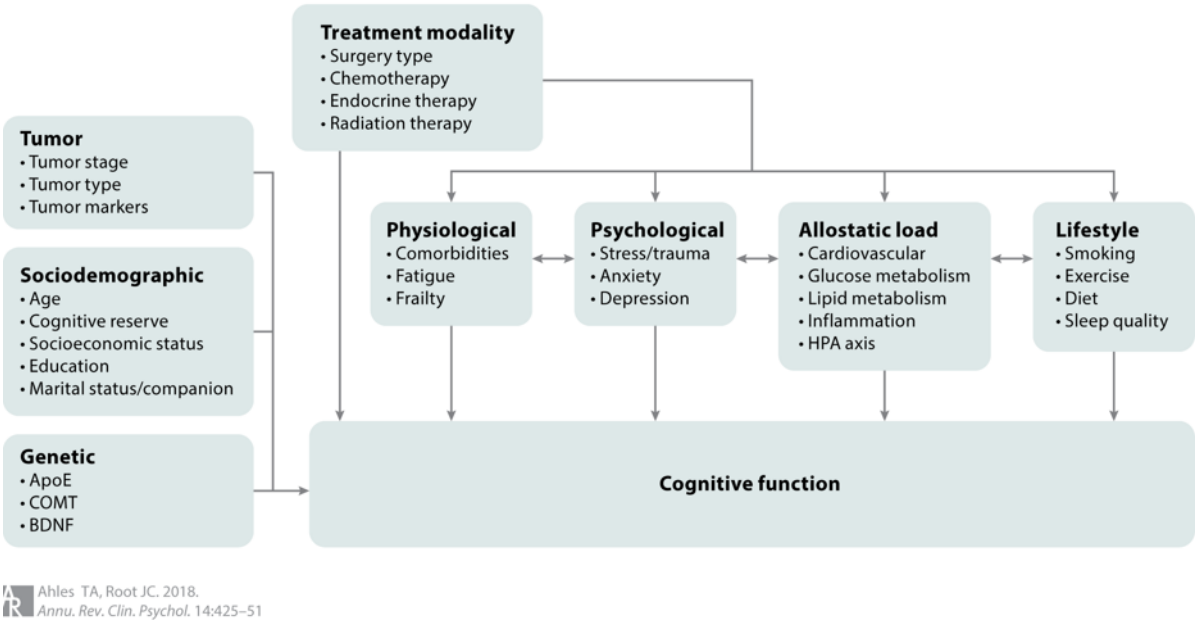
Cognitive reserve (i.e., the brain's ability to compensate for cerebral damage by establishing new neural connections or reactivating existing networks) and biological age (i.e., the cellular degradation

of the body) play a moderating role in the severity of CRCI (Ahles et al., 2010; Carroll et al., 2019). In other words, a low cognitive reserve or an advanced age would increase the risk of developing a persistent disorder (Kesler, 2011; Lange et al., 2016). However, even young patients (under the age of 45) can be affected by long-term cognitive difficulties (Von Ah et al., 2022).

Finally, lifestyle also influences the cognitive abilities of survivors. Obesity, physical activity, and sleep quality would particularly affect the scores on neuropsychological tests, although these scores remain above pathological thresholds (Hartman et al., 2015).

To summarize, CRCI is a multifactorial syndrome that likely results from an interaction between the tumor itself and its treatments, as well as other confounding factors. Ahles and Root (2018) propose an overarching model of the predictive factors of CRCI in cancer survivors (Figure 2).

Figure 2. Schematic of factors predictive of cognitive change in cancer survivors.



Note: Reprinted with permission from Ahles and Root (2018). 2018 Annual Reviews.

Structural and Functional Brain Changes Associated With CRCI

Many studies using neuroimaging have shown structural and functional changes in the brains of oncology patients. These studies are primarily conducted with women diagnosed with breast cancer and highlight the effects of chemotherapy on brain function or on brain structures (Feng et al., 2019).

Effects of Chemotherapy

The structural and functional changes in the brain induced by chemotherapy are increasingly well-identified (Whittaker et al., 2022). For instance, degradation of white matter integrity is observed in the superior longitudinal fasciculus (associated with visuospatial attention and complex motor behaviors) and in the corticospinal tract (responsible for motor commands) (Menning et al., 2018). Deprez et al. (2012) used diffusion magnetic resonance imaging (MRI) combined with a comprehensive cognitive assessment in a follow-up study. They showed a reduction in the activation of frontal, parietal, and occipital tracts after chemotherapy that was linked to impairments in attention and verbal memory abilities (Deprez et al., 2012).

Other research has noted a decrease in gray matter in the prefrontal and parahippocampal gyri as well as in the precuneus in breast cancer survivors who received chemotherapy (Inagaki et al., 2007). However, this effect was not found in patients evaluated several years after chemotherapy, suggesting a possible long-term recovery. These results, however, remain to be confirmed by further studies. Alterations in brain activity sometimes persist for years after the cessation of chemotherapy. For instance, deficits have been identified in the fronto-cortical regions, the basal ganglia, and the cerebellum five to ten years after the end of the treatment (Silverman et al., 2007). Similarly, a study compared the cognitive performance of patients who received high-dose chemotherapy ten years after the treatment to that of untreated patients (De Ruiter et al., 2011). In this study, participants performed tasks involving executive functions and attention processes while their brain activity was recorded using fMRI. Participants who had received chemotherapy exhibited long-term cognitive impairments characterized by hyporeactivity in the dorsolateral prefrontal cortex, parahippocampal gyrus, and posterior lateral parietal cortex (De Ruiter et al., 2011). Finally, Deprez et al. (2014) highlighted functional differences in breast cancer patients treated with chemotherapy through tasks involving executive functions and attention resources (Deprez et al., 2014). These tasks activated various brain regions, such as the bilateral fronto-parietal network, cerebellum, anterior cingulate, and occipitotemporal cortex in both the chemotherapy and the control groups. However, only patients who had received chemotherapy showed a general decrease in brain activity. This reduction was correlated with more severe cognitive complaints in these patients (Deprez et al., 2014).

Impact on Fronto-Parietal Regions and Executive Control

Research has shown the presence of diffuse and fine lesions in various brain regions. However, it is difficult to draw conclusions about a systematic localization of these lesions (Amidi & Wu, 2019). Studies using brain imaging highlight alterations that are primarily located in the parietal and frontal regions. These regions are involved in attention and executive functioning (Cristofori et al., 2019; Knudsen, 2007). A significantly reduced activation of the left dorsolateral prefrontal cortex and the premotor cortex is observed in breast cancer patients, compared to healthy controls. The breast cancer group showed deficits in inhibition and had a slower processing speed (Kesler,

2011). Compared to untreated breast cancer patients and healthy controls, patients who received chemotherapy showed a reduction in bilateral frontal hyperactivation during tasks recruiting executive functions (B. C. McDonald et al., 2012). In another study, brain activity was recorded by fMRI during declarative verbal memory tasks in patients with metastatic breast cancer (Kesler et al., 2009). The results indicated a weaker activation of the prefrontal cortex during memory encoding, compared to control subjects (Kesler et al., 2009). A few studies have also highlighted alterations in the temporal lobes of cancer patients that partially recovered after treatment (e.g., Conroy et al. [2013] and Lepage et al. [2014]). However, the results of these studies are varied and insufficient to signal the existence of a clearly defined lesion site (Amidi & Wu, 2019).

Bernstein et al. (2021) conducted a meta-analysis on structural brain changes after chemotherapy and their possible effects on executive and attentional deficits (Bernstein et al., 2021). This study emphasizes that two brain regions are often associated with CRCI despite the heterogeneity of the methods used: the left superior parietal cortex and the right superior prefrontal regions. In particular, the activation of the fronto-parietal network, involved in attentional processes, would be weaker after chemotherapy. As a result, chemotherapy would impact the regions that regulate attention and executive functions. The results of this meta-analysis are summarized in Figure 3 (Bernstein et al., 2021).

Figure 3. Synthesis of studies showing an association between brain regions and CRCI. Blue dots indicate cerebral hypoactivation. Red dots indicate cerebral hyperactivation.

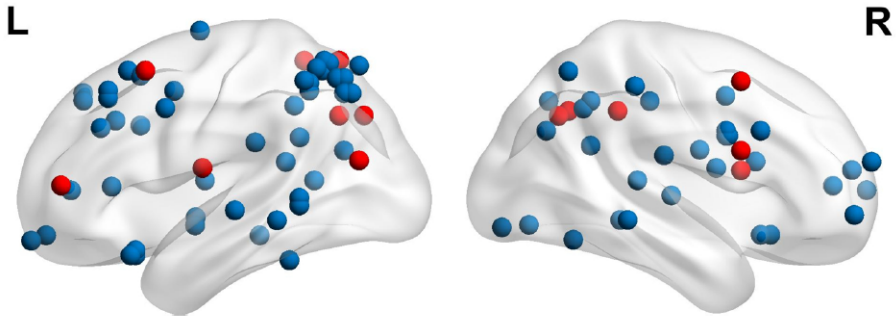


Fig. 2. Foci from studies showing greater activation (Red) and from studies reporting lower activation (Blue) associated with chemotherapy. L = Left; R = Right.

Note: Reprinted with permission from Bernstein et al. (2021). 2021 Elsevier Ltd.

An interesting point that emerges from this research is that differences in structural brain alterations are generally more distinct when cancer groups are compared to healthy controls, but more subtle when compared to other cancer patients (Amidi & Wu, 2019). In other words, brain differences are more evident between patients and healthy control subjects than between chemotherapy-treated patients and those not receiving chemotherapy. This phenomenon suggests that cancer itself might already induce brain changes prior to any treatment, and that chemotherapy could exacerbate these alterations.

Report of Cognitive Difficulties

CRCI may occur in patients with any type of cancer, regardless of tumor stage, treatments received, and the timing of the cancer. Despite the diversity of subpopulations, the difficulties described by patients and measured by tests are surprisingly similar (Demos-Davies et al., 2024). Cancer patients report cognitive symptoms affecting multiple domains, such as language, memory, concentration, attention, problem-solving, and prioritization (Hutchinson et al., 2012).

The linguistic difficulties described by patients are numerous. These include forgetting of words, which may be replaced by others (Bolton & Isaacs, 2018; Player et al., 2014; Skoogh et al., 2012). Patients have trouble to finish sentences and often reverse the order of words (Skoogh et al., 2012; Von Ah et al., 2013). In conversational situations, they struggle to follow speech turns (i.e., the successive turns of speech from different speakers within a conversation, [Sacks et al., 1974]) and to interact with a group (Bolton & Isaacs, 2018; Skoogh et al., 2012). More rarely, patients complain of forgetting the spelling of some words (Bolton & Isaacs, 2018).

The "tip-of-the-tongue" phenomenon, or anomia, is one of the most frequently described symptoms of CRCI by patients. The prevalence of this phenomenon ranges from one-third (Von Ah et al., 2013) to all patients (Myers, 2012). This variability could be explained by the different stages of care during which patients are questioned about their cognitive difficulties. For instance, Downie et al. (2006) conducted interviews with breast cancer patients during the curative phase of treatment and found that 78% of the twenty-one patients reported difficulties in naming words (Downie et al., 2006). Myers (2012), during interviews with patients six to twelve months after the curative phase, observed that all patients mentioned difficulties in retrieving words. However, the percentage of reported linguistic difficulties seems to decrease with the time elapsed after treatment completion (Myers, 2012). In a study by Von Ah et al. (2013), conducted with women one to twelve years after the curative phase, 36% of them continued to report lexical naming problems (Von Ah et al., 2013). This phenomenon is far from being exclusive to women with breast cancer. Among testicular cancer survivors, on average eleven years after diagnosis, 16% report "saying similar but incorrect words" at least once a week, and between 17% and 36% report difficulties in "finding words" or "encountering other language problems" (Skoogh et al., 2012, p. 191).

Memory dysfunctions also represent a significant source of concern for oncology patients, who experience difficulties remembering names (Bolton & Isaacs, 2018; Von Ah et al., 2013), familiar faces (Bolton & Isaacs, 2018), and important dates (Von Ah et al., 2013). Attention is also affected. Patients report troublesome concentration problems, such as being unable to focus on a task, becoming easily distracted, watching a movie without remembering the plot or the characters names (Bolton & Isaacs, 2018), or having to read the same page of a book multiple times without understanding its meaning (Player et al., 2014; Von Ah et al., 2013). Some patients report not remembering what they have been told (Tenda et al., 2022), repeating themselves, or forgetting the goal of a task they are currently performing (Von Ah et al., 2013). Regarding executive functions,

patients report difficulties remembering routine tasks, such as feeding animals or following a familiar route (Bolton & Isaacs, 2018; Player et al., 2014; Von Ah et al., 2013), struggling to switch between two simultaneous tasks (Bolton & Isaacs, 2018; Von Ah et al., 2013), and being unable to process new information or acquire new skills.

Subjective Assessment of CRCI

Self-administered questionnaires are the most common method to assess cognitive difficulties in cancer patients. It has the advantage of being close to the daily lives of patients. Among the existing questionnaires, four (i.e., FACT-Cog, EORTC-QLQ-C30, PAOFI, AFI) have been designed for oncology patients and are most often used in clinical studies.

FACT-Cog

The Functional Assessment of Cancer Therapy – Cognitive Function questionnaire (FACT-Cog v3 [Costa et al., 2018; Wagner et al., 2009]) was specifically developed to detect the presence of cognitive complaints in an oncology population. It is considered the reference tool for detecting cognitive difficulties in cancer patients (Henneghan et al., 2021). Originally written in English, it has been translated into French (Joly et al., 2012) and validated on a normal French population (Lange et al., 2016). Its current validated version comprises four subscales (Perceived Cognitive Impairments; Comments from Others; Perceived Cognitive Abilities; Impact on Quality of Life) and evaluates six cognitive dimensions (memory, verbal acuity, concentration/attention, mental acuity, functional interference, and multitasking) using thirty-seven questions. Four of these items are dedicated to language.

Costa et al. (2018) examined the responses from three groups of participants (healthy young adults, healthy older adults, and oncology patients) based on the cognitive domains included in the questionnaire. The patients' scores showed deficits across all cognitive domains, while those of the young adults and older adults were lower in only certain domains. This result suggests that the entire set of items in the FACT-Cog should be analyzed as a whole. In other words, the structure of the questionnaire does not allow for the identification of the most affected cognitive domains. Therefore, although four of the FACT-Cog items are dedicated to verbal abilities, their isolated score is not able to detect potential language difficulties. Furthermore, these items do not provide a precise understanding of the nature of the language complaint.

EORTC QLQ-C30

The QLQ-C30 questionnaire developed by the European Organisation for Research and Treatment of Cancer (EORTC) is a self-administered questionnaire designed to assess the quality of life of oncology patients (Aaronson et al., 1993). It includes five functional scales (physical, role, cognitive, emotional, and social) and three symptom scales (fatigue, pain, and nausea). The modularity of this questionnaire allows for a comprehensive view of the patients' quality of life. It is worth noting, however, that only two items out of thirty explicitly refer to cognition.

Patient's Assessment of Own Functioning Inventory

The Patient's Assessment of Own Functioning Inventory (PAOFI) is a questionnaire designed to assess cognitive difficulties perceived by patients, regardless of etiology (Chelune et al., 1986), and has been adapted for oncology patients (Van Dyk et al., 2016). Its reliability and validity have been demonstrated in women with breast cancer even before the initiation of treatment (Bell et al., 2013). It contains open-ended questions grouped by cognitive domains (memory, language, sensorimotor perception, higher-level cognitive functions) corresponding to the patients' complaints. Items are assessed using a 6-point Likert scale. Six of them focus on language, including two on anomia. Unfortunately, this questionnaire does not have a validated French translation.

Attentional Function Index

The Attentional Function Index (Cimprich et al., 2011) is a self-administered questionnaire designed to assess perceived efficiency during tasks that rely on attention and working memory in breast cancer survivors. It has the advantage of being quick to perform and of asking about the impact of attention difficulties on daily life. However, it does not take into account one of the most frequently reported complaints by these patients, namely word-finding difficulties. Additionally, this questionnaire does not have a validated French translation.

Neuropsychological Assessment

Logically, the most commonly tested cognitive domains tested are those that reflect the patients' complaints. These include memory, verbal abilities, and attention (Ahles et al., 2010; Kohler et al., 2020). The tests generally used in CRCI evaluation are the California Verbal Learning Test (Delis et al., 1987), the Hopkins Verbal Learning Test - Revised (Benedict et al., 1991), the Rey Auditory Verbal Learning Test (Rey, 1958), the Rey-Osterrieth Complex Figure (Osterrieth, 1944; Rey, 1941), the Wechsler Memory Scale (Wechsler, 1945), and the Trail Making Test (Armitage, 1946) (see Jung et al. [2023] for a review).

The literature on CRCI assessment reports inconsistent results regarding the affected cognitive domains. For instance, a decrease in verbal and visuospatial abilities after chemotherapy treatment has been observed in some studies (Jim et al., 2012), while others showed more mixed results (Demos-Davies et al., 2024). The discrepancy between results mainly arises from the assessment methods, which vary from one study to another. To help standardize the results, the International Cancer and Cognition Task Force and the *Groupe de Réflexion sur les Évaluations Cognitives en Oncologie* have proposed test batteries to aid in the diagnosis of CRCI.

ICCTF

The International Cancer and Cognition Task Force (ICCTF) is a multidisciplinary working group bringing together researchers, clinicians, and patients to address issues related to CRCI. In 2011, the group issued recommendations for the evaluation of CRCI (Wefel et al., 2011), and advised on using the Hopkins Verbal Learning Test-Revised, the Trail Making Test, and the Controlled Oral Word Association from the Multilingual Aphasia Examination (Benton et al., 1994). These tests measure the main cognitive abilities affected by CRCI, namely learning, memory, attention, processing speed, and executive functions (Demos-Davies et al., 2024). To compensate for the low sensitivity of these tests in detecting CRCI, the group redefined and adapted their pathological thresholds for the oncology population. Thus, for CRCI to be diagnosed, a patient must score below -1.5 standard deviations from the normative mean in at least two of the mentioned tests, or below -2 standard deviations in a single test. This battery has the advantage of being easy and quick to administer. However, it does not include questionnaires, as the ICCTF determined that subjective assessments (or patient-reported outcomes) would be more likely to be associated with confounding factors, such as fatigue, rather than neuropsychological tests. The scores of quality of life questionnaires and psychological measurement scales (e.g., anxiety and depression) have been found to be strongly correlated, but there was no correspondence between them and the scores from neuropsychological tests (Hermelink et al., 2010).

GREC-Onco

In France, the *Groupe de Réflexion sur les Évaluations Cognitives en Oncologie* workgroup (GREC-Onco) proposed a more comprehensive battery of tests, including cognitive screening (fNART, Mattis, MoCA) and a more in-depth examination of memory (RL-RI 16, HVLT, digit span forward and backward, Rey figure copy at 3 minutes, visual span forward and backward), language (BNT and Token Test), visuoconstructive functions (Rey figure copy), executive functions, and attention (Stroop, TMT A/B, phonemic and categorical verbal fluencies, WAIS coding). These tests were supplemented by behavioral assessments, as well as questionnaires for anxiety, depression, fatigue, and quality of life (Taillia et al., 2015). Rather than attempting to establish threshold scores, the objective of this battery is to detect cognitive changes in patients over time. However, the multitude of tests is time-consuming and may lead to the risk of overdiagnosis (Rick et al., 2024). Furthermore, it does

not address the gap between subjective complaints and test results.

Discrepancy between perceived impairment and neuropsychological tests

To date, the objective assessment of CRCI remains a clinical challenge. It is a "subtle" disorder in that it is difficult to detect through psychometric tests, even though its impact on patients' lives is tangible. Patients' scores often fall within the normal range on neuropsychological evaluations (Stewart et al., 2008). The lack of suitable tests results in a discrepancy between the difficulties described by patients and their objective assessment. This gap can lead to misdirected therapy and management strategies. In a meta-analysis, Whittaker et al. (2022) report that one-third of patients who received chemotherapy experience a significant cognitive decline, as confirmed by neuropsychological tools. However, this number is much higher in subjective assessments. Moreover, the inconsistency of results between studies complicates the precise identification of CRCI. For instance, in a 2013 study, Ganz et al. (2013) highlighted the correlation between subjective memory difficulties and a decline in verbal abilities among breast cancer patients. The same team did not observe any cognitive performance differences between assessments made before and twelve months after the completion of first-line treatments (chemotherapy and radiotherapy) (Ganz et al., 2013). Yet, the patients in this cohort reported difficulties with language and communication (Ganz et al., 2014). More recently, Veal et al. (2023) examined the relationship between patient-reported difficulties and their neuropsychological outcomes. The scores from the FACT-Cog questionnaire were not correlated with those from processing speed, visuospatial processing, and verbal memory tests. However, the description of memory lapses (via a provided list of symptoms) did correspond to the scores obtained from neuropsychological tests.

The discrepancy between responses to the questionnaire and the scores could be explained by several factors. First, neuropsychological tests have generally been developed and validated for populations with more severe neurological disorders, such as those with brain injuries. Although pathological thresholds have been redefined, these tests struggle to reveal cognitive impairments that are subtler in nature. Another explanation lies in the fact that patients are assessed under optimal clinical conditions, i.e., in an interview with a clinician in a calm and controlled environment. It is therefore possible that patients with CRCI manage to compensate for their cognitive difficulties, albeit with effort, in these ideal conditions and over a relatively short period. As a result, these tests fail to fully capture the functional impact of CRCI on patients' daily lives, where stimuli are multimodal and less controlled. Additionally, they do not account for confounding factors such as fatigue or anxiety, which can introduce variability into the results and cause inconsistencies in patient follow-up. Finally, these psychometric exams are designed to assess a single cognitive function in isolation, testing its efficiency independently of other components. However, this approach does not reflect the difficulties encountered by oncology patients. Therefore, it is likely that neuropsychological tests and questionnaires do not measure the same dimensions (Costa & Fardell, 2019). This hypothesis is supported by studies that have found an association between perceived cognitive difficulties

and neuroanatomical changes, without correlating the latter with neuropsychological test results. Questionnaires may thus be more sensitive to cognitive changes induced by everyday experiences.

In this respect, CRCI research should now develop more effective diagnostic tools. In addition to sensitivity and specificity criteria, these tools must also incorporate an ecological dimension, enabling an accurate assessment of patients' cognitive abilities in their daily environment. A holistic approach that considers the presence of co-factors is essential.

KEY SUMMARY — CANCER-RELATED COGNITIVE IMPAIRMENT

1. ***Cancer-Related Cognitive Impairment***

   - Cancers and their treatments can induce cognitive changes. These difficulties are subtle but nonetheless disabling. They decrease the quality of life of patients who have had or currently have cancer.

   - Cancer itself is believed to cause brain inflammation that leads to these cognitive difficulties. Treatments such as chemotherapy, along with other confounding factors (e.g., fatigue, depression, etc.), exacerbate the expression of these difficulties.

   - The temporal aspect of this cognitive syndrome is broad: it can appear before treatment begins, is usually very high during and immediately after treatment, and can persist beyond the cessation of treatment. It is essential to identify those patients who are at risk of developing long-lasting symptoms.

2. **Discrepancy Between Reported Difficulties and Objectified Performance**

   - Oncology patients report difficulties with memory problems, a reduced attention span, an inability to multitask, and persistent trouble finding the right words.

   - This set of difficulties is easily highlighted by questionnaires. However, current neuropsychological tests struggle to reveal the extent of the deficits. This gap between subjective and objective evaluations prevents some patients from accessing appropriate care. Moreover, the lack of a reliable psychometric tool slows down the advancement of knowledge about the functioning of CRCI.

   - The development of new approaches is necessary to identify oncology patients suffering from cancer-induced cognitive impairment. These methods will also shed light on the underlying mechanisms of CRCI.

4 Speech: A Gateway into Cognition

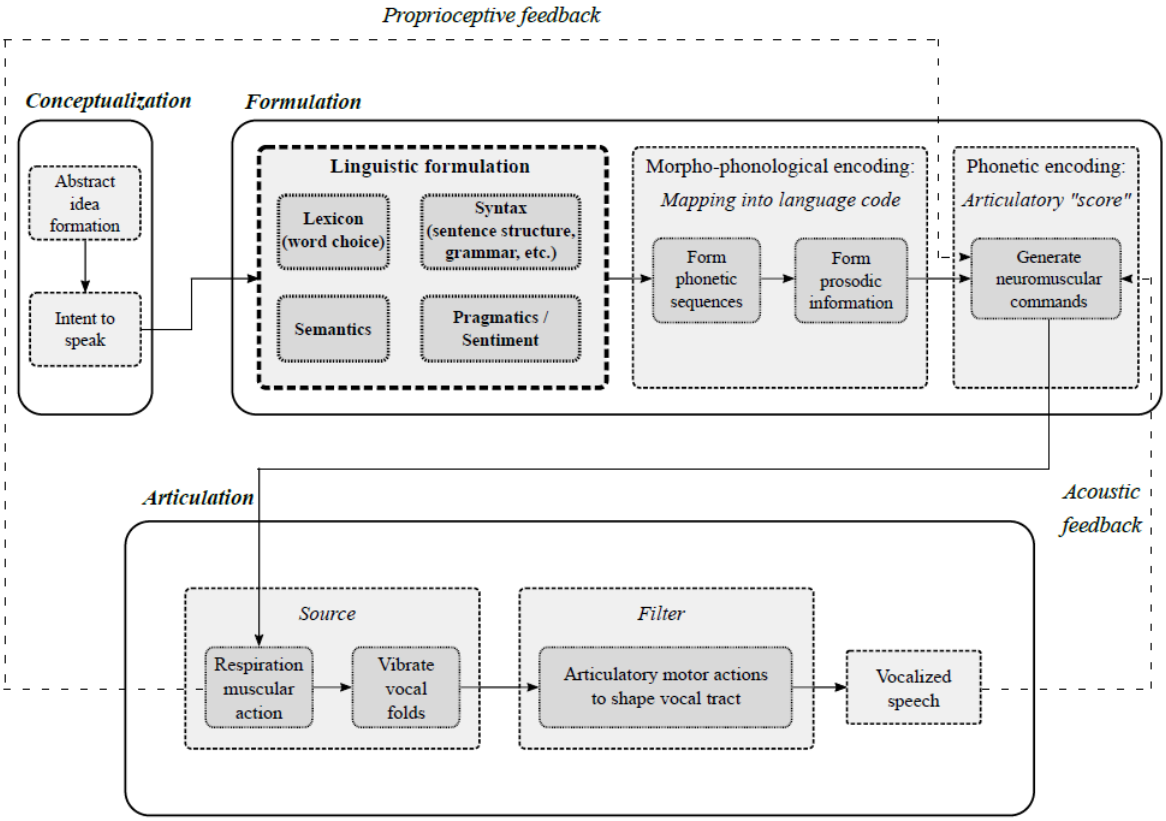
Speech, the audible production of language, is a faculty as rich as it is complex, and is used during daily interactions. It is expressed in the form of an acoustic signal that can be interpreted through segmental units, i.e., the phonemes of a linguistic system, and through supra-segmental elements, i.e., prosodic features that link a set of phonemes together. Prosody refers to three perceptible parameters of speech: pitch, volume, and rhythm, which are related to frequency in Hz, intensity in dB, and rate generally measured in syllables per second (De Jong and Wempe [2009], but also see Tilsen and Tiede [2023] for finer-grained measurements). It conveys structural information on the oral message, as well as the speaker emotions and intentions, whether real or simulated. Prosody thus stands at the interface of other linguistic domains, such as semantics (Ben-David et al., 2016) or syntax (Degano et al., 2024).

From Thought to Words: Models of Speech Production

For over a century, theories of speech production have contributed to the understanding of human language functioning. The Levelt's serial model (Levelt, 1983) was the first to include three main stages in speech production: the conceptualization of the message in thought, its mental formulation, and then its articulation. In contrast, the connectionist model proposed by Dell (1986) emphasizes multidirectional interactions between semantics, lexical selection, and phonology. This dynamic view of language helps explain phonemic and semantic errors. Today, theories of language have been enriched and refined through insights from behavioral neuroscience (see for instance Kemmerer [2022] for a comprehensive review on the topic).

The production of an oral message is generated by a sequence of linguistic and non-linguistic processes, sometimes simultaneous, that interact with each other. It involves retrieving concepts from semantic memory and then selecting lexical items from the available stock (Hagoort, 2017). The chosen lexical items are then subject to phonological encoding, i.e., the association of a concept with the sequence of phonemes assigned to it, followed by phonetic encoding, i.e., the transformation of speech sounds into articulatory motor gestures (Hagoort, 2017). Finally, during phonation, pulmonary activity coordinates with the mobility of the laryngeal and articulatory structures. Almost simultaneously, the speaker's own voice provides auditory feedback, enabling the detection of production errors. Figure 4, from Voleti et al. (2020), illustrates the linguistic processes involved in speech production, from the conceptualization of an idea to its vocal realization.

Figure 4. Diagram of the linguistic processes involved in speech production.



Note: Reprinted with permission from Voleti et al. (2020). 2019 IEEE.

Neural Substrates of Language

It is now indisputable that language is not limited to an isolated function confined to two brain regions, as suggested by neurologists in past centuries. The Broca-Wernicke model, along with its variants, has been largely surpassed, particularly due to its lack of spatial precision and its restricted localization (Tremblay & Dick, 2016). In contrast, language, with speech being its audible manifestation, interacts in a complex manner with other cognitive functions such as memory, executive functions, attentional processes, and motor functions. This assumption is supported by studies using fMRI, which have shown that language activates brain regions that are not exclusively associated with linguistic functions (Duffau, 2003). For instance, the programming of a message for articulation engages motor functions (Tate et al., 2014). Similarly, lexical selection has been found to activate the supplementary motor area. More specifically, the motor and executive control of language is thought to occur via the oblique frontal fasciculus, which connects the inferior lateral frontal gyrus to the supplementary motor area (Tremblay & Gracco, 2010).

The Semantic System and Lexical Selection

Lexical semantics refers to linguistic forms (a sequence of phonemes in the case of speech) arbitrarily associated with conceptual information (Reilly et al., 2025). The semantic system, an essential component of long-term memory, is thought to be responsible for storing knowledge related to concepts. However, several studies challenge this hypothesis: are the knowledge representations actually localized in a single brain region, or does their localization vary depending on the nature of the knowledge itself? It appears that concepts related to actions are more associated with frontal regions, while the lexical and semantic processing of objects primarily involves the lower temporal regions of the left hemisphere (Antonucci & Reilly, 2008). More recently, Borghesani et al. (2016) explored the multidimensional representation of concrete words, i.e., concepts with a perceptible physical referent. In an fMRI experiment, participants were asked to assess the degree of similarity between two written words. The primary distinction between these words lay in the actual size of the objects they referenced (e.g., "mouse" vs. "giraffe"). During silent reading, the ventral pathway of the visual cortex was activated first, followed by the activation of anterior temporal regions, which are involved in processing more abstract semantic information (Borghesani et al., 2016). In other words, the semantic processing of written words initially occurred perceptually through visual regions, regardless of word length, before involving the anterior temporal regions for more conceptual processing.

Lexical selection during message formulation is undoubtedly one of the most studied processes in the field of psycholinguistics. The work of Levelt and his team (notably Levelt et al. [1999]) introduced the concept of lexical competition, which refers to the activation of a semantic network of competing lemmas by a lexeme (i.e., the root of a "word"). The selection of the appropriate lemma (i.e., a semantic unit forming a "word") implies the existence of a unified conceptual representation for each word in the lexicon (or lexical concept). These representations would be organized within a semantic network, and the simultaneous activation of alternative lemmas would generate competition, thereby slowing down the selection process. In other words, the more semantic neighbors a lexical concept has, the longer its selection process would be.

However, many subsequent studies challenge the validity of the lexical competition selection theory. Through a series of experiments, Mahon et al. (2007) demonstrate that the semantic neighbors of a target word (e.g., "horse" and "zebra") accelerate the recognition of that word, in contrast to semantically distant words (e.g., "horse" and "whale"). Thus, semantic interference does not appear to occur at the level of lexical retrieval. Additionally, the analysis of simultaneous eye movements with speech production also provides evidence against the idea of lexical competition (Mahon et al., 2007). Huettig and Hartsuiker (2008) conducted an experiment where participants had to name a target object among a semantic distractor (co-hyponym), a visual distractor, and a distractor unrelated to the target. The researchers observed that participants fixated more the semantic and visual distractors than the unrelated distractors, which highlights the involvement of visual processing in word production. Furthermore, no additional latency was observed in the presence of visual or semantic distractors,

suggesting the absence of semantic interference in the lexical selection process (Huettig & Hartsuiker, 2008).

Semantic-lexical retrieval refers to the access of specific language-related information stored in memory. The retrieval of a concept would engage several brain regions, such as the left inferior frontal cortex, the left superior and middle temporal cortices, as well as the left inferior parietal cortex (Hagoort, 2017). This activation of multiple brain areas reflects the interaction of linguistic functions with other cognitive processes. Consequently, accessing the meaning of words would involve not only their storage in semantic memory but also the executive control of semantic knowledge (Antonucci & Reilly, 2008; Thompson et al., 2018).

Prosodic Aspects

Research on the neural bases of prosody has primarily focused on its pragmatic functions, particularly in relation to the expression of emotions. For this reason, it is frequently associated with the right hemisphere. For instance, a study by Sammler et al. (2015) examined the brain regions involved in the perception of intonation. Participants listened to isolated words produced with either an interrogative or declarative tone, in order to limit the activation of regions associated with syntactic processing. fMRI analyses revealed that the processing of intonation activates both the dorsal and ventral pathways of the right hemisphere, as well as the premotor cortex. However, lateralization to the right hemisphere for prosody is challenged by other studies suggesting an interhemispheric distribution instead. Some patients with brain lesions in the right hemisphere are still able to perform prosody identification tasks (Baum & Pell, 1999). The left hemisphere thus seems to be involved in controlling duration as a prosodic cue (Baum & Pell, 1999). The divergence between these studies may lie in the fact that the first study equates prosody with intonation, while the second also includes the parameter of duration. However, more research is needed to better understand the underlying processes involved in the production and perception of linguistic prosody, particularly within stress languages, where prosodic parameters play a structural role in conveying the message (Grice & Kügler, 2021).

Language is therefore a composite cognitive function that, depending on the speaker's needs, engages various aspects of cognition and recruits different brain regions and neural pathways at varying levels of involvement. While knowledge of language functioning is made possible through insights into cognition, the reverse is also true. From this point of view, the study of speech, as an audible manifestation of brain activity, could provide insights into the processes it engages.

Anomia: Understanding Cognition Through Error

Verbal production is often punctuated by semantic, phonetic, or morphosyntactic errors, i.e., the involuntary production of lemmas, phonemes, or morphemes that do not align with the expected context by the speaker or the listener. Their frequency and nature depend on various factors, such as the speaker's abilities, the communicative context, and the specific characteristics of the discourse genre. A rigorous analysis of these linguistic errors allows one to distinguish between those that fall within normal variation and those that indicate a pathology. Given the interaction of language with other cognitive functions, speech impairment is often the first sign of an underlying cognitive change.

The study of semantic and lexical errors has long attracted particular interest, especially in order to elucidate the mechanisms of language functioning. The tip-of-the-tongue phenomenon, or anomia, refers to a disturbance in lexical production, primarily affecting lexical items from open-class categories. The expression "having the word on the tip of one's tongue" embodies a commonly experienced phenomenon, whose intensity is often amplified by factors such as fatigue and stress. However, it is considered pathological when it occurs recurrently over a long period and causes discomfort to the speaker. It is therefore not surprising that anomia is an ordinary symptom in various cognitive disorders and syndromes. It is emblematic of non-fluent aphasias which results from a stroke or Alzheimer's Disease for instance. However, these types of anomia have distinct physiopathological origins, and their in-depth study provides valuable insight into the underlying cognitive mechanisms that are disrupted.

In the 1990s, Caramazza and Hillis (1990) proposed a theoretical model of semantic error production based on the analysis of clinical cases of aphasia. According to this model, language errors in spoken production may result from either a disruption in the semantic knowledge of a word or a dysfunction in the process of accessing its phonological form. In cases of semantic deficit, the information associated with the concept is "lost," leading to systematic errors: the individual consistently produces the same type of error and, unknowingly, confuses one word with another whose visual or phonological form is similar. This phenomenon is often accompanied by the impression of an extended delay between lexical selection and oral production, although, in reality, there is no true lexical selection, and the expected word is not produced. On the other hand, when anomia is induced by an executive deficit, the semantic knowledge related to the word is preserved, but the speaker experiences difficulty accessing its phonological form. The individual aware of their difficulties may then deliberately choose a word with a similar meaning, similar form, resort to a description or a synonym that fits relatively well within the context. The errors observed in this case are more variable, and the sought-after word may reappear after some time. An increased delay between lexical selection and oral production is also noted (Caramazza & Hillis, 1990, 1991). Research has furthermore highlighted the impact of age on anomia, particularly for words with low lexical frequency and proper nouns (Burke et al., 1991).

The examination of semantic errors in patients with Alzheimer's Disease and semantic dementia

helps to distinguish between these two pathologies (Reilly et al., 2011). Their symptoms seem similar, although their etiologies differ. Alzheimer's Disease is a neurodegenerative pathology that primarily affects episodic memory, i.e., the memory of events and lived experiences. Semantic dementia is often a precursor of fronto-temporal lobar degeneration and affects semantic memory, i.e., the knowledge of concepts. When these populations are compared, anomia is significantly more severe in cases of semantic dementia. This anomia is characterized by a less efficient ability to in lexical retrieval and a higher number of errors than in Alzheimer's Disease (Reilly et al., 2011). Additionally, patients with Alzheimer's Disease make more errors related to co-hyponym words (e.g., "lemon" and "strawberry" are two co-hyponyms of "fruit"), while those with semantic dementia tend to use more circumlocutions to describe the function of an object (Reilly et al., 2011). This study highlights that the neural bases of anomia vary depending on the underlying etiology. In the case of a cognitive syndrome whose precise etiology has not yet been diagnosed, analyzing anomia helps refine hypotheses about the underlying cognitive mechanisms. Regarding cancer patients, neuroanatomical data do not reveal brain lesions that could cause anomia (Bernstein et al., 2021). Since tip-of-the-tongue phenomena are commonly reported by this population, the relationship between CRCI and word-finding difficulties remains to be explored.

Lexical Interference on Prosody

When articulating a message, the time it takes to retrieve a concept or its phonemic form from memory may translate into one or more breaks in the spoken chain (Don & Lickley, 2015; Hartsuiker & Notebaert, 2010). These breaks directly affect the rhythm of speech, that is, the temporal "regularity" of the accents it carries (Astésano, 2016; Tilsen, 2019). Since the 1950s, numerous studies have aimed at understanding the concomitant mechanisms with lexical selection. For Goldman-Eisler (1958), for instance, choosing one term over others that share similar semantic properties would result in the production of a filled pause. Other studies suggest that the search for rare words is frequently associated with hesitations, which manifest as silent pauses (Beattie & Butterworth, 1979). A study on semantic interference, i.e., the naming delay caused by the presence of a semantic distractor, shows that lexical access difficulties in a healthy speaker generate more disfluencies, such as repairs or silent pauses, without increasing the number of semantic errors (Rapoeye et al., 2023). However, the results of this study rely on manipulated speech data, limiting its ecological validity. Similarly, the presentation of stimuli differed between the target, which was a written word form, and the distractor, which was an image. Therefore, it is also possible that the disfluencies observed are partially due to the visual processing of the stimuli rather than semantic interference.

These studies like many recent ones highlight the influence of lexical search on prosody. However, these studies lack of consistency in defining and measuring interruptions. The concepts of disfluencies, pauses, and hesitations remain vague within the literature. This point will be explored in detail in the next section.

Interruptions in the Speech Flow

in contrast to words in isolation, connected speech is characterized by a succession of adjacent phonemes perceived and processed as a regular sequence (Ash & Grossman, 2015). However, speech is not inherently rhythmic. The acoustic signal remains indeed an aperiodic signal (Turk & Shattuck-Hufnagel, 2013). Nevertheless, it exhibits "temporal regularities" that are perceived by the human ear as "quasi-rhythmic" (Poeppel & Assaneo, 2020, p.323), and processed similarly by the brain. It is worth noting that speech production is rarely linear. It consists of rhythmic and intonation elements that contribute to its structure to varying degrees. The term "oral message" (here considered as a sequence of linguistic features, regardless of the scale of the linguistic unit — phoneme, syllable, etc.) replaces the more abstract term "utterance" used commonly in this context. The oral message is made up of prosodic segments, i.e., rhythmic and intonational segments, whose lengths depend on the speaker's working memory and articulatory motor skills. These segments play a crucial role in the listener's understanding of the message. By segmenting syllabic units at lexical and syntactic levels, prosodic elements facilitate the semantic processing of the message.

Disfluencies

Some prosodic elements mark breaks in the flow of speech, thereby signaling a perceptible change in rhythm and intonation. A proposed classification distinguishes two types of interruptions: suspensive interruptions, which delay the progression of the message, and disfluent interruptions, which break the message and may potentially initiate a new one (Pallaud et al., 2019). In the latter case, the break in the message is interpreted by the listener as a sign of incompleteness.

Disfluencies have been widely studied, but there is no clear consensus regarding their terminology. Traditionally, the term "disfluency" encompasses filled pauses, silent pauses, sustained vowels and consonants, as well as false starts, repetitions, uncompleted phrases, and repairs (Pallaud et al., 2019). However, the very definition of disfluency varies significantly across authors. In summary, a disfluency is the interpretation of a break in the verbal flow perceived as unexpected, either by the listener (Lickley & Bard, 1998) or by the speaker themselves, who might then hear a break in their phonological loop (i.e., the auditory feedback of their own speech) (Huettig & Hartsuiker, 2010). Disfluencies are deemed unpredictable by nature, although certain regularities can be observed. In other words, disfluencies disturb the expected rhythmicity of speech. Disfluencies are considered "atypical" when they occur within the framework of speech sound disorders, such as stuttering (Lu et al., 2022). In this condition, they exhibit prosodic and phonetic specificities linked to the articulatory realization of phonemes. In contrast, "typical" disfluencies occur accross all speakers and are generally associated with high-level cognitive processes, such as lexical selection or syntactic planning (Lickley, 2017). Their high frequency may indicate the presence of underlying cognitive difficulties. According to studies, the rate of disfluencies ranges from 6 to 10 disfluencies per 100 words in healthy speakers (Bortfeld et al., 2001; Shriberg, 2001). Bortfeld et al. (2001) analyzed

disfluencies (repetitions, false starts, and *fillers*) in a conversational situation elicited by a role-playing game. In their analysis, the authors omitted silent pauses, as well as modality statements (i.e., comments on the production of the message), since these were too rare in the corpus. When several types of disfluencies were adjacent, each occurrence was counted separately. A disfluency rate was calculated based on one hundred words, although the exact definition of what stood as a word was not specified. The study highlights an effect of age, gender, and familiarity with the topic on the production of disfluencies. However, this study has a major limitation. It relies on an orthographic transcription of the productions, even though speech is primarily an acoustic phenomenon.

In the context of objective measures of speech, the concept of "disfluency" could be problematic at several levels. On the one hand, disfluencies form a subjective category, and their observation heavily depends on their definition. The inter-rater variability linked to this taxonomy prevents the reproducibility of experiments. On the other hand, most of the research on disfluencies has been conducted on English data, and to a lesser extent on French. However, their frequency and duration are likely to vary depending on the linguistic systems in which they are produced (Crible et al., 2017). Furthermore, the content and form of a verbal message are context-dependent, in the sense that they depend on the interactional situation (Kachkovskaia & Kocharov, 2024). They are also specific to the speaker, depending on their linguistic habits and the psychological factors influencing them (e.g., Cummins et al. [2015]). Finally, the idea that disfluencies are distinguishable due to their unexpected nature is debatable, as some disfluencies are predictable. For instance, a speaker may anticipate an interruption in order to take their turn in the conversation. Therefore, rather than speaking of disfluencies, it seems more appropriate to refer to these prosodic breaks as interruptions in the speech flow, as they are measurable physical phenomena using instruments for acoustic analysis.

Pauses

Pauses hold a particular place among speech interruptions. As a label, "pauses" encompasses both silent pauses and filled pauses, i.e., voiced elements that are semantically empty (e.g., "uh").

The threshold for silent pauses remains widely debated in the speech literature. Is every silence a "pause"? Studies carried out by Goldman-Eisler (1958, 1961) were among the first to focus on the duration of pauses. Goldman-Eisler suggested that a minimum value of 250 ms reflects ongoing linguistic processes (Goldman-Eisler, 1961). However, this threshold is based on read speech, and the methodology used in this paper has been criticized retrospectively (Boomer, 1970). This research was nonetheless seminal for further studies on the link between silent pauses and underlying cognitive processes. As Lickley (2015) notes, this arbitrary threshold was likely linked to the technological limitations of the time for speech analysis. Since then, different thresholds have been tested in various syntactic slots (e.g., pauses between phrases, pauses within words, etc.). To date, the minimal threshold of 200 ms set by Boomer (1965) has been adopted by many subsequent studies. Similarly, Goldman et al. (2010) proposed that pauses longer than 200-250 ms signal cognitive

effort, whereas shorter pauses are more likely related to physiological constraints (i.e., swallowing, breathing, articulation). However, it seems difficult to separate mechanical activity from cognitive processes, leading other studies to argue against the need of a minimal threshold (Campione & Véronis, 2002). Nevertheless, the absence of a minimal threshold increases the risk of including pauses that are related to phonetic characteristics, such as consonant bursts (Campione & Véronis, 2002). A few studies have focused on a maximal threshold for pauses (e.g., Candea [2000] and Grosjean and Deschamps [1972]). Very long pauses are rare and mostly occur in spontaneous speech. Removing them from a corpus would alter the general distribution of pauses. This would risk making their interpretation unreliable (Campione & Véronis, 2002). Furthermore, these long pauses can be particularly interesting in pathological contexts.

Several studies suggest that silent pauses have a grammatical or pragmatic function (Crible et al., 2022). Some pauses may have a demarcative function when they occur between two clauses, and other pauses may be interpreted as a sign of hesitation (Campione & Véronis, 2005). A silent pause would be perceived as "disfluent" when it disrupts the syntax of a verbal segment (Pallaud et al., 2019). However, assigning specific functions to silences has limitations. On the one hand, the fact that each pause is given a specific function neglects the possibility that multiple functions may coexist simultaneously. A speaker can for instance have a hard time accessing the lexicon while producing a major prosodic boundary. In this situation, the speaker would struggle to find a concept or the associated phonological form that begins the phrase. The resulting pause would then be labeled as a demarcative pause, and not as an hesitation pause. On the other hand, the identification of disfluent pauses primarily relies on their position within a rhythmic group. While there is a strong connection between prosody and syntax, the former shows a degree of independence from the latter (Ferreira, 1993). Moreover, prosodic elements are temporal parameters which escape the spatial localization of syntax. Finally, it is worth noting that not every silence in speech is related to a cognitive pause. Silences regularly occur in fluent discourse, between turns of speech, within phrases, or even at the end of the production of a consonantal phoneme (Lickley, 2017). From a perceptual standpoint, silences participate to speech segmentation by highlighting the prominence of syntactic structures (Duez, 1985). In fact, there is no physical difference between these silences, except for their duration (Cole, 2015).

As far as filled pauses are concerned, they have mostly been described as having a pragmatic function. The speaker is thought to express the desire to hold their turn in the conversation when producing "uh" or "uhm" (Clark, 2002). Filled pauses are thus the voiced marker of cognitive effort (Clark, 2002; Simon & Christodoulides, 2016). They can also contribute to stylistic effects. Moreover, as fillers, filled pauses are assumed to facilitate the processing of the message. Corley and Hartsuiker (2011) compared reaction times during a lexical recognition task across three experiments. Target words were immediately preceded by a filled pause, a silence, or a tone. Participant's reaction time to detect the target word was similar across the three experiments. According to the authors, this result suggests that any delay regardless of its nature, aids in lexical recognition (Corley & Hartsuiker, 2011).

Finally, sustained or lengthened vowels are vowels whose duration exceeds the phonetic time required for their production. In French, lengthened vowels that carry stress become a perceptible boundary signal and are called final lengthenings when they occur at the end of a rhythmic group. They are distinguished by their specific intonation contour, i.e., variations in the fundamental frequency (f0) which can be rising, falling, or rising-falling. Similarly to filled pauses, sustained vowels are perceived as markers of hesitation when they feature a flat intonation contour (Belz, 2023).

To conclude this section, pauses are defined in this dissertation as interruptions in the flow of speech at an abstract level. Just like the perception of phonemes is continuous (McMurray, 2022), we propose that, from an acoustic standpoint, pauses form a continuum ranging from a sharp drop in the fundamental frequency to an absence of phonemic realization. Filled pauses therefore encompass isolated vowels perceived as fillers (e.g., "uh" or "um") or syllables in which the nucleus duration is extended. The intonation contour of these phonemes remains flat throughout their emission (Campione & Véronis, 2004). At the extreme end of the interruption spectrum, silent pauses refer to situations where no phoneme emission occurs. This situation spans from audible inhalation to total silence, including mouth and pharyngeal sounds (Belz & Trouvain, 2019).

Limits of Speech Analysis

Lexical search, and by extension word-finding difficulties, not only slows down speech rate, but also leads to a drop in syllable pitch height and intensity for those produced during lexical search. Examining the acoustic signal allows generating hypotheses regarding the cognitive correlates of speech interruptions. However, speech analysis has several limitations.

First, prosody is subject to considerable inter- and intra-individual variation. Prosodic production is modulated by language systems (e.g., the differences between tonal language and non-tonal languages) to sociolinguistic variables (e.g., a speaker's membership of social and ethnic groups, Holliday [2021]), as well as physiological aspects (gender, age, integrity of oro-laryngeal structures) modulate the production of prosody. These variations are all the more apparent in connected speech, whose dynamic nature is difficult to manipulate in the laboratory.

Second, studies on the topic are numerous, but few of them are based on empirical acoustic data. These theoretical studies are difficult to generalize to clinical populations. Further, the analysis of language production requires additional investigative techniques such as eye-tracking to be more accurate. These techniques can validate or invalidate observations of prosodic phenomena. Combining speech and additional behavioral techniques offer a deeper insight into the cognitive mechanisms underlying speech interruptions.

KEY SUMMARY — SPEECH: A GATEWAY INTO COGNITION

1. **Production of Verbal Messages**

   - Speech production is a dynamic function of cognition. It involves a cascade of linguistic and non-linguistic processes, such as lexical search, phonological encoding, and so on. These processes are necessary for the conceptualization, formulation and articulation of a verbal message and recruit cerebral fasciculi beyond the regions associated with language.

   - Accessing the meaning of words requires their storage in semantic memory, as well as the executive control of knowledge. A deficit in one of these mechanisms can lead to lexical search difficulties (i.e., anomia).

   - Impaired speech is often the first indication of underlying cognitive changes. Speech analysis is a relevant tool for identifying and understanding subtle cognitive impairment. In the case of cancer survivors, anomia is the prevailing reported symptom of CRCI.

2. **Prosodic Parameters of Cognitive Change**

   - Anomia affects speech temporality. Prosodic variations such as longer pause are usually observed.

   - Pauses are speech interruptions that are often examined to investigate cognition. They form a continuum ranging from a sudden drop in fundamental frequency (sustained vowels and filled pauses) to an absence of phonemic realization (silent pause).

   - The multiplicity of speech parameters used in research makes it difficult to select potential CRCI markers. An inventory of existing parameters is necessary to select the most likely subset of pros cues to identify CRCI.

5 Investigating Linguistic Processes Through Eye Behavior

Vision is the most developed sensory system in humans (Leigh & Zee, 2015). The study of oculomotor behavior offers valuable insights into the functioning of cognitive processes. It sheds light on the links between certain neurological pathologies and eye movement abnormalities, such as nystagmus or Rapid Eye Movements, which are thought to be directly associated with biophysical

properties (Leigh & Zee, 2015). Some types of eye movement are specific to neurodegenerative pathologies, and their analysis helps in the diagnosis of these conditions (Sekar et al., 2024). From a cognitive point of view, oculomotor analysis provides information on processes such as decision-making (Orquin & Mueller Loose, 2013) and multitasking (Huestegge, 2011). In psycholinguistics, eye-tracking is particularly useful for deciphering the cascade of cognitive operations that are activated in verbal production (Griffin & Spieler, 2006).

Fundamentals of Oculomotor Behavior

Visual search, which mainly involves the frontal and parietal lobes, aims to minimize attentional demand while optimizing the acquisition of new foveal information. During this search, the fovea, i.e., the part of the retina where vision is most precise, must move sequentially to explore points of interest that appear in the periphery. There are two main ocular events: saccades and fixations (Mahanama et al., 2022).

Saccades are rapid movements of the eyeball. Voluntary saccades are Internally motivated and are intended to explore the visual environment. They direct the fovea towards a target without head movement, in response to a visual or auditory stimulus in the periphery. These movements are rapid and precise, so as not to interfere with visual processing. In general, there is a latency of around 200 ms between the appearance of the stimulus and the execution of the saccade. This delay corresponds to the neuronal processing time involving the retinal cerebral cortex, superior colliculus, basal ganglia, thalamus and cerebellum. Saccades are accompanied by corrective phases that adjust the position of the eye when an eye movement underestimates or exceeds the ideal amplitude for reaching the target (Mahanama et al., 2022).

Fixations, on the other hand, are made up of microsaccades in a delimited region of interest for a period long enough to process visual information. The fovea is never completely stable, especially when fixing small targets. Each fixation of the fovea delivers a tiny fraction of the information, so that details about the visual world are acquired through a succession of fixations griffinObserving2006. However, although its resolution is less precise, parafoveal processing also extracts the information needed to direct visual attention towards a new target. Fixation of an object is therefore not always essential for visual processing. On the other hand, the probability of fixating on a target increases when its quality is degraded or when its level of detail is high (Griffin & Spieler, 2006). Though, this effect is not found in speech production. In a study of the impact of degraded visual object quality on the production of disfluencies, Pistono and Hartsuiker (2023) not only found that blurred images did not lead to an increase in disfluencies, but that the time spent exploring the image decreased. The authors suggest that impeding access to the visual representation of an object does not slow down its identification or the naming process. The idea that poor visual quality slows down oral production is therefore called into question. What's more, even in simple contexts, individuals tend to fixate on a target during cognitively demanding tasks, such as producing a verbal message. The time spent

staring at the target would then reflect the processing time of this task (Griffin & Bock, 2000).

Interaction Between Speech and Visual Search

One of the first studies to explore the real-time relationship between language and cognitive functions using eye-tracking was that carried out by R. M. Cooper in 1974 (Cooper, 1974). This work gave rise to the "paradigm of the visual world", which suggests that when listeners hear a verbal message, they tend to direct his gaze towards objects sharing a visual similarity, in shape or color for instance, with the imageable items contained in that message (Eberhard et al., 1995). Huettig and Altmann (2005) have observed, e.g., that when listening to the word "snake", listeners turns their visual attention to the image of a cable (Huettig & Altmann, 2005).

Incremental Vision of Speech

As far as speech production is concerned, the visual world paradigm is commonly used to analyze lexical selection mechanisms. The high temporal and spatial resolution of eye-tracking makes it possible to study speech as it unfolds. Thus, the time spent looking at an informative element of an image would reflect the stages of lexical retrieval and phonological encoding of words (Griffin & Bock, 2000). Speaking would thus be an incremental process, i.e. words would be selected and phonologically encoded one after the other.

In unprepared speech, a speaker begins to articulate the first words of a message even before completing the phonological encoding of the following words (Ferreira & Swets, 2002). Recall of a lexical item is facilitated by its previous activation (Griffin & Spieler, 2006). Similarly, the frequency of a word, but not its length, affects the duration of its selection. In other words, the less frequent a word is, the longer it will take to retrieve it from memory. Meyer et al. (1998) compared the effect of image integrity (complete drawings vs. drawings with missing strokes) and the effect of lexical frequency on the naming delay of imageable lexical items. The poor quality of the images did not induce a longer naming time than that of complete images. Furthermore, latency was higher (between 750 and 800 ms) when participants were asked to name the object, than when they were simply asked to identify it without naming it (around 650 ms). The duration was increased by a lexical effect in the naming task, whereas this effect did not occur in the identification task. The authors conclude that the lexical frequency effect occurs mainly when accessing object names (Meyer et al., 1998).

In addition to lexical retrieval, syntactic planning can also be revealed by the order in which items are fixed during message production (Griffin & Spieler, 2006). Items are generally viewed in the order of their appearance in speech (Griffin & Bock, 2000). For instance, on presentation of an image that depicts an agent performing an action on a patient, a speaker will gaze first at the agent and then at the patient. Back and forth between elements could suggest that speakers have not yet fully decided

on the structure of their message (Griffin, 2004).

In view of these elements, the incremental vision of speech explains the presence of interruptions in the speech stream even in fluent speech. The study of eye movements could therefore prove particularly informative in understanding the anomia of a population with no language disorders to speak of. The eye behavior concomitant with the search for a word in cancer patients could shed light on the underlying mechanisms of the latter.

Visual Perception in Search of Meaning

Vision is largely involved in the orientation of attention (Leigh & Zee, 2015). The fixation of a visual element enables the information associated with it to be extracted and processed. Two competing theories attempt to explain the orientation of attention when perceiving a visual scene. On the one hand, the image guidance theory states that the perception of a scene is primarily guided by the prominence of properties of an image or scene, such as colors or contrasts. Cognitive guidance theory, on the other hand, suggests that the perception of a scene is guided more by the visual regions that carry meaning, and by their level of information (e.g., fruit on a table compared with the foot of the table).

To test these theories, J. M. Henderson and Hayes (2017) and J. M. Henderson et al. (2018) developed a method for creating semantic maps from scores assigned to different image extracts. The team then compared a heatmap of "visual attention" representing participants' fixation points, with semantic and saliency maps. The Figure 5, taken from J. M. Henderson et al. (2018), represents (from left to right) an instance of a scene presented to participants with the visual fixation map, the semantic map as well as the saliency map corresponding to it. A histogram matching analysis indicates that the semantic map explains the variance of the fixation point distribution better than that of saliency. This finding supports the theory of cognitive-guided visual perception. Note, however, that meaningful elements may well overlap with prominent features (Peacock et al., 2023). The design of semantic and salient maps also shows promise for studying the attentional processes of people with CRCI.

Figure 5. instance of a scene presented to participants (left) and its corresponding heat maps: visual fixation map (center left), semantic map (center right) and saliency map (right).

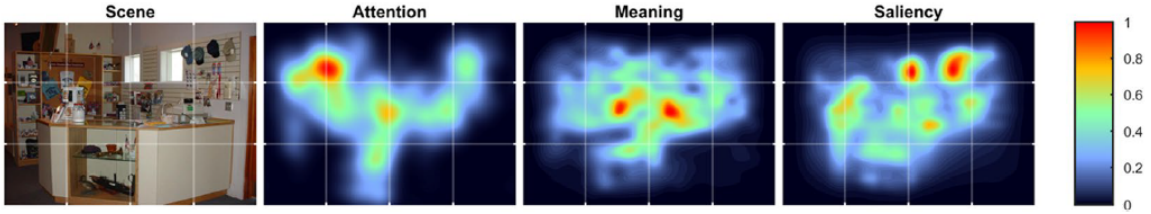


Figure 1. An example scene with its attention map empirically determined from viewer fixations, along with its meaning and saliency maps. (Grid lines are included for ease of comparison and were not present in the scenes).

Note: Reprinted with permission from J. M. Henderson et al. (2018). 2018 Springer Nature.

KEY SUMMARY — INVESTIGATING LINGUISTIC PROCESSES THROUGH EYE BEHAVIOR

Speech Guides Gaze

- Vision is a sensory system largely involved in the orientation of attention.
- Ocular behavior in composed of two main types of events: saccades, rapid movements that direct the fovea towards a target, and fixations, a succession of microsaccades on the same target within a common timeframe.
- The study of ocular behavior using eye-tracking techniques provides a better understanding of the cognitive processes underlying lexical search.
- When perceiving a visual scene, the gaze would be guided more by the search for meaningful elements (e.g., objects, animates, etc.) than prominent features (e.g., bright color).

CRCI Events Through Speech and Gaze

A growing number of studies are highlighting the value of speech analysis as a precise tool for the early detection of cognitive impairment. Speech-based linguistic cues contained in speech can be used to identify subtle changes in cognition, such as those observed in CRCI. How does CRCI express itself through speech and oculomotor behavior? To answer this question, appropriate speech parameters must first be selected. However, the literature on linguistic markers of cognitive disorders is vast, and the studied parameters remain poorly defined. As a result, the interpretation of their efficacy varies greatly from one study to another, making their reliability hazardous. This limitation can be overcome by completing a review of the parameters currently used in research. The results of this review will facilitate the choice of linguistic variables most likely to be associated with a CRCI.

In addition, only two studies have so far examined the linguistic aspects of CRCI (Myers et al., 2022; K. Williams et al., 2021). These studies were limited to the analysis of syntactic complexity, a narrow factor to capture the full complexity of cognitive alterations induced by cancer and its treatments. The high frequency of word-finding difficulties reported by cancer survivors highlights the need to understand the cognitive mechanisms underlying this phenomenon. This dissertation contributes to a better understanding of this impairment by identifying speech markers of CRCI. These new insights will help to design a speech-based automated diagnostic tool.

The **first article** in this dissertation, Linguistic markers of subtle cognitive impairment in connected speech: A Systematic Review, provides an overview of the linguistic parameters that are

used to detect subtle cognitive disorders in various etiologies and syndromes. The results of this synthesis will be used to pre-select potential CRCI markers. The **second article**, Using machine learning to identify speech markers of reported cognitive impairment in breast cancer survivors, aims to determine the predictive power of these markers using a machine learning model. This experimental paper compares speech samples from breast cancer survivors with cognitive complaints, survivors without cognitive complaints and control subjects. Finally, the **last article**, Oculomotor behavior of breast cancer survivors reporting cognitive changes, examines the cognitive processes of lexical retrieval in breast cancer survivors, using an eye-tracking technique.

General Discussion

The main objective of this dissertation was to gain a better understanding of the cognitive difficulties described by cancer survivors, and especially word-finding difficulties. This work first sought to identify prosodic features that could indicate the presence of a Cancer-Related Cognitive Impairment (CRCI). It then investigated the potential cognitive mechanisms leading to word-finding difficulties in CRCI.

First, this discussion focuses on prosodic markers of CRCI. It examines the extent to which confounding factors, such as fatigue, influence these cognitive difficulties and introduces the concept of a CRCI continuum. Second, it explores the cognitive deficits underlying word-finding difficulties in the context of CRCI. The dissertation then reviews the therapies that have been experimented with for CRCI rehabilitation. In conclusion, it presents the biases and limitations encountered in this research, before proposing directions for future studies.

1 Prosodic Signs of CRCI

The first experimental study of this dissertation aimed to identify potential prosodic markers of CRCI. The use of an automatic classification model revealed that the combination of speech-to-silence ratio, mean silent pauses duration, and mean filled pauses duration discriminate between cancer survivors with cognitive difficulties and control subjects.

Silent Interruptions

Two of these speech markers are related to the duration of silent interruptions. These results align with previous studies conducted on other populations with subtle cognitive impairments in different etiologies or syndromes. For instance, Vincze et al. (2021) compared individuals with Mild Cognitive Impairment (MCI) to patients diagnosed with Alzheimer's Disease and to healthy controls. MCI is a cognitive syndrome deemed an intermediate stage between normal aging and the onset of Alzheimer's Disease (Petersen, 2016). Like CRCI, MCI is characterized by cognitive complaints related to memory and word-finding difficulties. The expression of MCI is also subtle, affecting individuals' quality of life without hindering their autonomy. As a result, current neuropsychological tests struggle to detect this syndrome. Therefore, Vincze and her team aimed to identify individuals with MCI using

speech features, notably silent and filled pauses. According to their results, individuals with MCI produced more silent pauses, and these pauses were longer duration compared to both the controls and individuals with Alzheimer's Disease (Vincze et al., 2021). The results are consistent with our own findings, where silent-related variables proved the most representative of the population with CRCI.

Similarly, a study conducted by K. M. Smith et al. (2018) measured silent pauses duration during a picture-based description task. It aimed to identify speech markers of anomia in individuals with Parkinson's disease. The patients exhibited minor motor impairments, and two-thirds of them reported word-finding difficulties. However, their scores on cognitive tests remained within normal limits. The frequency of silent pauses was higher in patients compared to controls. Additionally, pause rate for those produced before a verb was correlated with the anomia reported by patients (K. M. Smith et al., 2018). Data from our studies also allowed for a more precise prediction of anomia as described by breast cancer survivors compared to the FACT-Cog questionnaire. The latter was not designed to identify word-finding difficulties, although it is a clinical reference tool for detecting at least the subjective presence of CRCI. The fact that the production of silence correlates with patients' complaints highlights the value of this method for assessing these cognitive difficulties.

Most studies exploring the relationship between mild neurocognitive disorders and speech focus on populations with age-related pathologies or neurodegenerative syndromes. However, the interest in silent pauses has also been observed in younger subjects, such as multiple sclerosis (Bóna, 2024) and individuals with mild symptoms of non-fluent aphasia following a stroke (Angelopoulou et al., 2018).

From a linguistic perspective, it is worth noting that the previously cited studies included speakers of Hungarian (Bóna, 2024; Vincze et al., 2021), English (K. M. Smith et al., 2018), Greek (Angelopoulou et al., 2018), and French (Richard, Lelandais, & Jacquin-Courtois, under review). Despite the linguistic diversity, silent pauses have been identified as a marker of cognitive impairment in all of these four languages. Silent pause duration or frequency could therefore serve as a common marker across multiple language systems. However, these cross-linguistic comparisons must be qualified, as silent pauses are often defined and measured differently across studies. The systematic review we conducted on linguistic markers of subtle cognitive impairment highlighted the variety of speech features that were used in studies on this topic. It seems essential to use features that are not influenced by cognitive or grammatical assumptions.

The speech-to-silence ratio could address this issue. Among the features we studied, this ratio stood out as one of the most promising for identifying cognitive difficulties. Outside clinical context, this parameter has proven effective in measuring fluency in second language learners (Cucchiarini et al., 2002; De Jong, 2023) or in understanding the relationship between vocabulary size and verbal fluency abilities (Jongman et al., 2021). However, this measure is still rarely used in studies on the detection of subtle cognitive impairments (see Pakhomov et al. [2010] on frontotemporal lobar degeneration). It did not emerge in our systematic review either (Richard, Lelandais, et al., 2024).

This is likely due to the fact that studies generally omit silent pauses that are related to grammatical functions (e.g., between two phrases) from their measurements.

The calculation of this ratio consists in dividing the total duration of silences by the total duration of the sample. A high ratio indicates fewer silent pauses, and therefore perceptibly more fluent speech. This ratio can be easily obtained using an Automatic Speech Recognition (*Automatic Speech Recognition*, ASR) system without requiring transcription (De Jong, 2018).

Unlike the average or median duration of silent pauses, which can be influenced by extreme values, the speech-to-silence ratio takes all silences into account. This measure is also less sensitive to stylistic variations in speech, as it does not rely on the qualitative aspects of the pauses. As a result, it limits issues related to the "nature" or "function" of pauses.

Although this ratio proves to be an effective way of detecting cognitive difficulties, it does not allow for deeper hypotheses to be formulated about any underlying disorders. The addition of another behavioral analysis technique such as eye tracking helps to better understand the cognitive processes underlying the difficulties experienced by breast cancer survivors when producing verbal messages.

Relationship Between Silence and Confounding Factors

Our results showed a trend for speech variables to be correlated with fatigue and depression scores. However, it seems premature to conclude a unidirectional relationship between the scores obtained on these tests and speech production for several reasons.

The presence of depressive symptoms may influence speech rate, particularly in the production of long pause duration (Albuquerque et al., 2021). Breast cancer survivors with cognitive complaints showed scores indicating higher levels of anxiety and depression compared to the control group. However, these scores did not reveal the presence of these symptoms in all survivors. Moreover, these scores remained below the alert thresholds (Brennan et al., 2010; Mitchell et al., 2010; Zigmond & Snaith, 1983). As such, the difference in anxiety and depression levels does not reflect a clinical reality, as these levels remain below the reference thresholds.

Regarding fatigue, there is no standardized threshold for detecting significant fatigue in breast cancer patients. Fatigue is a very common symptom among cancer survivors (Hofman et al., 2007). It is therefore possible that some of our participants felt tired. Nevertheless, it is important to note that the relationship between fatigue and decreased fluency is relatively under-documented and not clearly established to date. A well-established relationship concerns, however, the influence of fatigue on the motor control of speech (Hartelius et al., 2004). Yet the participants who had cancer yet did not exhibit slower articulation rates than those in the control group. This observation therefore rules out the hypothesis of fatigue influencing articulatory aspects in survivors.

In sum, depression and fatigue are frequently associated factors with CRCI (Gullett et al.,

2019). However, these factors are often related to treatments, particularly chemotherapy, which can exacerbate CRCI without directly explaining the observed cognitive dysfunctions (Meyers et al., 2005).

Furthermore, self-administered questionnaires are explicit because participants are aware of the criteria on which they are being evaluated. The understanding of the judgment criteria can thus influence their results, through an expectancy effect. Participants' responses are conditioned by their expectations regarding their membership in a specific group (cancer or control group). Breast cancer survivors may be inclined to exaggerate the extent of their difficulties, while control group participants may be prone to overestimate the efficiency of their cognition. In contrast, prosodic variables remain implicit. Combining both approaches is thus useful in this aspect.

Eventually, the observed relationship between speech production and depression or fatigue scores could actually capture the correlation between the FACT-Cog and these measures. CRCI is part of a set of symptoms that are associated with cancer (Kim et al., 2012; Starkweather et al., 2013), including anxiety, depression, and fatigue. In other words, the relationship between speech and psychological variables could be explained by the fact that survivors who reported cognitive impairment also reported higher levels of anxiety, depression, and fatigue (Veal et al., 2023). This observation is confirmed by our study using machine learning models. In this study, the population was divided into three groups: a first group of survivors whose FACT-Cog scores indicated cognitive problems, a second group of survivors without any cognitive complaint, and a third group of control subjects without any cognitive complaint either. However, several survivors without cognitive complaint were misclassified by the models into the CRCI group of individuals. These errors might be due to similarities between their prosodic features and those of survivors with a cognitive complaint.

Continuum of Cancer-Induced Cognitive Changes

The misclassifications made by our models are revealing. In general, the models were more successful at classifying participants based on their reported difficulty in word-finding than based on their FACT-Cog score. Among the survivors in the corpus, one third of them had an inferior score regarding the FACT-Cog threshold for cognitive complaints. However, six of these eleven participants were incorrectly classified into the CRCI group. In five of these cases, this misclassification mirrored the linguistic difficulties they perceived. The discrepancy between the FACT-Cog score and the production of speech interruptions could indicate a cognitive status that lies on a continuum of perceived and objective cognitive difficulties, ranging from normal cognition to manifest cognitive impairments.

Generally speaking, pause duration creates a clear distinction between survivors with complaints and the control group. In contrast, the variables of the group without complaints show greater spread. On the one hand, only three participants out of the eleven in this group reported no linguistic difficulties. On the other hand, the majority of their silent pauses are closer to those of the survivors

with complaints, while some are more aligned with observed pause range for the control subjects. The group of survivors without cognitive complaints thus shows significant spread, which would require further exploration. This variability could be explained by several factors, including treatments. Due to the diversity of oncological treatments and the small sample size, we were unable to control the treatments received by participants. The differences may be related to these treatments, as some of them such as tamoxifen (Ahles et al., 2010; Bender et al., 2006; Schilder et al., 2010) or chemotherapy (Quesnel et al., 2009), have a well-documented impact on verbal performance (but see Rick et al. [2018] for no effect of oncological treatment on cognition).

CRCI could be associated with a continuum of cognitive difficulties, with varying impacts not only depending on their severity and persistence (Ahles & Root, 2018), but also according to external factors such as daily life events. In a recent presentation, Ahles (2025) highlighted the evolving nature of cancer, with varying phenotypes from one patient to another. Similarly, CRCI would be subject to interindividual variability. Ahles (2025) suggests that lesions on the same brain regions can lead to diverse consequences. Each individual presents variations in cognitive performance over time. This variation is intensified by cancer and its treatments. In other words, the expression of CRCI could be modulated by each patient's daily life experiences. CRCI could thus be located on a variable spectrum of cognitive changes. Therefore, speech markers could reflect a specific facet of this continuum. A complementary analysis conducted at different time points is thus essential to validate this hypothesis.

Cognitive Correlates of Anomia in the Context of CRCI

Speech feature spread may reflect a continuum of cognitive difficulties encountered by patients. What exactly is the nature and extent of these difficulties? The analysis of oculomotor behavior during speech production has helped to better understand some underlying mechanisms of CRCI.

Naming Latency

Word-finding difficulties are a common symptom of CRCI, although their cognitive correlates remain poorly understood. Generally, lexical difficulties can result either from a degradation of the semantic system or from a deficit in access to the lexicon, which is often related to an impairment of executive control. A third hypothesis also suggests that attention plays a key role in the lexical selection processes.

As a reminder, in the case of a word-finding difficulty of semantic origin, individuals confuse or have forgotten the concepts (Caramazza & Hillis, 1990). In contrast, in the case of a word-finding difficulty of executive origin, individuals have access to the concept, but not to its phonological form (the sounds of words) (Caramazza & Hillis, 1990). In this regard, the verbatim speech of patients offers valuable insight into the underlying processes of the specific anomia they experience. In a

study conducted by Kanaskie and Loeb (2015), semi-structured interviews with post-breast cancer patients highlight the cognitive difficulties they face. As the only symptom unanimously shared, anomia is mentioned by all seven patients in the study. One patient describes a feeling of emptiness when speaking; she visualizes what she wants to say but is unable to find the right word. As reported by the authors, the patients expressed the feeling of not having access to words. This phenomenon is illustrated by a participant in a study by Player et al. (2014): "I'll say 'that purple plant there' but I can't tell you it's lavender ... until much later" (Player et al., 2014, p. 5). In a study by Bolton and Isaacs (2018), another patient testifies the following: "Oh can you pass me the writing thing, the writing stick with the ink in it'. I know what I want but I can't find the word..." (Bolton & Isaacs, 2018, p. 1226). Our participants report a similar complaint: "It's not that I don't know it, but I say the other word because it's easier. [...] I'll tell my children, 'go take that from the white shelf'. Even though I know it's brown, for instance." or "To uh t-t-try to explain, for instance, uh... to uh... overcome this problem where I'm searching for my words, well, I can detour with a definition, so that it's not obvious."[1].

In light of these elements, the patients' word-finding difficulties do not seem to result from a degradation of the semantic system. This hypothesis is supported by the fact that neuroimaging studies do not reveal any apparent lesions or abnormal activation in the areas that are associated with the semantic memory of words (Bernstein et al., 2021), such as the left temporal regions (Vigneau et al., 2006). However, results regarding regions associated with memory are more divergent, although not clearly established (Sousa et al., 2020). In contrast, neuroimaging studies conducted in cancer patients clearly highlight lesions or abnormal activations in brain regions that are associated with attentional and executive difficulties Kesler et al. (2009).

Our eye-tracking data are in agreement with the idea of an executive deficit rather than a memory deficit. On the one hand, the spatial and temporal analysis of fixation points indicated that cancer survivors do not need to look at the images more to retrieve their words. On the other hand, the delay between the first gaze at a visual item and its naming showed that survivors with cognitive complaints took longer than controls to search for their words. Taken together, these results suggest a lexical difficulty related to the selection or phonological encoding of words. In other words, they rule out the possibility of a degradation of the semantic system as the primary cause of the word-finding difficulty. Instead, they point more towards a problem related to the word's form: the concepts are preserved, but the phonological form remains temporarily inaccessible. When differences are detected, the scores on memory tests would reflect more of an executive and attentional deficit than a true memory disorder.

Connected Speech vs. Verbal Fluency

Word-finding difficulty is a common complaint among cancer patients. This issue is hardly detected by verbal fluency tasks, whereas our prosodic measures seem to reflect this difficulty more accurately.

[1] Unpublished dissertation data

Disparity in Verbal Fluency Results

According to the ICCTF recommendations (Wefel et al., 2011), the assessment of CRCI should include a verbal fluency task. This task serves to evaluate both lexical access and executive functions. In the instructions, the participant are typically asked to generate as many words as possible within a given time (usually one minute), beginning with a specific letter (letter or phonemic fluency) or belonging to a specific category (category or semantic fluency).

Several studies on CRCI evaluation include verbal fluency tests, but the results are divergent. Some of them confirm a difference in verbal fluency between cancer patients and controls (Boele et al., 2015), while others observe no difference (Mehlsen et al., 2009; Nguyen et al., 2013). Similarly, some studies highlight the impact of chemotherapy on verbal fluency performance (Durán-Gómez et al., 2022; Janelsins et al., 2022), while others find no effect (Biglia et al., 2012; Hurria & Lachs, 2007). A longitudinal study even presents a counterintuitive result: Andryszak et al. (2017) report an improvement in phonemic fluency in both controls and chemotherapy-treated patients over time. However, the authors attribute the improvement in performance to a training effect.

Cognitive Functions Engaged by Verbal Fluency

In clinical settings, verbal fluency tasks are administered either to estimate lexical access abilities or to assess executive functions. The fluency of isolated words is underpinned by the interaction between attentional processes, long-term verbal memory, and semantic knowledge (Ruff et al., 1997).

Verbal fluency tasks engage both strictly linguistic processes and executive functions (Aita et al., 2019; Shao et al., 2014; Whiteside et al., 2016). The use of neuroimaging showed that the activation of temporal and frontal regions overlap during both phonemic and categorical fluencies. However, greater involvement of frontal regions is observed during phonemic fluencies, while increased activation of temporal regions is noted during categorical fluencies (Katzev et al., 2013; Schmidt et al., 2019). In other words, verbal fluencies require both executive and linguistic functions: phonemic fluencies recruits executive functions more, while semantic fluencies mobilize linguistic functions more.

Isolated Word Fluency vs. Connected Speech Fluency

Fluency tasks are far from reflecting daily abilities. First, phonemic and semantic fluencies are not equivalent. Lexical semantic knowledge plays a more determining role in categorical fluency than in literal fluency (Shao et al., 2014). Moreover, categorical fluency would be easier to perform than literal fluency because the retrieval of a word activates other associated concepts. Second, isolated word fluencies would primarily activate relatively identifiable and restricted cognitive processes (Birn et al., 2010).

In contrast, connected speech relies on the complex activation of interacting cognitive processes (Hagoort, 2019). The production of a verbal message is thus more challenging than producing isolated words. In cancer survivors, this multiple activation could lengthen the lexical retrieval process. Compensatory mechanisms have been observed in neuroimaging studies during simple psychometric tasks, while scores on these tasks remained within normal ranges (Andryszak et al., 2017). These compensatory mechanisms manifest as hyper- or hypo-activation of the brain, and have mainly been found in the frontal regions that are associated with the attentional network and executive control (B. C. McDonald et al., 2012). The activation of frontal regions is necessary to avoid interference with language production (Dhooge & Hartsuiker, 2012). This finding could explain the discrepancy between the normal scores obtained during verbal fluency tasks for cancer patients and the speech-related variables, as these two measures are distinct. The naming delay in connected speech marked by silent or filled pauses would thus be more sensitive than verbal fluencies in detecting subtle anomia.

Anomia Reveals a Broader Disorder

In summary, anomia is a distinctive feature of CRCI. It thus serves as an entry point to investigate this syndrome. Although patients primarily complain of memory problems, an increasing number of studies suggest that there could be an issue upstream of the encoding process, particularly related to an attentional deficit. Anomia in CRCI could be a consequence of this attentional deficit, especially during lexical retrieval. The relationship between speech production and CRCI warrants further exploration.

2 Clinical Interventions for CRCI Treatment

This section examines therapeutic options that can be proposed for CRCI. It then suggests potential speech-language pathology interventions for the cognitive-linguistic disorders associated with CRCI. Like other cognitive disorders, the clinical management of CRCI relies on evidence-based practice. Several types of interventions aimed at treating CRCI are regularly evaluated in clinical trials (Zeng et al., 2020). However, a systematic review conducted by Merceur et al. (2024) reveals that about one in two studies fails to demonstrate the effectiveness of therapies, including cognitive rehabilitation, physical rehabilitation, and combined therapies.

Current Therapies

Although the focus of this dissertation is not on the management of CRCI, its findings contribute to enriching the existing knowledge on the mechanisms of this disorder. The originality of this work lies in the behavioral data provided through the analysis of speech and eye movements. Up

to now, studies on the objective evaluation of CRCI have relied on classical neuropsychological tests, whose scores were only weakly associated with the difficulties perceived by patients. In contrast, our behavioral data seem to capture the extent of these deficits. Furthermore, our results appear to introduce a continuity between the observed brain changes and the way they manifest in the behavior of survivors. These analyses highlight deficits that are primarily executive in nature. These results are similar to those of Edelman et al. (2024), although they need to be confirmed by larger studies. To our knowledge, this study is one of the few to include behavioral data for the evaluation of CRCI (see Huang et al. [2025] for a study using an eye-tracking technique). These new insights into the underlying mechanisms of CRCI can guide the choice of clinical interventions.

Pharmacological Treatments

Pharmacological treatments tested for CRCI include drugs used to treat attention deficits and dementias (Karschnia et al., 2019). For instance, the use of dopaminergic and noradrenergic antagonists such as methylphenidate and modafinil has been tested in cancer patients to regulate the prefrontal network involved in attentional dysfunctions. However, these treatments have not shown effectiveness in reducing CRCI. Regarding anti-dementia drugs, acetylcholinesterase inhibitor donepezil, which acts primarily on the hippocampus (i.e., the memory center), has shown marginal effects. Finally, the use of other substances such as ginkgo biloba or erythropoietin (i.e., a hormone stimulating red blood cell production) has also not shown significant effects. In summary, the results of pharmacological treatments remain inconclusive at present (Yang & Ah, 2024).

Cognitive Therapies

Cognitive therapies aimed at treating CRCI are the most well-documented interventions and those with the best rates of efficiency (e.g., [Ercoli et al., 2015]). Bray et al. (2017) demonstrate the efficiency of a remote cognitive rehabilitation program with 242 participants. This intervention resulted in a reduction of cognitive symptoms as assessed by the FACT-Cog, although no effect was observed through neuropsychological tests.

Cognitive-behavioral therapies, such as psychoeducation on accepting cognitive losses, showed improvements in cognitive performance. However, additional evidence is needed to confirm these results (e.g., [Ferguson et al., 2012]).

Mindfulness-based cognitive therapies, which are primarily used to reduce stress, have also shown encouraging functional changes at the brain level during working memory tasks (Melis et al., 2025). However, another study does not confirm the effects of these therapies when compared to a physical activity intervention or a waitlist control group (Melis et al., 2023). In contrast, other stress-reduction techniques aimed at improving cognitive performance, such as Qigong, Tai Chi, yoga, supportive therapy, and acupuncture, have shown little (Yang & Ah, 2024) or no effects (Zeng et al., 2020).

Physical Activity

The effects of physical exercise on the brain could represent a valuable resource for the management of CRCI. Engaging in physical activity promotes neurogenesis (the formation of new neurons) in the hippocampus, a brain structure involved in memory (Fabel & Kempermann, 2008). Additionally, exercise has shown significant effects on improving the physical state of cancer survivors, as well as their levels of anxiety, stress, and depression. The interest of physical activity to restore cognitive efficiency in cancer survivors is therefore logical. However, the benefits of such interventions remain limited. In clinical trials involving physical activity, survivors reported a positive change in their cognitive status (Hartman et al., 2024). On the other hand, exercise had no impact on deficits measured by tests (Hartman et al., 2024). In another study, aerobic exercise did not show a significant difference compared to being on a waiting list (Brunet et al., 2025). Therefore, physical activity alone appears ineffective in treating CRCI (Campbell et al., 2020). These results are nonetheless limited, partly due to the diversity of offered exercises and the lack of appropriate cognitive evaluation. More rigorously controlled studies are required to draw reliable conclusions regarding the effects of physical activity on CRCI (Yang & Ah, 2024). However, it is beneficial for mitigating the impact of confounding factors such as fatigue, anxiety, and depression (Campbell et al., 2019). For this reason, physical activity is at least recommended for improving the psychological and physical well-being of cancer survivors (Campbell et al., 2019). The recommendations for regular physical activity are, however, not specific to cancer. The World Health Organization recommends daily exercise for the entire population (*WHO Guidelines on Physical Activity and Sedentary Behaviour*, 2020). While sedentary behavior and inactivity increase the risk of cardiovascular diseases, physical exercise protects against the onset of these conditions and reduces the risk of psychiatric diseases (*WHO Guidelines on Physical Activity and Sedentary Behaviour*, 2020).

Combinatory Approaches

Yang and Ah (2024) suggest that combining physical exercises and cognitive rehabilitation could have promising effects. However, the samples in these studies were too small for the results to be generalized. Despite the lack of solid evidence, a Delphi study on the detection and rehabilitation of CRCI indicates that a multimodal approach is nevertheless recommended by the consulted expert panel (Duivon et al., 2024).

Interventions in Speech-Language Pathology

Given the prevalence of linguistic difficulties associated with CRCI, the role of the speech-language pathologist in its management is well justified. A study conducted by Geels et al. (2024b) assessed the knowledge of American speech-language pathologists regarding CRCI and their clinical practice. Three-quarters of participants thought that managing CRCI fell within the scope of speech-language

pathology, and 65% reported having already provided care to this population. In France, speech-language pathologists are the healthcare professionals responsible for the rehabilitation and adaptation of cognitive-linguistic disorders, which affect language, speech, and communication. In cancer populations, language and communication disorders might be more related to an executive-attention deficit rather than a true language disorder. The observed anomia in cancer patients nevertheless generates anxiety during their social interactions (Geels et al., 2024a; Kanaskie & Loeb, 2015). It seems relevant to integrate speech-language pathology interventions into oncological support care.

However, no scientific study has yet explored speech-language pathology interventions in the context of CRCI. Development of speech-language pathology methods could be slow down by the lack of documentation on linguistic disorders encountered by cancer patients. This results in the absence of validated program for the management of cognitive-linguistic disorders in this context. Given the cognitive mechanisms and inflammatory processes underlying CRCI, it could be interesting to test validated speech-language pathology interventions on populations with similar difficulties associated with brain inflammation, such as individuals with multiple sclerosis or post-COVID-19. An early and short-term intervention that combines rehabilitation and adaptation could be considered. The aim would be to facilitate the transfer of results, given that the disorder mainly manifests in daily life. Therefore, speech-language pathology rehabilitation for the management of CRCI could be integrated into a holistic care approach.

3 Limitations and Perspectives

This section presents the main limitations that were encountered during this dissertation, as well as the perspectives for future research. This preliminary work opens the way for further research on speech characteristics and oculomotor behavior in CRCI.

First, it is important to highlight the exploratory nature of this study. The research aimed for an ecological objective, that is, an analysis of speech and its cognitive correlates in as natural a context as possible. However, this goal comes with certain methodological limitations, namely the low control of temporal features such as speech duration and linguistic variations in the discourse production itself. Unlike more controlled experimental paradigms, such as those analyzing isolated words or reading, speech in a discourse context can be influenced by multiple factors (emotions, environment, etc.). The lack of control over discourse features could introduce additional variability, complicating the interpretation of the cognitive mechanisms underlying the prosodic features.

Moreover, the sample size constitutes another major limiting factor. This pilot study is based on a few participants, which prevents any generalization of the results to a larger population of cancer survivors. This small sample size does not allow for definitive conclusions. In particular, it is not possible to obtain an effect size for prosodic variables with a limited number of participants. In order for the results to be considered reliable, this study would need to be replicate with a larger number of

participants and with more speech samples per participant.

Moreover, this dissertation focused exclusively on breast cancer survivors. This approach precludes any generalization of the results to other types of cancer. The cognitive differences observed in this specific population may be influenced by specific factors to this pathology, such as the treatment type or the physiology of breast cancer. The lack of clinical diversity currently limits the scope of the conclusions of our work. Future research should include participants who have undergone various types of treatments to better understand the variations between different cancers. The goal will be to determine if the prosodic markers identified in this dissertation are also present in populations with other types of cancer.

Finally, an important limitation of this study is the lack of control over the administered treatments to participants. The small sample size makes it difficult to account for this variable, which is essential because the effects of treatments on cognitive functions can be numerous and varied. The impact of chemotherapy, radiotherapy, or other specific treatments could interfere with the observed cognitive symptoms, but it is difficult to disentangle these effects in the absence of a larger sample and more precise data on the treatments. To overcome this limitation, it would be beneficial to conduct a multicenter study involving a larger number of research centers and participants. Such an approach would allow for the collection of more diverse data and better control over the effects of different treatments on cognitive disorders, while also increasing the representativeness of the results.

Given the timeline of cancer, it seems necessary to assess speech and eye movements at several time points throughout the evolution of the disease and its recovery. An ongoing study conducted by our team evaluates the speech of participants with breast cancer at two time points: immediately after the end of treatment and at a distance, more than a year after treatment cessation, in order to observe any changes. In the long term, this technique will help identify those patients at risk of developing persistent disorders.

A long-term goal is to develop a tool combining automatic speech recognition and a classification model to allow rapid identification of CRCI and different patient profiles. It is common for the obtained patients scores not to fall within pathological thresholds. However, cognitive deterioration can be observed when these scores are compared over time (Stewart et al., 2008). Establishing a baseline before the start of treatments is essential to assess the extent of cognitive changes induced by cancer (Janelsins et al., 2011). This process can be carried out by collecting a speech sample. Analyzing speech could thus facilitate the monitoring of cognitive changes. This easy and quick to use tool could be used by various healthcare professionals, such as oncologists, rehabilitation doctors, nurses, and speech therapists. It could also be introduced at different time points, before treatments, or even before diagnosis, such as before a biopsy.

Conclusion

Cancer is a heterogeneous and widespread disease. Thanks to advances in medical knowledge and techniques, certain types of cancer now offers better survival rates, such as breast cancer. Quality of life after cancer has become a major issue. Cancer can cause numerous dermatological, psychological, and even cognitive sequelae, such as Cancer-Related Cognitive Impairment (CRCI).

CRCI is a side effect of cancer and its treatments. Patients report difficulties with memory, attention, executive functions, and language. These cognitive difficulties can significantly affect their daily lives. Along with memory problems, word-finding difficulties feature among the most common cognitive symptoms. CRCI prevents some patients from returning to work and limits their social interactions.

CRCI is generally observed during or immediately after treatments (i.e., surgery, chemotherapy, radiotherapy, hormone therapy). It can sometimes appear before treatments are administered. In most cases, symptoms decrease during the first few years following the cessation of treatments. However, some patients experience persistent symptoms that may last up to ten years after treatment ends.

The pathophysiological origin of this disorder remains poorly understood. Recent research suggests that the cancer itself may cause neurological inflammation, which in turn leads to cognitive difficulties. The role of treatments such as chemotherapy, radiotherapy, and hormone therapy is well established in the worsening of this cognitive disorder. Other factors, such as fatigue, anxiety, and depression, are often associated with CRCI and contribute to the worsening of its symptoms. However, they alone cannot explain the development of this disorder.

These difficulties are often highlighted during interviews with patients, as well as by standardized questionnaires such as the FACT-Cog. These questionnaires show a wide range of experienced difficulties on a daily basis. While useful for detecting the presence of cognitive complaints, they do not allow for specifying the nature of these difficulties or identifying their cognitive origin.

In contrast, neuropsychological tests yield more mixed results. Patients' scores generally fall within the normal range. This discrepancy between perceived difficulties and normal scores may be due to a lack of sensitivity in the tests or to a divergence between subjectively and objectively evaluated performance. The lack of appropriate diagnostic tools prevents some patients from accessing proper care and slows down the understanding of the mechanisms behind CRCI. New approaches are therefore needed to better understand this disorder.

Language is a complex cognitive function that interacts with other functions, such as memory, attention, executive functions, and motor functions. If one of the necessary mechanisms for the normal functioning of language is deficient, speech production will be affected. In fact, connected speech (i.e., the chaining of phonemes) contains many prosodic cues that can reveal cognitive difficulties.

Speech production is rarely linear. It is often interrupted by pauses of varying duration, which disrupt its fluency. However, not all of these interruptions necessarily indicate cognitive difficulties. A detailed analysis of "abnormal" interruptions allows for the identification of certain populations with cognitive disorders. Lexical retrieval, for instance, can be laborious. In this case, it often results in pauses before the sought-after words. When these difficulties are relatively frequent, they are referred to as word-finding difficulty or anomia. Anomia can result from three distinct cognitive deficits: first, memory and the semantic system degradation, as seen in patients with Alzheimer's Disease; second, executive functions and difficulties accessing the phonological form of a word; and third, attentional deficits that would disrupt the inhibition of co-activated concepts during lexical retrieval. One of the current challenges in speech analysis lies in the absence of a method or linguistic criteria that have reached consensus for evaluating cognitive disorders. Although speech analysis can indicate cognitive variations, it is necessary to use a complementary technique to identify the processes involved in word-finding difficulties and to understand their cognitive origin.

In this regard, eye-tracking is a valuable complementary technique for studying speech and cognition. Research indeed has shown that gaze generally directs toward elements that are semantically related to the words heard. On the production side, the visual world paradigm has provided evidence on the incremental process of speech. In other words, a speaker begins articulating the phonemes of a word before fully encoding the rest of their message. This process would explain the presence of interruptions in speech. Moreover, a less frequent word results in a longer retrieval and lexical search time. Finally, the visual elements that carry semantic information are looked at in the order of their naming. The calculation of the delay between fixations on a visual element and its naming could thus provide information about the ease of retrieval of this word.

Vision plays a major role in the orientation of attention. Two main theories oppose each other on this subject. One proposes that our gaze is attracted to the prominent elements of a visual scene, such as a bright color (i.e., image-guidance theory). The other suggests that vision is guided by our search for meaning (i.e., cognitive-guidance theory). Comparing saliency and semantic heatmaps to participant fixations supports the hypothesis of the search for meaning. This method could be useful for studying potential attentional disorders in cancer patients.

The first article of this dissertation reviewed the literature on potential markers of subtle cognitive disorders. The goal was to list the features currently in use and assess their potential for detecting CRCI.

The second article assessed the discriminative power of fourteen preselected prosodic features. The goal was to determine the optimal combination of parameters to classify cancer patients with

CRCI, survivors without complaints, and controls, using machine learning models.

The third article tested three hypotheses regarding the possible origin of word-finding difficulties in the context of CRCI. The first hypothesis suggested that word-finding difficulties would be related to a degradation of the semantic system. If this were the case, the patients would have spent more time than the controls fixating on the visual elements whose naming proved complicated. The second hypothesis proposed that word-finding difficulties were due to a problem accessing the phonological form of the word. In this case, the patients would have shown a longer naming delay between the first fixation of an element and its naming, compared to the controls. The third hypothesis considered that word-finding difficulties might be related to an attentional deficit. The patients would have been more attracted to visually salient elements of the images than to those carrying meaning, compared to the controls.

The first objective of this dissertation was to identify characteristic speech features for CRCI. Among a set of linguistic features, the combination of three prosodic parameters allowed for the identification of cancer survivors with cognitive complaints, those without cognitive complaints, and controls with an accuracy of 73.7%. The model's errors primarily concerned survivors without cognitive complaints. These patients scored within the normal range on the FACT-Cog, although they still reported difficulties. A possible explanation is that CRCI represents a continuum of difficulties that are subject to performance variations not captured by the FACT-Cog, but which were captured by speech analysis. Two of the three prosodic markers were related to silence. Silence might, more than other speech parameters, reflect cognitive difficulties, as can be observed in other pathologies or syndromes.

The second objective was to understand the underlying mechanisms of CRCI through the study of anomia. The results suggested that the naming delay would indicate an executive origin of anomia rather than a memory-based one. The hypothesis of an attentional deficit was not validated by our data, but seems difficult to entirely rule out, given the research indicating attentional deficits in cancer patients. The visual material used in this dissertation likely lacked the complexity needed to capture these difficulties. Further research on the topic is necessary to confirm or reject this hypothesis.

Through the combination of speech analysis and an eye-tracking technique, this dissertation also contributes to the identification of patients at risk of developing CRCI. Moreover, it provides insights into the mechanisms behind the origin of word-finding difficulties within the context of CRCI.

From a linguistic perspective, our dissertation points out the absence of homogeneous markers. The wide diversity of parameters and the lack of consensus limit the reproducibility of studies. However, it proposes a terminology that could serve as a reference for future studies. It also offers a rigorous methodology for speech analysis that can be applied in clinical settings.

In conclusion, this dissertation lays the first foundation for the development of automated diagnostic tools for CRCI. The work initiated in this dissertation should be continued with further

research on clinical populations reporting subtle cognitive difficulties. It would, for instance, be relevant to study other oncological populations (lung cancer, testicular cancer, etc.), as well as other pathologies or syndromes with similar symptoms.