



**HAL**  
open science

# Établissement de normes francophones d'un outil d'évaluation quantitative du langage naturel

Emma Boyer-Gibaud, Mélanie Rousseaux

► **To cite this version:**

Emma Boyer-Gibaud, Mélanie Rousseaux. Établissement de normes francophones d'un outil d'évaluation quantitative du langage naturel. Sciences cognitives. 2016. dumas-01357237

**HAL Id: dumas-01357237**

**<https://dumas.ccsd.cnrs.fr/dumas-01357237>**

Submitted on 17 Sep 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ACADÉMIE DE PARIS  
UNIVERSITÉ PARIS VI PIERRE et MARIE CURIE  
MÉMOIRE POUR LE CERTIFICAT DE CAPACITÉ  
D'ORTHOPHONISTE

**ÉTABLISSEMENT DE NORMES  
FRANCOPHONES D'UN OUTIL  
D'ÉVALUATION QUANTITATIVE DU  
LANGAGE NATUREL.**

Sous la direction de Sylvie HÉBERT professeure à l'École  
d'orthophonie et d'audiologie de l'Université de Montréal et  
d'Anna ZUMBANSEN docteur en orthophonie.  
BRAMS : Laboratoire international de recherche sur le cerveau, la  
musique et le son.

ANNÉE UNIVERSITAIRE 2015-2016

**BOYER-GIBAUD Emma**

Née le 20 janvier 1993

**ROUSSEAUX Mélanie**

Née le 9 avril 1990

# REMERCIEMENTS

Nous remercions tout d'abord Madame Sylvie HÉBERT de nous avoir accompagnées dans toutes nos démarches pour venir effectuer notre dernière année d'études au Québec et de nous avoir permis de réaliser notre travail de recherche au sein du laboratoire BRAMS.

Merci beaucoup à Anna ZUMBANSEN pour son aide et sa présence tout au long de la réalisation de notre mémoire. Sa disponibilité et son soutien à chaque instant nous ont permis de nous sentir accompagnées et rassurées dans chacune des étapes du mémoire.

Merci à l'association « Aphasie Rive Sud » de Saint Hubert, association de personnes aphasiques grâce à laquelle nous avons eu une expérience humaine et professionnelle unique.

Nous remercions enfin nos familles et nos proches pour leur soutien tout au long de cette année au Québec, et au cours de toutes nos années d'études.

# ENGAGEMENT DE NON-PLAGIAT

Je soussignée Emma BOYER-GIBAUD, déclare être pleinement consciente que le plagiat de documents ou d'une partie d'un document publiés sur toutes formes de support, y compris l'Internet, constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée. En conséquence, je m'engage à citer toutes les sources que j'ai utilisées pour écrire ce mémoire.

Signature :

# **ENGAGEMENT DE NON-PLAGIAT**

Je soussignée Mélanie ROUSSEAUX, déclare être pleinement consciente que le plagiat de documents ou d'une partie d'un document publiés sur toutes formes de support, y compris l'Internet, constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée. En conséquence, je m'engage à citer toutes les sources que j'ai utilisées pour écrire ce mémoire.

Signature :

# TABLE DES MATIÈRES

<b>INTRODUCTION</b> .....	<b>1</b>
<b>PARTIE THÉORIQUE</b> .....	<b>3</b>
<b>A. Le manque d’outil pour mesurer les progrès langagiers chez les personnes aphasiques</b> .....	<b>3</b>
(Emma BOYER-GIBAUD)	
1. Importance de la mesure des progrès langagiers.....	3
2. Les tests actuels.....	4
3. Un nouveau test répondant au besoin d’évaluation des progrès des patients.....	5
<b>B. Le <i>Montreal Assesment of Connected Speech</i> (MACS)</b> .....	<b>6</b>
(Emma BOYER-GIBAUD)	
1. Introduction.....	6
2. Les mesures choisies.....	6
3. Les stimuli utilisés.....	8
<b>C. Propriétés psychométriques du MACS</b> .....	<b>10</b>
(Mélanie ROUSSEAUX)	
1. Validité.....	10
a. Validité d’apparence.....	10
b. Validité de contenu.....	11
c. Validité de construit.....	12

2. Fiabilité.....	13
a. Fiabilité inter-juges.....	13
b. Fiabilité des versions parallèles.....	14
c. Fiabilité test-retest.....	14
3. Sensibilité.....	16
<b>D. Nécessité d'établir des normes.....</b>	<b>18</b>
(Mélania ROUSSEAUX)	
<b>PARTIE PRATIQUE.....</b>	<b>19</b>
(Emma BOYER-GIBAUD et Mélania ROUSSEAUX)	
1. Objectifs.....	19
2. Méthodologie.....	19
a. Participants.....	19
b. Matériel.....	21
c. Déroulement.....	23
d. Analyse.....	24
3. Résultats.....	25
<b>DISCUSSION.....</b>	<b>35</b>
<b>CONCLUSION.....</b>	<b>38</b>

## LISTE DES TABLEAUX

**Tableau 1** : Caractéristiques des participants de la tranche d'âge 31-50 ans.

**Tableau 2** : Répartition des quinze images en trois groupes d'images.

**Tableau 3** : Caractéristiques des témoins francophones québécois du MACS.

**Tableau 4** : Statistiques descriptives et distribution des scores d'Informativité par groupe d'image.

**Tableau 5** : Statistiques descriptives et distribution des scores de Grammaticalité par groupe d'image.

## LISTE DES FIGURES

**Figure A** : Distribution des scores d'Informativité pour le groupe 1.

**Figure B** : Distribution des scores de Grammaticalité pour le groupe 1.

**Figure C** : Distribution des scores d'Informativité pour le groupe 2.

**Figure D** : Distribution des scores de Grammaticalité pour le groupe 2.

**Figure E** : Distribution des scores d'Informativité pour le groupe 3.

**Figure F** : Distribution des scores de Grammaticalité pour le groupe 3.

# LISTE DES ANNEXES

**Annexe A** : Image 1 Banque

**Annexe B** : Image 2 Campagne

**Annexe C** : Image 3 Chambre

**Annexe D** : Image 4 Cinéma

**Annexe E** : Image 5 Cuisine

**Annexe F** : Image 6 Épicerie

**Annexe G** : Image 7 Fête

**Annexe H** : Image 8 Montagne

**Annexe I** : Image 9 Parc

**Annexe J** : Image 10 Pêche

**Annexe K** : Image 11 Plage

**Annexe L** : Image 12 Restaurant

**Annexe M** : Image 13 Rue

**Annexe N** : Image 14 Salon

**Annexe O** : Image 15 Zoo

# GLOSSAIRE ET ABRÉVIATIONS

AVC : Accident Vasculaire Cérébral

BDAE : Boston Diagnostic Aphasia Examination

ECVB : Échelle de Communication Verbale de Bordeaux

ET : Écart-type

M : Moyenne

MACS : Montreal Assesment of Connected Speech

MIT : Melodic Intonation Therapy

MMSE : Mini Mental State Examination

MT86 : Protocole Montréal- Toulouse d'examen linguistique de  
l'aphasie

OMS : Organisation Mondiale de la Santé

TLC : Test Lillois de Communication

UIC : Unité d'Information Correcte

# INTRODUCTION

« L'aphasie désigne l'ensemble des troubles de la communication par le langage secondaires à des lésions cérébrales acquises entraînant une rupture du code linguistique » (Chomel-Guillaume et coll., 2010, p.61 [5]). Ce trouble du langage se caractérise par des atteintes diverses tant au niveau de la production que de la réception du langage dans les modalités orale ou écrite. Réduction de l'expression orale, trouble de la compréhension, déformations phonétiques, difficultés syntaxiques constituent les principales difficultés de l'aphasie selon les aires lésées et le type d'aphasie.

Les chiffres témoignent de la proportion de personnes touchées. Selon l'*Organisation Mondiale de la Santé* (OMS), chaque année près de 6,7 millions de personnes dans le monde meurent d'AVC (*World Health Organization*, 2012. [29]). En France, 150 000 personnes sont victimes d'un AVC chaque année et la moitié en garderont de graves séquelles (*France AVC*, s.d. [14]).

L'évaluation et la rééducation orthophoniques des personnes aphasiques sont vivement encouragées. L'orthophoniste accompagne le patient dans l'acceptation du trouble du langage et permet une stimulation immédiate et intensive pour favoriser la récupération et l'acquisition de stratégies de compensation (Brady et coll., 2012 [2]). L'évaluation orthophonique permet un état des lieux des difficultés du patient nécessaire pour aborder la rééducation du trouble du langage (Chomel-Guillaume et coll., 2010 [5]). Les cliniciens peuvent s'aider de batteries de tests standardisées et normées pour évaluer la sévérité et le type d'aphasie, telles que le *Protocole Montreal-Toulouse d'examen linguistique de l'aphasie* (MT86, Nespoulous et coll., 1992 [22]) ou l'*Échelle d'évaluation de l'aphasie adaptée du Boston Diagnostic Aphasia Examination* (Mazaux et Orgogozo, 1982 [18]).

Le bilan de renouvellement tente par la suite de quantifier les progrès du patient. Proposer une même épreuve à quelques mois d'écart présente cependant un risque de biais d'apprentissage et de ce point de vue, les orthophonistes ne disposent pas d'outils adéquats pour mesurer l'efficacité de la thérapie et les progrès du patient. De plus, la réussite aux tâches langagières de dénomination ou de répétition ne garantit pas une amélioration de la communication dans les situations de communication du quotidien comme une conversation.

Le présent mémoire de fin d'études a été réalisé autour d'un test orthophonique en cours d'élaboration, le *Montreal Assessment of Connected Speech* (MACS), répondant au besoin de tests orthophoniques permettant l'évaluation des progrès langagiers de personnes aphasiques après rééducation. Cet outil propose une mesure des progrès langagiers dans le discours naturel grâce à une tâche de description d'images.

Notre travail avait pour but d'achever le travail d'établissement des normes des témoins francophones québécois, en examinant la performance de sujets âgés de 31 à 50 ans.

Après avoir mis en évidence le manque d'outil pour mesurer les progrès langagiers chez les personnes aphasiques, nous présenterons le MACS. Nous aborderons ensuite les critères psychométriques qui s'y rattachent pour garantir la fiabilité et la validité de l'outil, puis nous nous pencherons sur la nécessité d'établir des normes. Nous présenterons enfin notre travail de recherche ayant mené à l'établissement des normes des témoins francophones québécois pour la tranche d'âge 31-50 ans du MACS. Nous terminerons par une discussion critique des résultats obtenus et évoquerons les pistes ouvertes par cette recherche.

# PARTIE THÉORIQUE

## A. Le manque d'outil pour mesurer les progrès langagiers chez les personnes aphasiques.

### 1. Importance de la mesure des progrès langagiers

L'aphasie est un déficit langagier entraînant des difficultés de communication qui engendrent une souffrance chez les personnes atteintes. Le manque de facilité et de fluidité de la communication rend l'échange difficile avec l'entourage et peut causer un repli sur soi. La perturbation du langage entraîne ainsi un handicap psychosocial à prendre en compte (Daviet et coll., 2002 [7]).

Les orthophonistes francophones disposent d'outils d'évaluation de l'efficacité communicationnelle des patients aphasiques. *L'Échelle de Communication Verbale de Bordeaux* (ECVB) est un questionnaire sur les situations de communication quotidiennes rempli par le patient ou son entourage (Darrigrand et Mazaux, 2005 [6]). Le Test Lillois de Communication (TLC), est lui destiné aux orthophonistes qui doivent observer des échanges marqués par différents degrés de contraintes communicationnelles (Rousseaux et coll., 2001 [26]). Néanmoins, ces tests ne font pas état des capacités purement langagières de la personne mais de sa communication par le biais de tous les canaux dont elle dispose. De plus, l'amélioration de la communication globale peut subir l'influence de facteurs subjectifs comme la confiance en soi ou de nouvelles stratégies de communication non-verbale efficaces. Face à une communication plus fonctionnelle, les difficultés pour transmettre une information de manière strictement verbale demeurent un obstacle douloureux, c'est pourquoi l'évaluation des progrès strictement langagiers reste capitale. Pour ce faire, mesurer l'efficacité d'une thérapie lors d'une tâche de production de langage naturel paraît plus écologique que lors de tâches de répétition ou de dénomination ne reflétant pas une réelle situation de communication (Menn et coll., 1994 [19]). Les tests actuels présentés ci-après, évaluent des productions de langage naturel mais ne permettent pas une évaluation des progrès langagiers des patients aphasiques.

## 2. Les tests actuels

Actuellement, les orthophonistes francophones disposent de deux batteries de tests standardisées du langage pour l'aphasie. Ces deux tests utilisent notamment la description d'image comme un indicateur de la qualité et la quantité de la production linguistique naturelle du patient.

Le *Protocole Montréal-Toulouse d'examen linguistique de l'aphasie* (MT86) contient une description d'image pour évaluer le discours narratif oral. L'orthophoniste cote ensuite plusieurs aspects du discours du patient sur une échelle. Il est ensuite possible de faire une évaluation plus détaillée de cet échantillon de discours en utilisant des tableaux d'analyse syntaxique et d'analyse lexico-sémantique. L'évaluation du discours oral effectuée grâce à cette description est ensuite relue conjointement avec les résultats des autres domaines évalués dans le test (expression écrite, compréhension orale et écrite, transpositions) afin de pouvoir définir le profil du patient.

Le Boston Diagnostic Aphasia Examination (BDAE) contient une image que le patient doit décrire. Le discours transcrit est divisé en énoncés par le professionnel, qui détermine le type d'énoncé auquel chacun correspond. Un indice de complexité est ensuite calculé avec certains types d'énoncés. Cette épreuve permet de déterminer un profil d'aphasie lorsqu'on l'associe à toutes les autres épreuves contenues dans le BDAE.

L'ensemble des sous-tests de ces batteries incluant une description d'image, a l'avantage de tester différents aspects de la communication (orale, écrite, compréhension, expression...). Ces batteries de tests ont pour objectif de comprendre le type d'aphasie que le patient présente ainsi que les domaines dans lesquels il se trouve le plus en difficulté et ceux dans lesquels il a conservé une partie ou toutes ses capacités. Elles correspondent au besoin qu'ont les professionnels de faire un état des lieux au début de la prise en charge.

L'orthophoniste peut, par la suite, s'appuyer sur ce qui constitue les points forts du patient afin de stimuler ce qui est lésé. Cependant, ces outils n'ont pas été conçus pour évaluer une évolution des performances du patient. Or, lors de la création d'un test, il est nécessaire de déterminer précisément l'objectif de ce test afin de pouvoir vérifier que chaque étape (telle que le choix des items, la construction de la tâche, la définition de la cotation) s'accorde avec le but principal (Ivanova et Hallowell, 2013 [15]). Trois problèmes majeurs peuvent être soulevés à cet égard. D'abord, chaque test ne contient qu'une image à décrire. Or, ceci pose problème car l'échantillon produit serait trop court et donc aurait une fiabilité test-retest insuffisante. Ensuite, aucune version parallèle à ces tests n'a été conçue.

Finalement, la version existante de chacun des deux instruments n'a pas été validée psychométriquement pour l'évaluation des améliorations. Les tests diagnostiques de l'aphasie présentent donc le risque important d'être inadéquats pour évaluer les progrès des patients.

### **3. Le Montréal Assessment of Connected Speech (MACS) : un nouveau test répondant aux besoins d'évaluation des progrès des patients.**

Le MACS a pris naissance dans l'objectif de percevoir l'évolution linguistique des patients aphasiques dans leur discours naturel grâce à un test de description d'images normé et standardisé comportant plusieurs versions. Ce type d'évaluation est primordial dans un métier en construction et en constante évolution comme l'orthophonie car il permet de tester l'efficacité des techniques de stimulation et de rééducation utilisées.

Afin d'évaluer la progression des capacités langagières de communication d'un patient, il est nécessaire de prendre en compte plusieurs aspects. Tout d'abord, le patient ne doit pas être testé sur le même matériel pour éviter le biais d'apprentissage. Un effet d'apprentissage (ou effet test-retest) est une amélioration de la performance liée à la simple répétition d'une tâche dans le temps. Le score du sujet va donc être meilleur sans que le sujet ne se soit forcément amélioré dans le processus évalué.

Ensuite, il faut pouvoir représenter les capacités du patient dans la vie de tous les jours c'est-à-dire obtenir une mesure qui ait une validité dite « écologique », pour que le test soit le plus proche des performances réelles du patient. En effet, il semble superflu de connaître les capacités du patient dans un domaine qui ne va pas l'aider dans sa communication quotidienne. Dans cette optique, il a été décidé que l'évaluation par le MACS porterait sur la modalité orale du langage puisque celle-ci est prioritaire, par rapport au langage écrit, dans la vie quotidienne.

## **B. Le MACS**

### **1. Introduction**

Le MACS a pour objectif de mesurer l'évolution des capacités linguistiques et comporte trois versions différentes. Ce test a été conçu par Anna Zumbansen dans le cadre de sa thèse (Zumbansen, 2014 [31]), dans le laboratoire du Professeur Sylvie Hébert, pour analyser un éventuel bénéfice du chant dans la rééducation de l'aphasie.

Dans cette recherche, il a été nécessaire de créer un outil d'évaluation quantitative de la progression du discours naturel des patients aphasiques dans le but d'analyser l'efficacité de la Melodic Intonation Thérapie (MIT), ou Thérapie par Intonation Mélodique (Sparks et coll., 1974 [29]), méthode utilisant le chant pour améliorer le discours naturel des patients. Nous développons ci-après les mesures retenues comme indicateurs de la performance langagière ainsi que le cadre scientifique qui soutient ce choix.

### **2. Les mesures choisies**

Traditionnellement, les analyses du langage naturel d'adultes aphasiques ont surtout porté sur la façon dont le discours est conforme aux règles et aux modèles de langage standards plutôt qu'à la façon dont l'information est communiquée à l'interlocuteur (Yorkston et Beukelman, 1980 [30]). Cependant, les recherches qui souhaitent quantifier les changements du caractère informatif du langage naturel chez les personnes aphasiques, en réponse au traitement ou à la manipulation de variables expérimentales, ont été entravées par le manque de mesures standard pour caractériser cet aspect du langage naturel (Yorkston et Beukelman, 1980 [30]). Plusieurs auteurs se sont cependant penchés sur l'élaboration de mesures du discours naturel de personnes aphasiques.

La mesure d'Unité d'Information Correcte (UIC) a été introduite par Yorkston et Beukelman (1980 [30]). Une unité d'information correcte est composée d'un seul mot qui doit être exact, pertinent, instructif mais pas obligatoirement correct sur le plan grammatical. En mesurant le débit de parole (nombre de syllabes par minute) et le débit d'information (nombre d'UIC par minute), ils ont abordé le problème de vitesse de transmission de l'information chez les personnes aphasiques, à différencier d'avec le

nombre d'information transmises. La mesure d'UIC a été reprise par Nicholas et Brookshire (1993 [23]) qui utilisent trois mesures (nombre de mots par minute, pourcentage d'UIC, nombre d'UIC par minute) pour évaluer le caractère informatif et l'efficacité du langage naturel.

Les recherches ont démontré que le pourcentage d'UIC produites en condition d'échantillonnage structurée est un prédicteur très juste des performances en condition conversationnelle (Doyle et coll., 1995 [8]). De plus, les procédures d'incitation structurée associées aux mesures d'Informativité avec la mesure des UIC prédisent l'Informativité de la communication d'adultes aphasiques en interaction conversationnelle.

Le développement de la capacité à transmettre un message reste central en rééducation orthophonique du langage avec un patient aphasique. En effet, la première mission de l'orthophoniste est de permettre au patient de communiquer de manière efficace et le plus clairement possible. Pour cela, il peut utiliser de nombreux moyens comme le langage oral ou écrit, les gestes, l'aide du contexte. L'orthophoniste privilégie cependant l'amélioration du langage oral tant qu'elle est jugée possible. C'est pourquoi le MACS se centre sur l'information transmise par le langage oral.

Dans le MACS, l'Informativité est utilisée comme une des deux mesures quantitatives du discours naturel du patient. Ce score repose sur le nombre total de mots de l'échantillon produit et sur le nombre d'Unités d'Information Correcte que contient l'échantillon. Il est possible, grâce à ces deux variables, de calculer le pourcentage d'Informativité du discours, ou score d'Informativité, et l'analyse des scores permet de mesurer avec quelle efficacité le patient transmet l'information

Quand l'axe principal de prise en charge vise le langage oral et qu'une certaine quantité d'information parvient à être transmise, il devient nécessaire d'améliorer la forme du discours, en parallèle à l'information transmise, en faisant travailler les usages grammaticaux au patient. Les troubles morphosyntaxiques, ou agrammatisme, sont en effet souvent présents dans les aphasies. Il s'agit de perturbations du langage par l'omission ou l'utilisation erronée de morphèmes grammaticaux, un style télégraphique, l'omission de mots outils et une prédominance des mots à faible contenu dans le discours (Rivard, 2013). Dans le MACS, il a été décidé de mesurer l'adéquation grammaticale (Grammaticalité) afin de quantifier l'évolution du patient dans cette dimension de son discours naturel. Pour cela, le score de Grammaticalité est fondé sur les mots outils bien utilisés, mal utilisés ou absents. On considère comme mot outil tout mot qui n'est ni nom, ni adjectif, ni adverbe, ni verbe. Ce choix résulte de l'observation que ce sont majoritairement eux qui sont mal

utilisés en cas d'agrammatisme. Les erreurs audibles d'accords adjectivaux ou de conjugaison ne sont donc pas comptabilisées ici.

Ainsi, les mesures retenues pour le MACS sont l'Informativité et la Grammaticalité, dont les formules sont les suivantes :

**Score d'Informativité = 100 x Nombre d'Unités d'Information  
Correcte / Nombre de mots de l'échantillon**

**Score de Grammaticalité = 100 x Nombre de mots-outils bien utilisés  
de l'échantillon / Nombre de mots total de l'échantillon**

## **2. Les stimuli utilisés**

Lors d'un test, faire produire au participant un échantillon de discours naturel représentatif de ses capacités langagières est un défi. Pour cela, il faut déterminer quel moyen de stimulation fera produire au patient l'échantillon voulu. Plusieurs recherches se sont penchées sur la question. Nous allons évoquer ici les principales avancées de ces dernières années et expliquer quels stimuli ont été choisis pour le MACS.

Le MACS a été fondé sur une description d'images car cette tâche de production de discours a été démontrée comme faisant peu appel au système mnésique et aux fonctions exécutives (Doyle et coll., 1995 [8]). De plus, elle est standardisable et semble permettre la production d'un discours qui est représentatif des capacités de conversation naturelle des sujets (Doyle et coll., 1995 [8]). Enfin, l'utilisation d'une image comme support de discours dans une épreuve où le patient doit raconter une histoire qu'on lui a présentée privilégie une production orale plus efficace et présentant moins de perturbations. (Doyle, 1998 [9])

Après avoir choisi la meilleure modalité de stimulation, il reste à déterminer combien de stimuli vont être présentés au patient pour obtenir l'incitation de production de discours naturel la plus efficace possible. Selon une étude (Nicholas et Brookshire, 1994 [4]),

présenter plusieurs stimuli au lieu d'un seul et donc obtenir un plus grand échantillon de discours permettrait d'obtenir des scores plus stables dans le temps, ce qui augmente la fiabilité test-retest. Puisque les capacités des personnes aphasiques peuvent fortement varier, les échantillons ainsi obtenus sont plus représentatifs des capacités du patient, ce qui est fondamental dans la validité et la fiabilité d'un test. Dans une autre étude, Nicholas et Brookshire (1994 [3]) ont cherché à connaître le nombre de stimuli minimum à partir duquel on obtient un échantillon avec une bonne stabilité test-retest et qui soit significativement représentatif des capacités langagières. Ils ont conclu que l'échantillon était suffisamment long à partir d'un groupe de quatre à cinq stimuli, l'augmentation de la fiabilité étant moindre au delà de cinq stimuli. Dans le MACS, il a donc été choisi d'utiliser cinq images par passation pour récolter un échantillon représentatif des capacités linguistiques du sujet.

Les images du MACS présentent des situations de vie quotidienne dans différents environnements avec plusieurs personnages (voir Annexes A à H). Elles sont dessinées au trait, en noir et blanc et ont été réalisées par l'illustratrice Claudine Parra. Chaque image présente une situation, un environnement, des personnages à chaque fois différents ainsi qu'un élément saillant de façon à faire produire le plus possible de langage naturel aux participants lors de leur description. Ainsi, afin de contrôler l'effet de familiarité, les scènes proposées sont des scènes de vie quotidienne. La présence de plusieurs personnages (trois à sept personnages selon les images) en action permet de favoriser la production de verbes d'action. Les personnages sont en interaction avec leur environnement pour inciter la production de phrase type sujet-verbe-complément.

Dans leur conception, les quinze images ont été pensées pour être réparties en trois groupes afin de créer des versions parallèles du test de description d'images (Tableau 2).

En somme, le MACS propose une tâche de description d'images représentant des scènes de la vie quotidienne. Le patient a pour consigne "Racontez-moi tout ce qui se passe sur cette image". Ce test comporte trois versions de cinq images chacune pour pouvoir évaluer l'évolution du discours du patient tout en évitant un éventuel biais d'apprentissage. Après avoir recueilli et transcrit l'échantillon de discours naturel du patient, l'orthophoniste doit évaluer les deux niveaux d'analyse du discours que sont l'Informativité et la Grammaticalité.

## **C. Propriétés psychométriques du MACS**

Lors de la création d'un nouvel outil d'évaluation, il est nécessaire d'éprouver ses propriétés psychométriques, afin d'assurer une qualité objective aux cliniciens et aux chercheurs qui l'utiliseront. Cela crédibilise le test et permet de déterminer sa finesse et sa qualité. Une standardisation rigoureuse du test permet l'application de procédures fiables. Les informations concernant les méthodes statistiques utilisées pour l'élaboration de l'évaluation doivent être mises à la disposition des cliniciens (dans le manuel d'utilisation). L'utilisateur pouvant se référer à la fiabilité et à la finalité de l'outil à tout moment, ses inférences concernant les résultats au test seront renforcées (Ivanova et Hallowell, 2013).

Tel que l'ont souligné Ivanova et Hallowell (2013), très peu de tests destinés à l'aphasie ont des preuves psychométriques de validité, fiabilité et de sensibilité au changement. Nous présentons ici les différentes mesures de la validité et de la fiabilité détaillées dans la littérature, en précisant lesquelles ont déjà été réalisées et lesquelles sont pertinentes dans le cas du MACS, en fonction de son objectif : évaluer l'évolution des capacités langagières. La sensibilité au changement est un point fondamental à aborder dans ce test, contrairement au cas des évaluations à visée diagnostique déjà existantes dans le domaine de l'aphasie.

### **1. Validité**

#### **a. Validité d'apparence**

La validité d'apparence est accordée à un test par un utilisateur après examen des aspects visibles de l'évaluation tels que la longueur, les items, les modalités des stimuli et des réponses. Ce jugement est subjectif et reste superficiel, il est une première étape à la validation d'un test. Cette validité ne suffit pas à déterminer si l'instrument est valide ou non, l'examen des autres types de validité sont nécessaires (Fermanian 2005 [10]).

Concernant le MACS, la validité apparente a été jugée par des orthophonistes qui répondaient aux questions suivantes : " Le MACS est-il facile d'utilisation ? " ; " Pensez-vous que cet outil évalue correctement le langage dans le discours naturel des personnes aphasiques ? " ; " Les images favorisent-elles une production verbale spontanée ? " .

Les professionnels ont conclu que le MACS se révèle en apparence une mesure pertinente du discours naturel des personnes aphasiques. En outre, lors des premières utilisations du

MACS durant les essais pilotes puis avec davantage de participants aphasiques et sains, il a été relevé des commentaires et réactions démontrant que les images suscitaient, comme espéré, un intérêt significatif et une production du discours de la part des participants.

### **b. Validité de contenu**

La validation de contenu d'un test consiste à vérifier que les différents aspects du test sont représentatifs du concept qu'il cherche à évaluer. Pour cela, on se focalise sur le matériel proposé mais aussi sur les consignes, les modalités dans lesquelles le matériel peut être présenté, la méthode de cotation et tous les éléments qui rentrent en compte dans la passation et qui pourraient potentiellement induire un biais. Pour valider le contenu d'un test, les experts doivent définir précisément et dans son intégralité le concept visé. Cette définition initiale sert alors de socle pour déterminer quels aspects du concept sont importants dans le test, ainsi que les variables pouvant influencer les résultats. La validité de contenu est accordée ou non en tenant compte du concept visé, de la population ciblée et de la fonction du test. Cette validité est déterminée par des juges à un moment précis, elle reste donc relative à l'instant et à l'endroit où elle a été faite (Laveault et Grégoire, 2002 [17]).

Pour le MACS, deux caractéristiques ont été choisies afin de refléter les capacités du sujet à utiliser le langage dans son discours naturel : l'Informativité et la Grammaticalité. Ces caractéristiques ont pour origine des travaux antérieurs menés sur une trentaine d'années et sont reconnues pour être représentatives du langage des individus aphasiques (Yorkston et Beukelman, 1980 [30] ; Nicholas et Brookshire, 1993 [23] ; Doyle et coll., 1995 [8] ; Rivard, 2013 [25]). Le MACS étant établi sur de multiples recherches menées par différentes équipes concernant notamment le type de tâche, le nombre de stimuli ou la méthode de cotation, sa validité de contenu est supposée bonne. Cependant, le processus de validation devra être réalisé formellement en recherchant le consensus des experts concernant les items, la consigne, les modalités des stimuli et des réponses, les contraintes de temps, les critères de cotation, ceci tout en tenant compte du concept ciblé.

### **c. Validité de construit**

Lors de la création d'un outil de mesure, il faut définir un cadre conceptuel précisant les caractéristiques du concept visé et générant des hypothèses sur les liens présents ou non entre ce qu'on mesure et d'autres concepts existants. La validité de construit provient du fait que l'on peut émettre des hypothèses concernant le concept mesuré. Ces hypothèses sont ensuite étudiées par l'expérimentation et l'analyse statistique, afin de juger si elles sont confirmées ou non (Anastasi 1994 [1]). L'hypothèse peut être affirmative (on suppose qu'un lien existe) ou négative (on suppose qu'il n'existe pas de lien). On parle alors respectivement de validité convergente et de validité divergente. Si, après expérimentation, une ou plusieurs hypothèses vont dans le sens que l'on attendait, alors il est dit que la validité de construit est renforcée. En revanche, une hypothèse importante infirmée peut parfois faire fortement douter de la validité de construit du test. Ce processus de validation s'établit au fur et à mesure que des hypothèses sont formulées et vérifiées. Il n'est jamais terminé car d'autres hypothèses peuvent être posées à tout moment (Fermanian 2005 [10]).

Dans le cas où l'on veut créer un test pour en remplacer un autre (dans le but, par exemple, de mesurer le même phénomène mais avec un temps de passation et de cotation réduit) il faut évaluer la validité contre-critère, qui est une validité de construit convergente permettant de vérifier que les deux variables reflètent le même concept. Si on observe une forte corrélation entre les deux mesures, alors le nouveau test peut en effet remplacer l'ancien (Ivanova et Hallowell, 2013 [15]).

La validité contre-critère ne s'applique pas pour le MACS pour le moment car aucun test orthophonique déjà créé ne vise l'évaluation de l'évolution des capacités langagières chez les personnes aphasiques. Les autres formes de validité de construit (convergentes et divergentes) sont donc très importantes à renforcer. Celles-ci ont ainsi été récemment testées par Frachon et Quiquempois (2015). D'autres recherches devront poursuivre ces travaux.

## **2. Fiabilité**

L'estimation de la fiabilité d'un test permet de savoir dans quelle mesure les résultats obtenus sont cohérents. De nombreux facteurs peuvent amener les scores à varier, comme le risque d'erreur, les différences entre les participants ou encore les différences induites par les juges s'occupant de la cotation (Shewan, 1988 [28]). Il est dit d'une fiabilité qu'elle est élevée lorsque les scores obtenus chez les participants sont proches lorsqu'une variable est fixe. Pour évaluer la distance entre les scores, on doit utiliser une mesure quantitative ou qualitative en fonction de la nature de la mesure réalisée. L'évaluation de la proximité entre les scores est possible grâce à un coefficient de corrélation, compris entre 0 et 1. Plus le coefficient est proche de 1, plus la corrélation est forte, donc la variance est davantage expliquée par la variance des scores réels au test (Fermanian 2005 [10]).

Si la fiabilité est bonne, alors les sujets obtiennent des scores similaires lors de deux passations du test par un participant, malgré la variation d'un ou plusieurs paramètres non contrôlables. La variabilité due à ces paramètres est l'erreur de mesure. Les scores à un test sont toujours affectés par l'erreur de mesure, même dans des conditions de passation idéales (Anastasi 1994 [1]).

La fiabilité d'un test ne peut donc pas être parfaite par le fait même de l'existence de l'erreur de mesure. Cependant, un test avec une bonne fiabilité permet de réduire la marge d'erreur possible dans l'interprétation des résultats et assure une stabilité dans les résultats (Laveault et Grégoire, 2002 [17]). Elle doit donc être testée dans plusieurs situations expérimentales, que nous abordons ci-après.

### **a. Fiabilité inter-juges**

Cette fiabilité est éprouvée lorsque le score de chaque patient est coté au même moment mais indépendamment par deux juges différents (Fermanian 2005 [10]). L'objectif est de déterminer dans quelle mesure le juge fait varier le score obtenu par le participant, en calculant le coefficient de corrélation entre les différents scores cotés par les juges au même test pour le même patient. Lorsque les instructions du guide de cotation sont précises, cela permet d'augmenter ce coefficient et donc de diminuer la part de variance imputable à l'examineur (Ivanova et Hallowell, 2013 [15]). Un test avec une fiabilité inter-juge forte permet de diminuer le risque d'erreur de mesure due à la subjectivité de l'évaluateur.

Dans le MACS, la fiabilité inter-juge a été prouvée sur cent quatorze échantillons de discours, cotés indépendamment par deux juges. Les coefficients de corrélation de l'Informativité et de la Grammaticalité ont révélé une fiabilité inter-juge satisfaisante à la fois pour le score d'Informativité ( $r = .84$  ;  $R^2 = .70$ ) et le score de Grammaticalité ( $r = .87$  ;  $R^2 = .75$ ) (Zumbansen, 2012 [31]).

### **b. Fiabilité des versions parallèles**

La fiabilité des versions parallèles permet d'assurer que deux formes parallèles du même outil sont équivalentes (Fortin 1994 [12]).

Pour déterminer si cette fiabilité est forte, il faut faire passer les versions différentes du test à un même individu, coté par un même examinateur, en évitant de faire varier d'autres éléments tels que le cadre dans lequel le sujet est soumis aux deux passations. Ensuite, un coefficient de corrélation est calculé entre les scores de chacune des versions, afin de déterminer le coefficient de fiabilité du test.

Le MACS a été élaboré avec des versions parallèles. L'équivalence de ces versions a été vérifiée statistiquement sur un échantillon de soixante-dix personnes de différents âges (Milette-Monier, 2012 [20]).

La fiabilité par la méthode de la bissection peut être une forme de fiabilité des versions parallèles. Elle permet, lorsqu'un test n'a pas de version parallèle préconçue, de tester la corrélation des scores entre les deux moitiés d'un test. La stabilité dans le temps n'entre alors pas en jeu, puisque le sujet ne fait qu'une passation (Anastasi 1994 [1]).

Le MACS étant un test avec des versions parallèles, cette démarche n'a pas été nécessaire.

### **c. Fiabilité test-retest**

La fiabilité test-retest est la mesure de la fiabilité dans le temps. Elle requiert donc deux passations (Ivanova et Hallowell, 2013 [15]). Elle permet de s'assurer que, si un patient réalise deux passations identiques à deux moments différents, les scores resteront significativement stables. Cela garantit que si toutes les variables autres que le temps sont stables, les résultats sont similaires et donc que le test ne varie que si un autre élément peut l'expliquer.

Elle est calculée en effectuant un coefficient de corrélation entre les scores de chaque passation. Plus le laps de temps est long entre les deux passations, plus la corrélation risque d'être faible entre les deux scores. Ceci est particulièrement vrai si les passations s'effectuent auprès d'un patient aphasique car l'aphasie est un trouble qui peut évoluer rapidement dans le temps (Ivanova et Halloxell, 2013 [15]).

Lorsque la fiabilité test-retest est évaluée, plusieurs éléments sont à prendre en compte.

Tout d'abord, il faut apprécier le biais d'apprentissage, puisque le sujet effectue deux passations du même test séparées par un intervalle de temps. Plus l'intervalle est court, plus l'effet d'apprentissage peut biaiser la seconde passation, et donc les scores ainsi que le coefficient de corrélation entre ceux-ci. En ce qui concerne le MACS, une version parallèle du test a été utilisée, ce qui a permis d'éviter un fort effet d'apprentissage.

Ensuite, on doit vérifier que les conditions d'évaluation soient les plus identiques possibles lors des deux passations, pour diminuer les sources de variations non contrôlables.

Il faut également pouvoir présupposer que les participants n'aient pas des capacités fluctuantes. C'est pourquoi la fiabilité test-retest du MACS a été testée sur un échantillon de population dont les aptitudes avaient été reconnues comme stables dans le temps.

Enfin, il faut éviter le biais de l'examineur en faisant coter le test "à l'aveugle" afin que les participants restent anonymes et que l'ordre des passations ne soit pas connu des juges. Concernant le MACS, les cotations ont pu se faire à l'aveugle grâce à l'enregistrement sonore des échantillons de discours.

La fiabilité test-retest du MACS a été évaluée lors du travail effectué l'année passée dans le cadre d'un mémoire d'orthophonie (Frachon et Quinquempois, 2015). Lors de cette étude, la fiabilité test-retest a pu être prouvée pour le score d'Informativité et pour celui de Grammaticalité. Les scores de personnes aphasiques supposées stables (faisant preuve d'une pathologie au stade chronique et n'étant pas suivies sur le plan orthophonique) se sont révélés significativement corrélés sur des passations avec deux semaines d'intervalle.

### 3. Sensibilité au changement

Un instrument de mesure est dit sensible au changement s'il peut évaluer précisément les variations en plus ou en moins d'une aptitude cliniquement manifeste. La sensibilité au changement est le concept théorique et statistique sur lequel s'appuie la démonstration de la modification d'un état (Fermanian 2005 [10]). Cette sensibilité n'est pas nécessaire pour les outils diagnostiques tels que les tests existants pour les patients aphasiques francophones mais elle est un aspect fondamental pour un test d'évaluation standardisée de progression langagière tel que le MACS. Ce dernier, créé pour objectiver une amélioration linguistique quelle que soit la thérapie suivie par le patient, sera aussi utile lors de la comparaison de l'efficacité de deux méthodes de rééducation différentes.

Le MACS ne sert pas à décrire de manière exhaustive les aspects atteints et conservés du langage du participant. Il est destiné à relever dans quelles mesures les perturbations dues à l'aphasie auront des répercussions dans le discours naturel du patient et vise à quantifier les changements existants de compétences linguistiques.

Pour tester la sensibilité au changement d'un outil, il est nécessaire de sélectionner un échantillon de participants réputés "en progrès". Pour cela, on peut se fier à une amélioration observée dans une mesure proche du construit du test.

Frachon et Quinquempois (2015 [13]) ont vérifié la sensibilité du MACS au changement en sélectionnant un échantillon de patients aphasiques suivant une rééducation orthophonique intensive et qui s'étaient améliorés à un test de dénomination, le *Boston Naming Test* (Kaplan et coll., 2001 [16]). En utilisant les versions parallèles et en mettant en place des conditions de passation semblables pour éviter de mesurer un effet d'apprentissage, ainsi qu'en utilisant une cotation à l'aveugle afin de diminuer le biais de la subjectivité du juge, le score d'Informativité a été relevé et analysé. L'évolution de ce score se révèle en corrélation significative avec celui de la dénomination, l'hypothèse de la sensibilité au changement du score d'Informativité est prouvée convergente.

Une mesure des habiletés grammaticales, compréhension de phrases du *Token Test* (McNeil et Prescott, 1978 [24]), a aussi été effectuée sur cet échantillon de population, mais l'évolution du score de Grammaticalité n'a pas été démontrée significativement corrélée avec l'évolution des habiletés grammaticales.

Des hypothèses divergentes entre l'Informativité et les habiletés grammaticales ainsi qu'entre la Grammaticalité et la dénomination ont été émises, afin de déterminer si le MACS fait preuve de spécificité dans les changements mesurés.

L'évolution de l'Informativité présente une corrélation divergente avec celle du score aux

habiletés grammaticales, c'est donc un indice de mesure de changement spécifique.

Une étude sur la spécificité du score de Grammaticalité doit encore être menée.

Pour conclure, le processus de validation du MACS est en cours.

La validité d'apparence a été démontrée et nous pouvons dire que la validité de contenu est appuyée par les recherches précédant la création du MACS, mais il est nécessaire de soumettre le test au jugement d'experts afin de confirmer une bonne validité de contenu. La validité de construit a été explorée lors de la recherche de Frachon et Quinquempois (2015 [13]). Ces auteurs ont pu constater une corrélation significative du pourcentage d'amélioration en Informativité avec le pourcentage obtenu au score de dénomination mais aucune corrélation significative relevée avec celui de la compréhension de phrases au *Token Test*. La validité de construit a donc pu être renforcée lors de l'examen de la sensibilité au changement. Le MACS n'ayant pas été créé dans un but de diagnostic, la validité diagnostique ne s'applique pas à ce test.

Pour ce qui est de la fiabilité, la fiabilité inter-juges a été démontrée pour le MACS.

La fiabilité des versions parallèles a été évaluée lors de la création des différentes versions. L'analyse des résultats lors de l'établissement des différentes données normatives permet d'observer si les trois versions sont significativement équivalentes et la fiabilité test-retest a été démontrée pour le score d'Informativité comme pour le score de Grammaticalité.

Enfin, la sensibilité au changement du score d'Informativité et sa spécificité ont été démontrées. La sensibilité au changement de celui de Grammaticalité reste à déterminer.

#### **D. Nécessité d'établir des normes**

Afin de pouvoir dire d'un test qu'il est étalonné il faut établir ses normes. Pour cela, un échantillon de population passe le test et les scores de ce groupe font office de référence.

Les normes sont constituées des résultats du groupe de référence (Anastasi, 1994 [1]).

Le choix de la population de référence n'est pas fait par hasard, le chercheur doit définir cette population de personnes en tenant compte de l'objectif de l'évaluation. Ce groupe d'individus se doit d'être homogène afin que toute personne susceptible d'être testée y soit incluse. Toute la population de référence ne pouvant généralement pas être testée, les tests sont souvent basés sur un échantillon plus ou moins grand. On relève les différents scores de l'échantillon choisi, puis les moyennes et variances sont calculées afin de situer ultérieurement les scores individuels des sujets qui passeront le test et déterminer s'ils sont ou non dans la norme (Laveault et Grégoire, 2002 [17]).

Les normes réalisées avec une population sont spécifiques à sa culture et à son utilisation de la langue. Les normes ne peuvent donc pas servir à analyser les résultats au test d'un individu évalué dans un langage différent de la population de référence, même s'il présente les mêmes troubles langagiers. Les scores chez deux personnes n'ayant pas la même langue ne sont pas comparables car leurs performances sont influencées par les caractéristiques psycholinguistiques de leur langue maternelle. Cet aspect représente un défi lors de la traduction d'un test déjà existant (Ivanova et Hallowell, 2013 [15]).

En ce qui concerne le MACS, l'échantillon de population témoin de référence a été divisé en trois tranches d'âge différentes. Les normes des tranches d'âge 18-30 ans et 51-70 ans ont été établies avant notre étude (voir Tableau 3). Il restait donc à effectuer la passation du test pour les individus de la population de référence entre 31 et 50 ans, puis de transcrire leurs discours, de relever les scores d'Informativité et de Grammaticalité et, enfin, d'en faire les moyennes et écart-types qui serviront de repères pour le MACS lors de ses prochaines utilisations.

# PARTIE PRATIQUE

## 1. Objectif

Notre objectif était d'établir les normes des témoins francophones québécois de la tranche d'âge 31-50 ans du MACS.

## 2. Méthodologie

### a. Participants

Trente-neuf participants ont été recrutés pour l'élaboration des normes 31-50 ans du MACS. Parmi eux, vingt-deux femmes et dix-sept hommes. Ils devaient avoir entre 31-50 ans au moment de l'évaluation. Ils étaient tous québécois francophones avec le français comme langue maternelle. Les participants n'ont pas été retenus pour l'étude s'ils avaient fait un accident cérébral vasculaire ou une dépression dans la dernière année. Un trouble du langage, un trouble cognitif, un trouble visuel non corrigé étaient des critères d'exclusion également. Les participants devaient avoir un score minimal de 26 sur 30 au *Mini Mental State Evaluation* (MMSE), (Folstein et McHugh, 1975 [11]). Le recrutement des participants a été réalisé au moyen d'affiches sur le campus de l'Université de Montréal, ainsi que par l'envoi d'un courriel à la banque de participants du BRAMS.

Un participant a été exclu avant la transcription de son échantillon car il ne respectait pas le critère d'inclusion du français comme langue maternelle. Deux participants ont eu des scores respectifs de 25/30 et 24/30, ils ont donc été exclus. Le nombre de participants a finalement été de trente-six, avec vingt-deux femmes et quatorze hommes (voir Tableau 1). La moyenne d'âge était de 37,7 (ET= 5,8) et la moyenne du nombre d'années d'étude était de 17,9 (ET= 3,4).

**Tableau 1 - Caractéristiques des participants de la tranche d'âge 31-50 ans**

<b>Population</b>	<b>Nombre</b>	<b>Age</b>	<b>Sexe</b>		<b>Latéralité</b>		<b>Nb années scolarité</b>	<b>Score MMSE</b>
<b>Sujets québécois francophones sains</b>	36	M= 37,7	22F	14H	2G	34D	M=17,9	M= 28,9
		ET = 5,8					ET= 3,4	ET= 1,3

## **b. Matériel**

Un questionnaire a été proposé aux participants pour collecter les informations nécessaires pour leur inclusion à cette étude.

Un test de dépistage des troubles cognitifs, le MMSE, était également nécessaire pour vérifier l'absence de trouble cognitif des participants.

Le MACS est composé de quinze images différentes imprimées en noir et blanc sur une feuille classique en format lettre US (27,94 x 21,59 format standard nord-américain). Les quinze images représentent des situations de vie quotidienne dans différents environnements avec plusieurs personnages. Chaque image présente une situation, un environnement, des personnages à chaque fois différents ainsi qu'un élément saillant. Dessinées par l'illustratrice Claudine Parra, les images ont été élaborées de façon à faire produire le plus possible de langage naturel aux participants lors de leur description. Ainsi, afin de contrôler l'effet de familiarité, les scènes proposées sont des scènes de vie quotidienne. La présence de plusieurs personnages (trois à sept personnages selon les images) en action permet de favoriser la production de verbes d'action. Les personnages sont en interaction avec leur environnement pour inciter la production de phrases type sujet-verbe-complément. Enfin, un élément saillant a été ajouté à chaque scène pour une plus grande richesse du discours.

Dans leur conception, les quinze images ont été pensées pour être réparties en trois groupes afin de créer des versions parallèles du test de description d'images. Trois groupes de cinq images ont été conçus. Le choix du nombre d'images par groupe a été fait en fonction des recherches effectuées dans la littérature. Tel que nous l'avons mentionné dans la section *Les stimuli utilisés*, un meilleur rapport effort/stabilité est retrouvé avec quatre à cinq stimuli (Nicholas et Brookshire, 1994 [3]). Le Tableau 2 présente la répartition des quinze images en trois groupes d'images.

Pour les passations, un enregistrement audio a été effectué grâce à un dictaphone.

Le logiciel Microsoft Word a été utilisé pour la transcription des échantillons.

Le logiciel Cordial Analyseur (Synapse- développement, 2010) a permis de réaliser la cotation à partir des échantillons (nombre total de mots, pourcentage de mots outils par rapport au nombre total de mots).

Le logiciel SPSS a été utilisé pour l'analyse statistique.

**Tableau 2 - Répartition des quinze images en trois groupes d'images.**

Groupe	Image (Numéro)	Nombre de personnages vivants	Thème
1	Cuisine (5)	3	Vie de famille dans la maison
	Épicerie (6)	4	Courses en ville
	Fête (7)	5	Sorties-Loisirs
	Parc (9)	6	Sorties de ville
	Plage (11)	7	Vacances
2	Chambre (3)	3	Vie de famille dans la maison
	Restaurant (12)	4	Sorties de ville
	Rue (13)	5	Courses en ville
	Campagne (2)	6	Vacances
	Zoo (15)	7	Sorties-Loisirs
3	Pêche (10)	3	Sorties-Loisirs
	Banque (1)	4	Courses en ville
	Montagne (8)	5	Vacances
	Salon (14)	6	Vie de famille dans la maison
	Cinéma (4)	7	Sorties de ville

### **c. Déroulement**

Pour établir les normes, il était nécessaire de réaliser la passation des trois versions de l'épreuve avec les trente-six participants retenus. L'enregistrement des passations permettait alors la transcription des échantillons pour en extraire par la suite les échantillons interprétables et informatifs. Grâce au logiciel Cordial Analyseur, les scores d'Informativité et de Grammaticalité ont pu être extraits et les moyennes et écarts-types, nécessaires à la création des normes, calculés.

L'examineur et le participant étaient installés dans une pièce fermée et calme pour la bonne concentration du participant et la qualité de l'enregistrement audio. Ils étaient assis face à face à une table. Le questionnaire était donné au participant et une fois rempli par celui-ci, l'enregistrement audio avec un dictaphone était lancé. Le test débutait avec le test cognitif (MMSE). La seule consigne donnée était de répondre aux questions posées. Pour la bonne réalisation du test, toutes les réponses données, correctes ou fausses, étaient notées par l'examineur et la feuille du test était pliée de façon à être hors de vue du participant. Le test de langage commençait ensuite avec la description des quinze images. L'examineur montrait chaque image au participant, l'une après l'autre. La consigne simple « Racontez-moi tout ce qui se passe sur cette image » était réitérée avant chaque image. Aucune autre indication n'était donnée au participant. À la fin de chaque description, l'examineur demandait : « Avez-vous autre chose à ajouter ? ». Cette question n'était posée qu'une seule fois après chaque description d'image. Lors de ce test, l'examineur écoutait calmement le participant sans prendre de notes, ni risquer de donner des indications non-verbales qui pouvait perturber le participant dans sa tâche. De plus, avec quelques participants parmi les plus jeunes ou les personnes familières, un tutoiement poli était utilisé pour créer une ambiance agréable de passation.

Enfin le questionnaire d'inclusion était rempli avec l'aide de l'examineur pour s'assurer de la bonne compréhension des questions. L'enregistrement était ensuite arrêté et sauvegardé.

#### **d. Analyses**

L'enregistrement des passations a permis la transcription des échantillons de langage de chaque participant. Les quinze échantillons de chaque participant ont été transcrits avec des règles préétablies afin d'avoir un échantillon brut à partir duquel travailler.

Chaque échantillon a subi une correction pour en extraire l'échantillon interprétable par le logiciel Cordial Analyseur. La correction consistait à supprimer les hésitations, les amorces, les segments ininterprétables et à la correction des mots déformés ou inachevés. L'ensemble de ces modifications a été noté dans un tableau Excel prévu à cet effet.

Toujours avec des règles préétablies, les mots outils absents ou mal utilisés dans chaque échantillon ont été comptés et ajoutés au tableau Excel.

Pour extraire l'échantillon informatif de chaque description d'image, ont été supprimés les segments ne concernant pas l'image ou erronés, toujours en suivant des règles bien précises.

Pour chaque participant, chaque échantillon par image a été modifié pour en extraire deux nouveaux échantillons un interprétable et un informatif. Finalement, 1080 fichiers ont été créés (36 participants x 15 images x 2 échantillons). L'analyse des échantillons par le logiciel Cordial Analyseur a permis d'obtenir le nombre total de mots et la proportion de mots-outils pour chaque échantillon. Les scores d'Informativité et de Grammaticalité pour chaque image et participant ont ainsi pu être calculés avec les formules suivantes :

$$\text{Informativité} = 100 \times \text{UIC} \div \text{Mots interprétables}$$

$$\text{Grammaticalité} = 100 \times \text{Mots-outils corrects} \div \text{Mots interprétables}$$

La moyenne et l'écart-type des scores obtenus ont été calculés pour chaque image et groupes d'image. Pour chaque image, la normalité de la distribution de chacun des trois scores a été vérifiée à l'aide des tests statistiques de Kolmogorov-Smirnov et Shapiro-Wilk (Shapiro et Wilk, 1965 [27]).

### 3. Résultats

Les statistiques descriptives de chacun des scores pour chaque groupe d'images sont indiquées dans les Tableaux 3 et 4.

La moyenne d'Informativité pour le premier groupe d'images est  $M= 73,4$  ( $ET= 9,3$ ), pour le deuxième groupe d'images,  $M= 73,4$  ( $ET= 8,8$ ) et pour le troisième groupe d'images,  $M= 75,6$  ( $ET= 8,6$ ).

La moyenne de Grammaticalité pour le premier groupe d'images est  $M= 53,8$  ( $ET= 2,6$ ), pour le deuxième groupe d'images,  $M= 53,3$  ( $ET= 2,4$ ) et pour le troisième groupe d'images,  $M= 53,2$  ( $ET= 2,4$ ).

Concernant la distribution des scores d'Informativité pour chaque groupe d'images pris séparément (voir Figures 1, 2, 3), aucun n'est significativement différent d'une distribution normale ( $p>0,05$ ). Une valeur  $p$  inférieure à 0,05 obtenue aux deux tests de normalité (Kolmogorov-Smirnov et Shapiro-Wilk) indiquerait que la distribution est significativement différente d'une distribution normale. Toutes les valeurs de  $p$  se situent au-delà de 0,005.

Concernant la distribution des scores de Grammaticalité pour chaque groupe d'images (Figures 4, 5, 6) aucun n'est significativement différent non plus d'une distribution normale. Toutes les valeurs de  $p$  se situent au-delà de 0,005.

Le Tableau 3 présente les normes des témoins francophones québécois établies pour les trois tranches d'âge 18-30 ans, 31-50 ans et 51-70 ans.

**Tableau 3 - Caractéristiques des témoins francophones québécois du MACS.**

<b>Témoins francophones québécois</b>	<b>Nombre</b>	<b>Age</b>	<b>Sexe</b>		<b>Nb années scolarité</b>	<b>Score MMSE</b>
<b>18-30 ans</b>	31	M= 22,87	20 F	11 H	M= 15,87	M= 28,84
		ET= 3,71			ET= 2,45	ET= 1,16
<b>31-50 ans</b>	36	M= 37,7	22 F	14 H	M= 17,9	M= 28,9
		ET= 5,8			ET= 3,4	ET= 1,3
<b>51-70 ans</b>	34	M=61	22 F	12 H	M= 14,94	M = 28,85
		ET= 5,3			ET= 3,91	ET= 1,28

**Tableau 4 - Statistiques descriptives et distribution des scores d'Informativité par groupe d'image.**

Scores d'informativité	Moyenne	Ecart-type	Etendue	Significativité (p) des tests de normalité	
				Kolmogorov-Smirnov	Shapiro-Wilk
<b>Groupe 1</b>	74,3	9,3	45,9 - 90,2	.200	.118
<b>Groupe 2</b>	73,4	8,8	58,0 - 89,82	.200	.260
<b>Groupe 3</b>	75,6	8,6	61,4 - 90,92	.097	.063

**Tableau 5 - Statistiques descriptives et distribution des scores de Grammaticalité par groupe d'image.**

Scores de Grammaticalité	Moyenne	Ecart-type	Etendue	Significativité (p) des tests de normalité	
				Kolmogorov-Smirnov	Shapiro-Wilk
<b>Groupe 1</b>	53,8	2,6	49,1 - 58,8	.200	.591
<b>Groupe 2</b>	53,3	2,4	48,7 - 58,6	.200	.827
<b>Groupe 3</b>	53,2	2,4	48,2 - 58,1	.200	.967

Figure A- Distribution des scores d'Informativité pour le groupe 1.

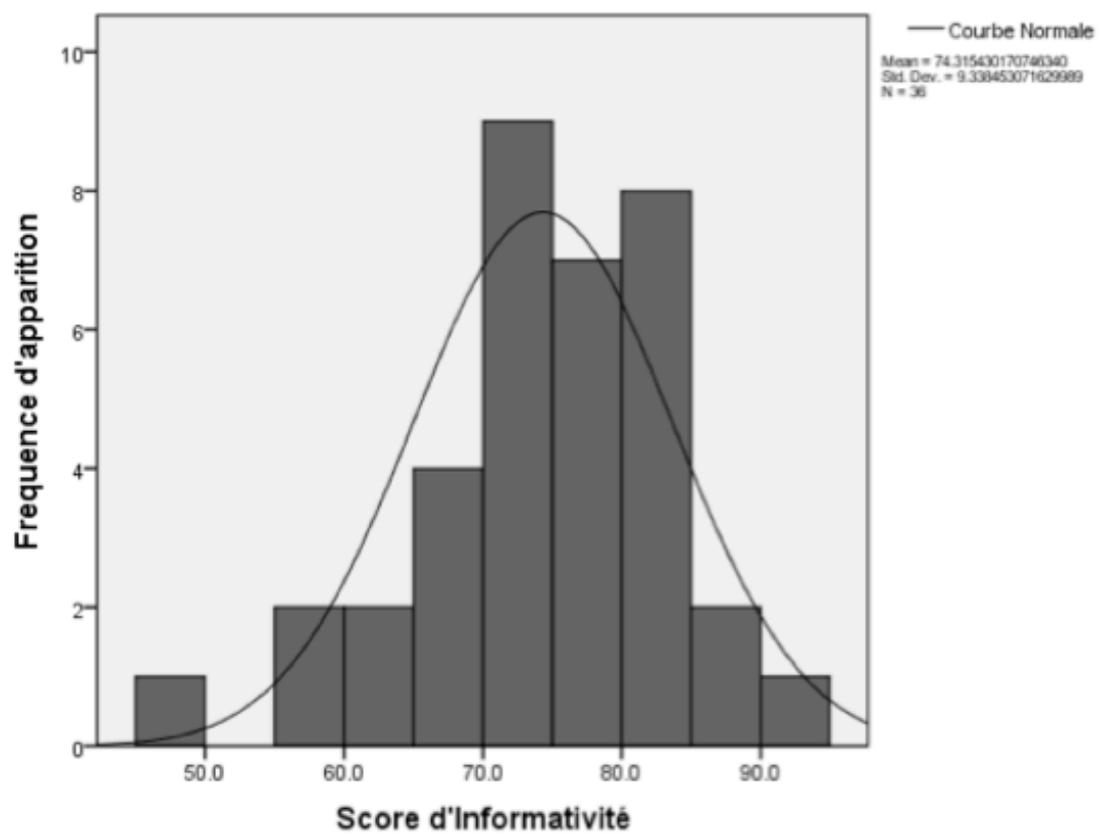


Figure B - Distribution des scores de Grammaticalité pour le groupe 1.

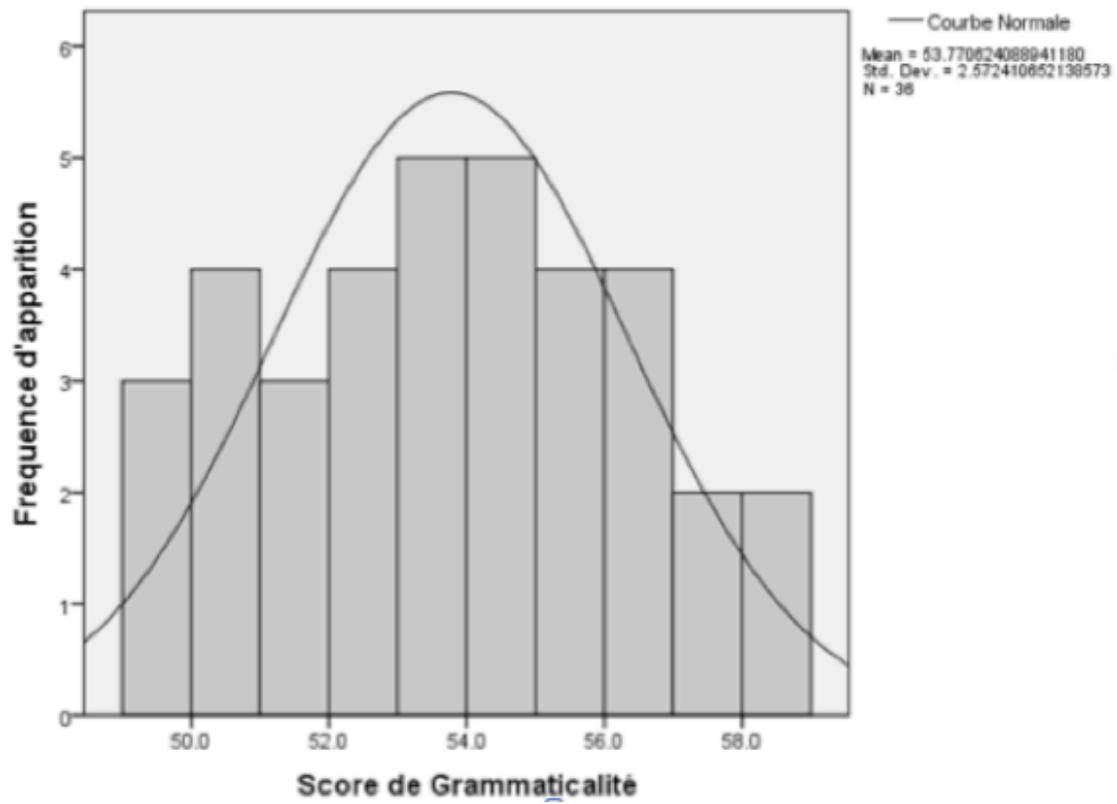
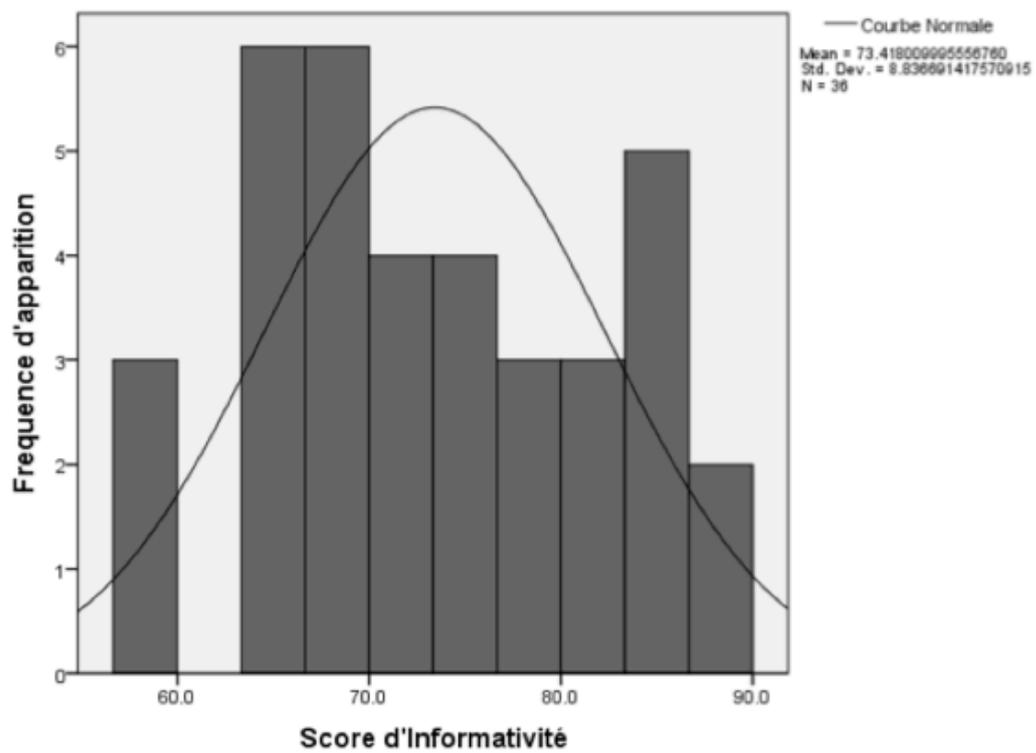


Figure C - Distribution des scores d'Informativité pour le groupe 2.



**Figure D - Distribution des scores de Grammaticalité pour le groupe 1.**

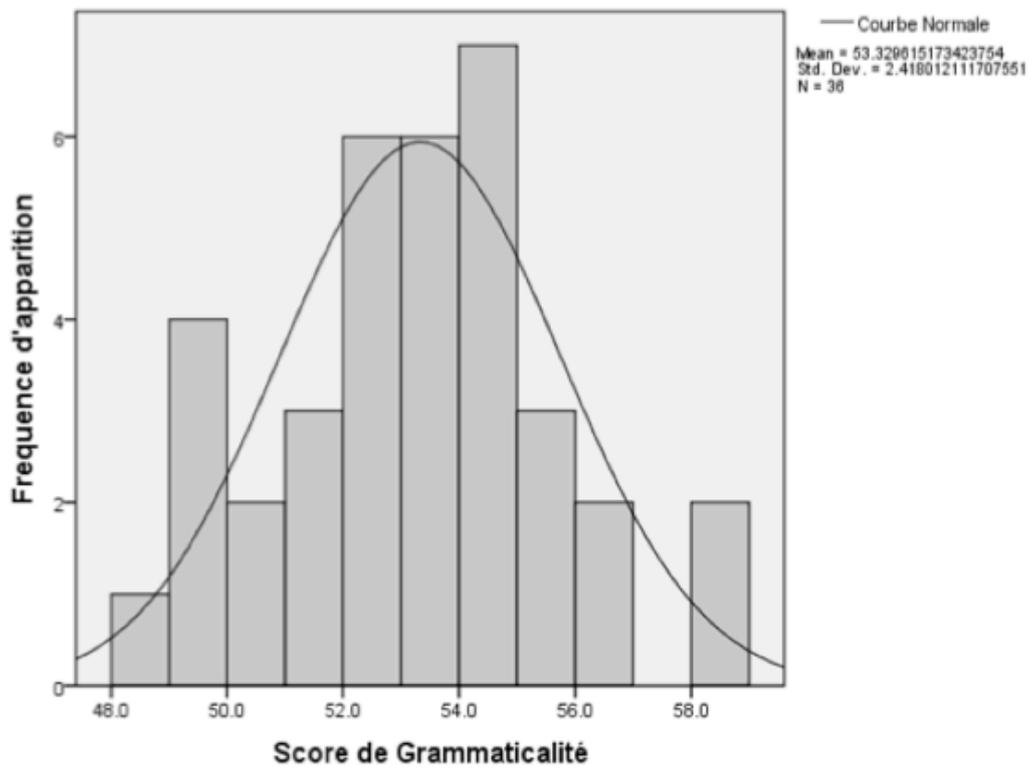


Figure E - Distribution des scores d'Informativité pour le groupe 3.

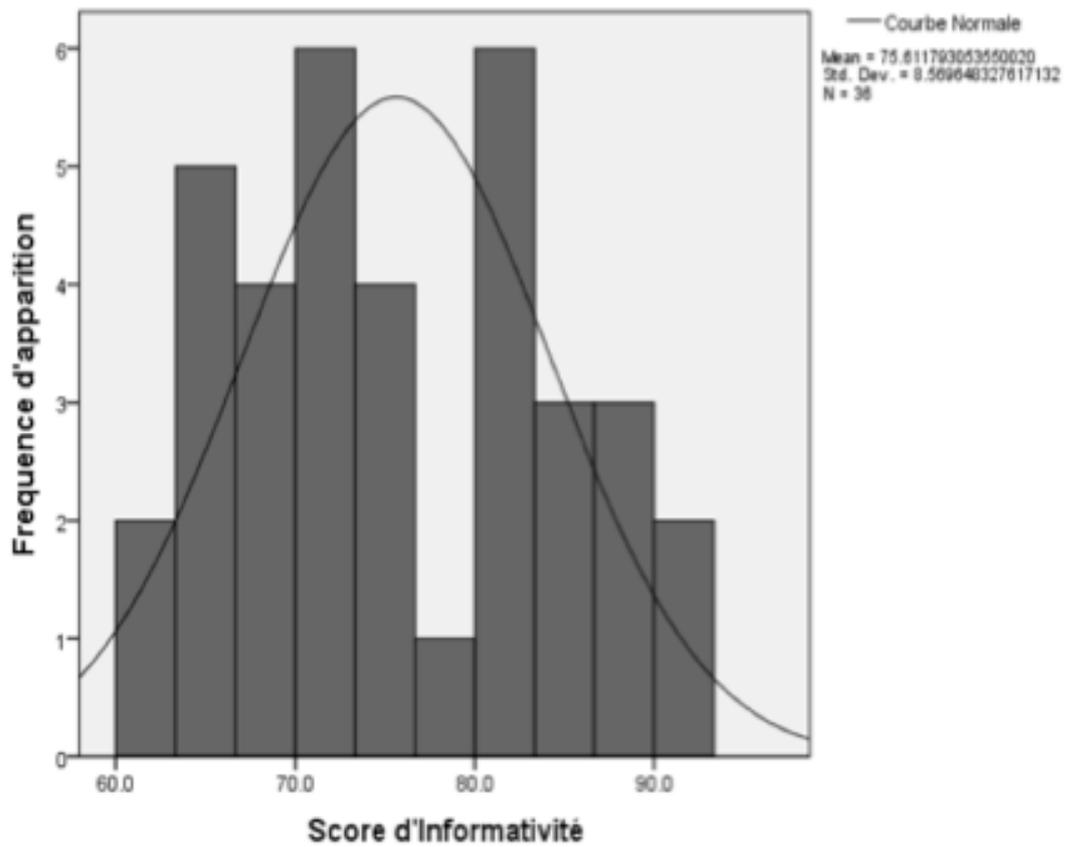
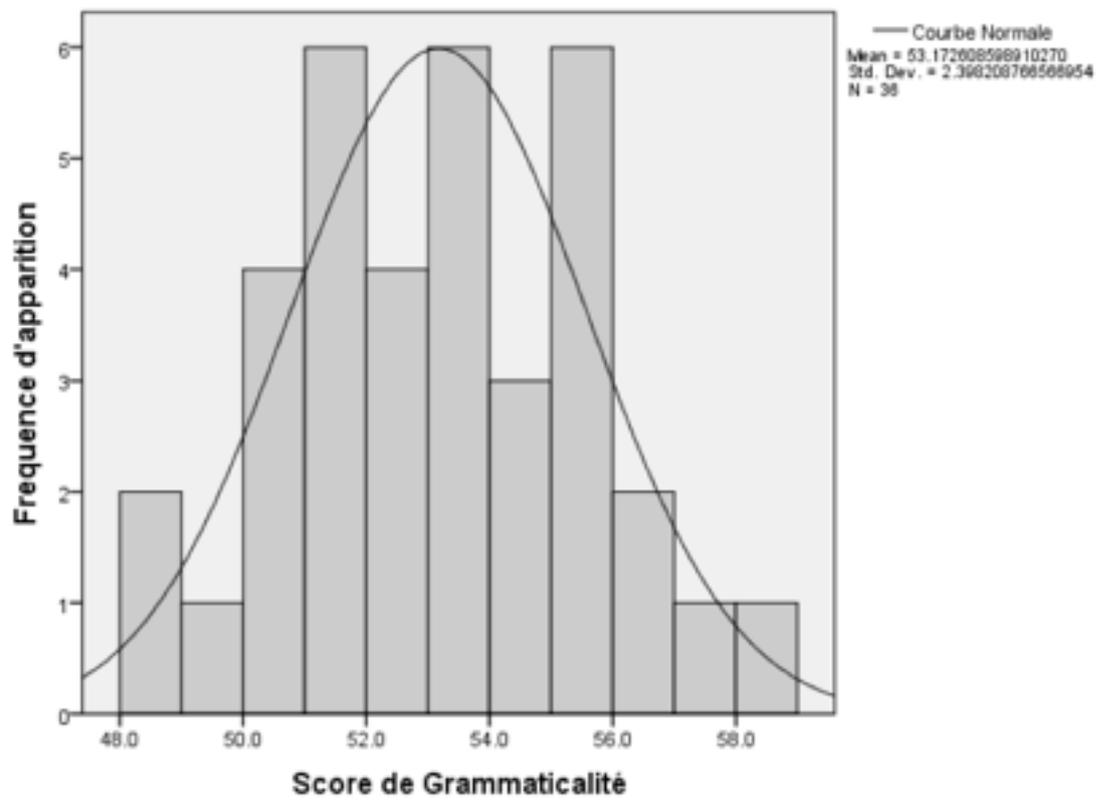


Figure F - Distribution des scores de Grammaticalité pour le groupe 3.



## DISCUSSION

Le MACS est un outil d'évaluation orthophonique en cours de conception. C'est un outil novateur dans son ambition d'objectiver l'évolution des capacités linguistiques de personnes aphasiques dans leur langage naturel. Tout test se doit d'avoir des normes établies sur les résultats d'un échantillon de population répondant aux critères d'exclusion définis. Les normes des témoins francophones québécois du MACS avaient été réalisées pour les tranches d'âge 18-30 ans et 51-70 ans. Notre travail consistait à compléter les normes des sujets sains du test. Nous avons pu établir les normes de la tranche d'âge manquante, 31-50 ans.

Notre expérience de recherche au Québec nous a sensibilisées à l'importance d'utiliser des normes adéquates à la population clinique testée. En effet, même si nous avons transcrit des échantillons de discours de personnes francophones, la différence d'accent et d'expressions entre le français québécois et le français de France métropolitaine a exigé quelques adaptations. Vivre au Québec pendant toute la durée de notre travail de recherche a été un avantage et nous a permis de faire notre oreille plus rapidement. Nous avons également pu demander l'aide de personnes québécoises pour la compréhension d'extraits particuliers. Il était important de bien connaître le vocabulaire québécois pour éviter de l'associer au sens premier véhiculé par notre culture française et ainsi risquer une mauvaise interprétation. Par exemple, le verbe « piger » qui signifie pour nous « comprendre » en langage familier, signifie en parler québécois « piocher ». Autre exemple, pour l'image de la plage, le mot « popsicle » revenait de manière récurrente, il s'agissait de « glace à l'eau ». Ainsi, créer des normes auprès de personnes issues d'autres cultures francophones est nécessaire pour l'utilisation du MACS hors du Québec. Malgré la francophonie, le vocabulaire pouvant parfois être différent, les normes établies sur une population francophone québécoise pourraient ne pas convenir à une population francophone non québécoise. Le matériel (i.e., les images) semble cependant adapté à d'autres cultures occidentales. Au cours des transcriptions, nous avons remarqué que les participants associaient les images à des lieux en Europe autant qu'en Amérique du Nord. Actuellement, le MACS est en cours d'adaptation dans d'autres langues, notamment l'anglais. Enfin, une quatrième tranche d'âge pourrait être créée pour les personnes de plus de soixante-dix ans puisqu'avec le vieillissement de la population, les personnes aphasiques sont de plus en plus susceptibles de faire partie de cette catégorie d'âge. Les

normes du MACS suivraient ainsi le type de patientèle aphasique rencontrée en clinique orthophonique.

Un autre point de discussion concerne le temps de cotation du MACS. Comme expliqué dans la partie pratique dans le paragraphe *Analyses*, après avoir recueilli les échantillons de discours produits lors de la description d'images, nous avons procédé à la transcription de 15 échantillons par participants. De plus, nous avons transcrit des échantillons de participants sans troubles du langage et la durée des échantillons était variable selon les participants mais pouvait parfois être longue. La transcription était une tâche nouvelle pour nous et a nécessité beaucoup de temps. Au fur et à mesure, notre vitesse de transcription s'est accélérée et nous avons acquis des mécanismes plus fonctionnels. Dans le cadre de la future passation clinique du MACS, le praticien n'aurait à transcrire que cinq échantillons bruts d'une personne aphasique par passation, ce qui réduira le temps de cotation en comparaison de celui qui nous a été nécessaire avec nos participants témoins. Cependant, on peut supposer que même pour un praticien de même langue et culture que le patient, qui est également habitué à la tâche de transcription, la cotation du MACS peut demeurer trop longue pour la pratique clinique. C'est pourquoi plusieurs pistes sont explorées pour faciliter l'utilisation de ce test par les cliniciens. Par exemple, l'utilisation d'une retranscription automatique par un logiciel de reconnaissance vocale pourrait faciliter la transcription. Toutefois, le calibrage du logiciel sur les voix de chaque patient et la complexité de reconnaissance vocale des productions de personnes aphasiques constituent des obstacles à surmonter. Un outil informatique pourrait aussi être programmé pour détecter les erreurs grammaticales pour le jugement des mots-outils mal utilisés.

Enfin, la réduction de l'échantillon à la première minute de discours pourrait être envisagée. Il faudrait préalablement étudier si seulement une minute de discours par stimulus serait assez représentative des capacités linguistiques du sujet. Il faudra donc évaluer la validité contre-critère avec la version des scores obtenus en prenant en compte les échantillons de discours complets (Ivanova et coll., 2013 [15]).

La particularité du MACS est qu'il propose trois versions parallèles afin de réduire l'effet d'apprentissage. Précisons que le biais d'apprentissage ne peut être entièrement supprimé puisque le sujet acquiert l'expérience de la forme du test dès sa première passation. Il garde donc une certaine connaissance des consignes et du protocole de passation qui pourrait influencer sur les deuxième et troisième passations. Toutefois, la fiabilité test-retest du MACS a été examinée et s'est avérée satisfaisante, ce qui renforce l'idée que le biais d'apprentissage a été réduit avec succès dans ce test (Frachon et Quiquempois, 2015 [13]).

La validation psychométrique d'un test est toujours à améliorer (Ivanova et Hallowell, 2013 [15]). Certaines étapes sont fondamentales et ont été réalisées pour le MACS. Toutefois, la validation de contenu devra être formellement établie par le consensus des experts en aphasiologie. De plus, tel que le souligne Fermanian (2005 [10]), la validité de construit n'est jamais aboutie car d'autres hypothèses peuvent être posées à tout moment. Celle-ci est renforcée ou affaiblie à mesure que le test est utilisé dans des recherches et qu'apparaît des convergences et des divergences avec d'autres mesures supposées plus ou moins proches du concept mesuré par le test (Anastasi, 1994 [1]).

# CONCLUSION

Notre mémoire de fin d'études avait pour but d'établir les normes des témoins francophones québécois de la tranche d'âge 31-50 ans du MACS (Montreal Assesment of Connected Speech), outil d'évaluation du discours naturel en cours d'élaboration.

Nous avons pu recueillir les données normatives de 36 participants et établir ainsi les normes des sujets sains entre 31 et 50 ans pour le MACS.

Le MACS dispose maintenant de normes de plus d'une centaine de témoins francophones québécois entre 18 et 70 ans, répartis en trois tranches d'âge.

Une nouvelle étape constitutive du MACS franchie, cet outil fait un pas de plus vers un avenir très prometteur. Il constitue un outil novateur dans l'évaluation des progrès langagiers et de l'efficacité de la thérapie.

# BIBLIOGRAPHIE

Anastasi, A. (1994). *Introduction à la psychométrie* (Montréal: Guérin universitaire. Traduit par François Gagné). [1]

Brady, M. C., Kelly, H., Godwin, J., & Enderby, P. (2012). Speech and language therapy for aphasia following stroke. *Cochrane Database Syst Rev*, 5(22), 420-425. [2]

Brookshire, R. H., & Nicholas, L. E. (1994). Speech sample size and test-retest stability of connected speech measures for adults with aphasia. *Journal of Speech, Language, and Hearing Research*, 37(2), 399-407. [3]

Brookshire, R. H., & Nicholas, L. E. (1994). Test-retest stability of measures of connected speech in aphasia. *Clinical aphasiology*, 22(1), 19-133. [4]

Chomel-Guillaume, S., Leloup, G., Bernard, I., Riva, I., & François-Guinaud, C. (2010). *Les aphasies: évaluation et rééducation*: Elsevier Masson Issy-les-Moulineaux. [5]

Darrigrand, B., & Mazaux, J.-M. (2005, janvier 1). Échelle de communication verbale de Bordeaux. Université Victor Segalen-Broché. [6]

Daviet, J.-C., Dudognon, P. J., Salle, J. Y., Munoz, M., Lissandre, J. P., Rebeyrotte, I., & Borie, M. J. (2002). *Rééducation des accidentés vasculaires cérébraux: bilan et prise en charge*. Editions Scientifiques et Médicales Elsevier SAS. [7]

Doyle, P. J., Goda, A. J., & Spencer, K. A. (1995). The communicative informativeness and efficiency of connected discourse by adults with aphasia under structured and conversational sampling conditions. *American Journal of Speech-Language Pathology*, 4(4), 130-134. [8]

Doyle, P. J., McNeil, M. R., Spencer, K. A., Goda, A. J., Cottrell, K., & Lustig, A. P. (1998). The effects of concurrent picture presentations on retelling of orally presented stories by adults with aphasia. *Aphasiology*, 12(7-8), 561-574. [9]

Fermanian, J. (2005). *Validation des échelles d'évaluation en médecine physique et de réadaptation: comment apprécier correctement leurs qualités psychométriques*. Paper presented at the Annales de réadaptation et de médecine physique. [10]

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. [11]

Fortin, F. (1994). Méthodologie. *Recherche en soins infirmiers*, 39. [12]

Frachon-Bizot, J., & Quiquempois, D. (2015). *Fiabilité test-retest et sensibilité au changement d'un nouvel outil de mesure du langage dans le discours naturel des personnes aphasiques*. Mémoire de l'Université Paris VI-Pierre et Marie Curie. [13]

France AVC. (2003). Page AVC infos.

Repéré à [http://www.franceavc.com/?rep=avc\\_info&rub=prevenir&comp=7](http://www.franceavc.com/?rep=avc_info&rub=prevenir&comp=7) [14]

Ivanova, M. V., & Hallowell, B. (2013). A tutorial on aphasia test development in any language: Key substantive and psychometric considerations. *Aphasiology*, 27(8), 891-920. [15]

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston naming test*. Pro-ed. [16]

Laveault, D., & Grégoire, J. (2002). *Introduction aux théories des tests: en psychologie et en sciences de l'éducation: De Boeck Supérieur*. [17]

Mazaux, J. M., & Orgogozo, J. M. (1982). Échelle d'évaluation de l'aphasie adaptée du Boston Diagnostic Aphasia Examination. Paris: Editions et applications psychologiques. [18]

Menn, L., Ramsberger, G., & Estabrooks, N. H. (1994). A linguistic communication measure for aphasic narratives. *Aphasiology*, 8(4), 343-359. [19]

Millette-Monier, J. (2012, mai). Normalisation préliminaire d'un outil d'évaluation du langage propositionnel. Données de recherche inédites. [20]

Nespoulous, J.-L., Lecours, A. R., Lafond, D., Lemay, A., Puel, M., Cot, F., . . . Giroux, F. (1992). Protocole Montréal-Toulouse d'examen linguistique de l'aphasie MT-86 M1bêta Module standard initial. (2e éd.). Ortho-Edition. [22]

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech, Language, and Hearing Research*, 36(2), 338-350. [23]

McNeil, M. R., & Prescott, T. E. (1978). *Revised token test*. Pro-ed. [24]

Rivard, J. (2014). Agrammatisme: effet potentiel d'un entraînement musical sur le traitement syntaxique. Mémoire de l'Université de Montréal. [25]

Rousseaux, M., Delacourt, A., Wyrzykowski, N., & Lefevre, M. (2001, juin 1). Test Lillois de Communication. OrthoÉdition. [26]

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591-611. [27]

Shewan, C. M. (1988). The Shewan Spontaneous Language Analysis (SSLA) system for aphasic adults: Description, reliability, and validity. *Journal of communication disorders*, 21(2), 103-138. [28]

Sparks, R., Helm, N., & Albert, M. (1974). Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex*, 10(4), 303-316. [29]

World Health Organization. (2016). Page maladies cardiovasculaires.  
Repéré à <http://www.who.int/mediacentre/factsheets/fs317/fr/> [29]

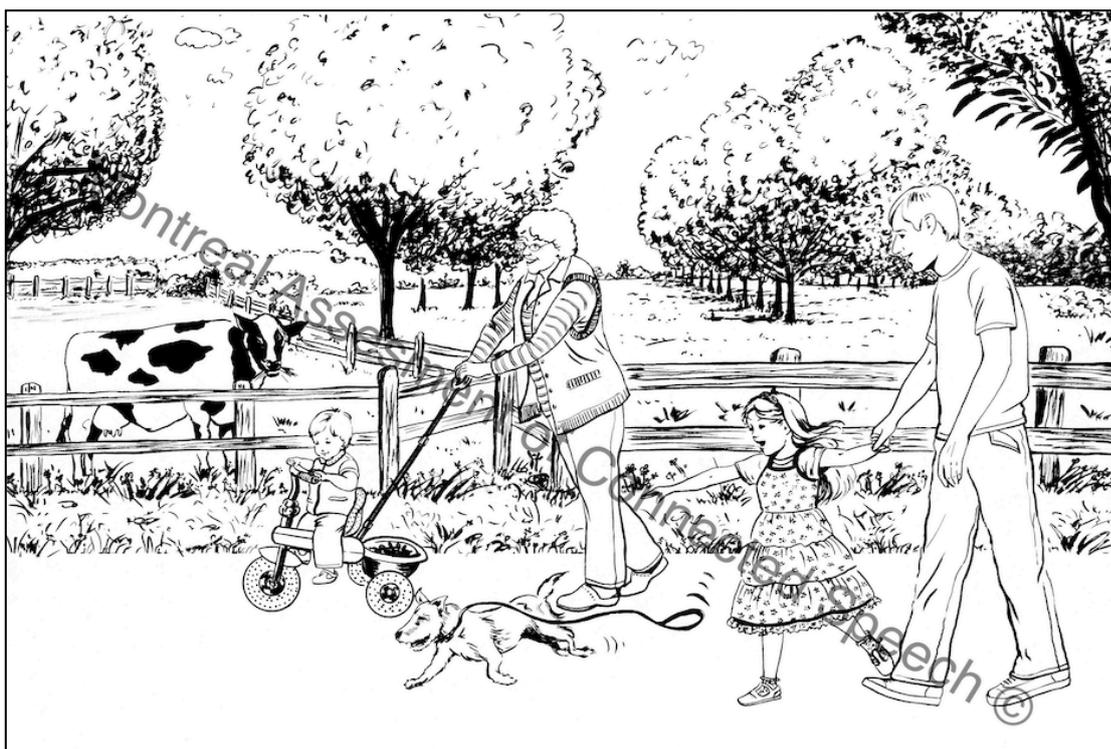
Yorkston, K. M., & Beukelman, D. R. (1980). An analysis of connected speech samples of aphasic and normal speakers. *Journal of speech and hearing disorders*, 45(1), 27-36. [30]

Zumbansen, A. (2015). Les bénéfices du chant dans la réadaptation de l'aphasie. Thèse de l'Université de Montréal. [31]

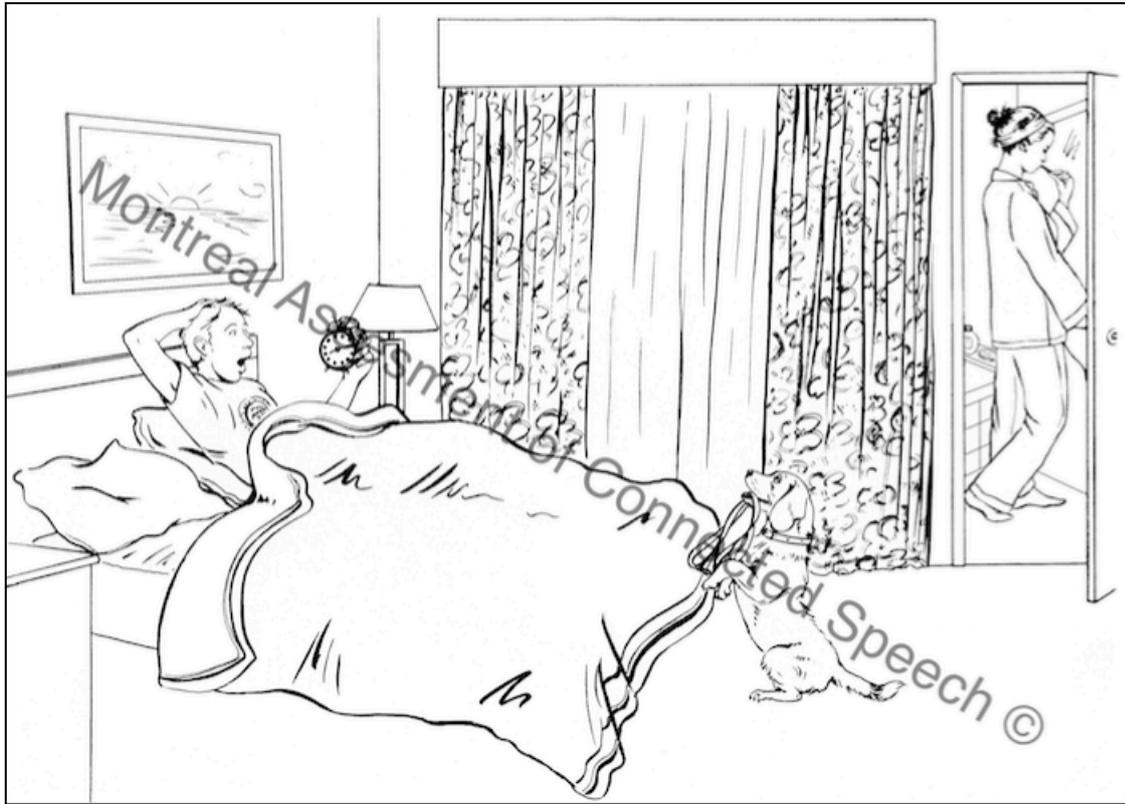
**Annexe A - Image 1 Banque**



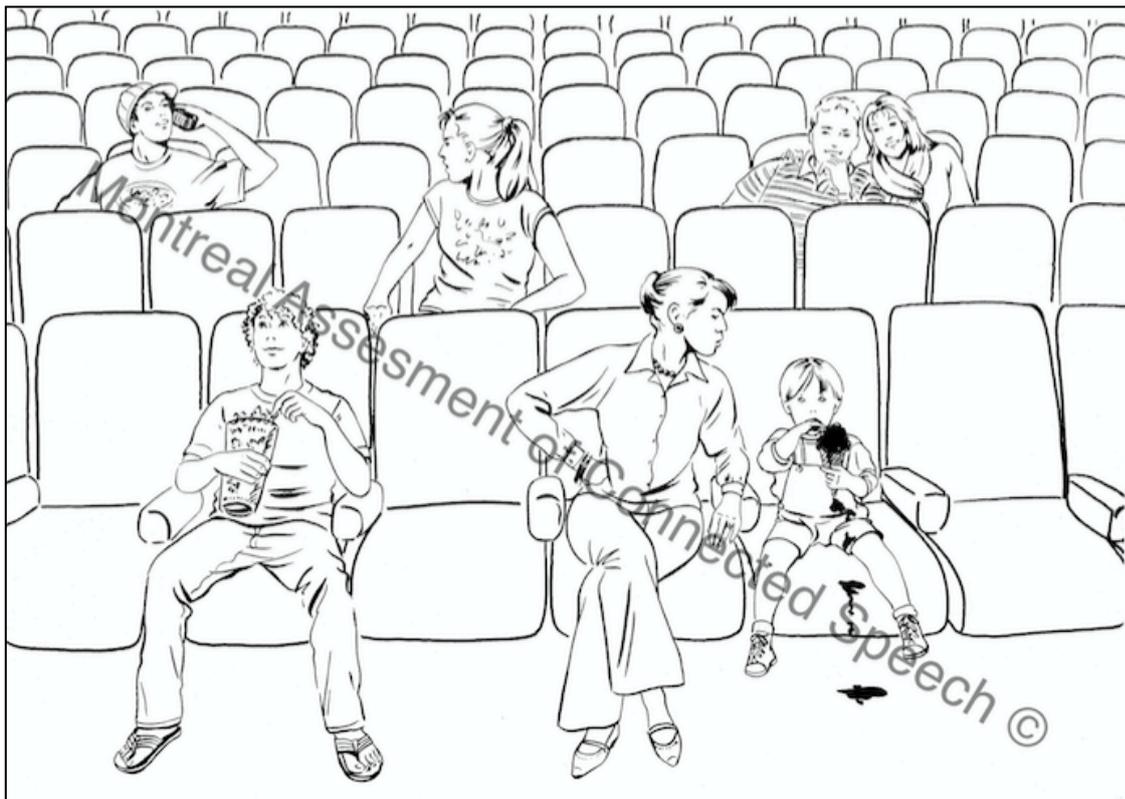
**Annexe B - Image 2 Campagne**



Annexe C - Image 3 Chambre



Annexe D - Image 4 Cinéma



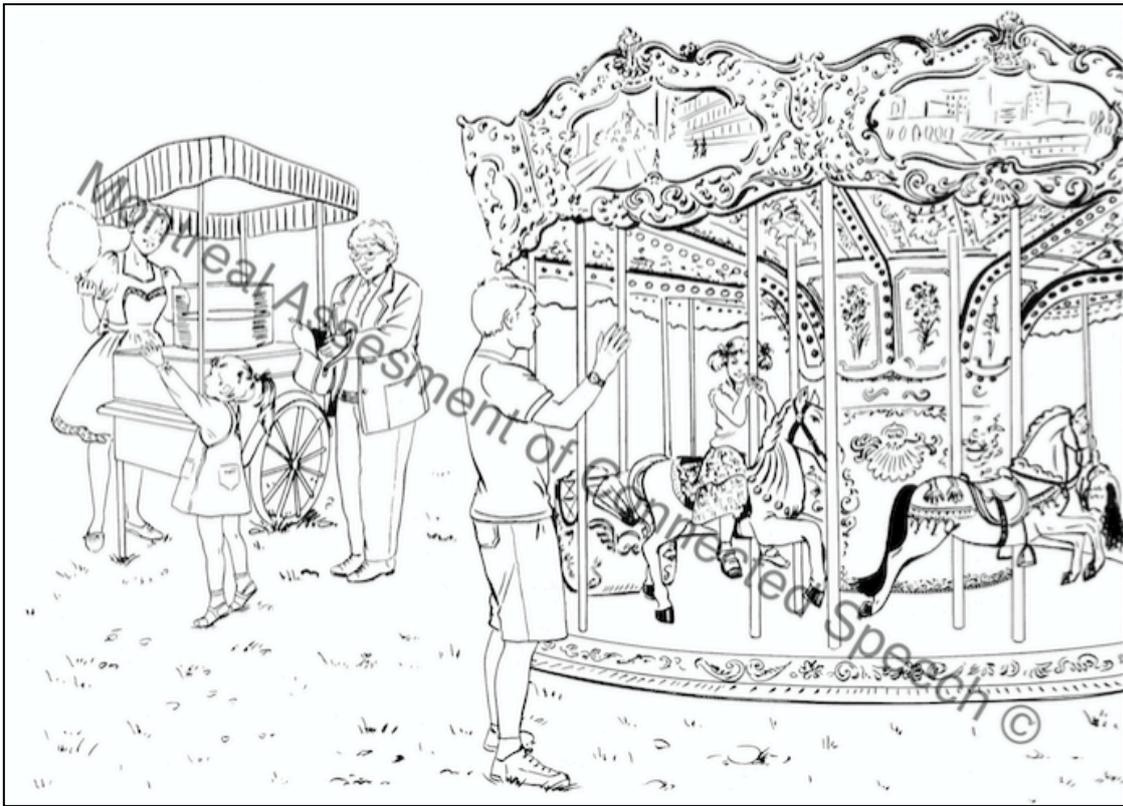
Annexe E - Image 5 Cuisine



Annexe F - Image 6 Épicerie



Annexe G - Image 7 Fête



Annexe H - Image 8 Montagne



Annexe I - Image 9 Parc



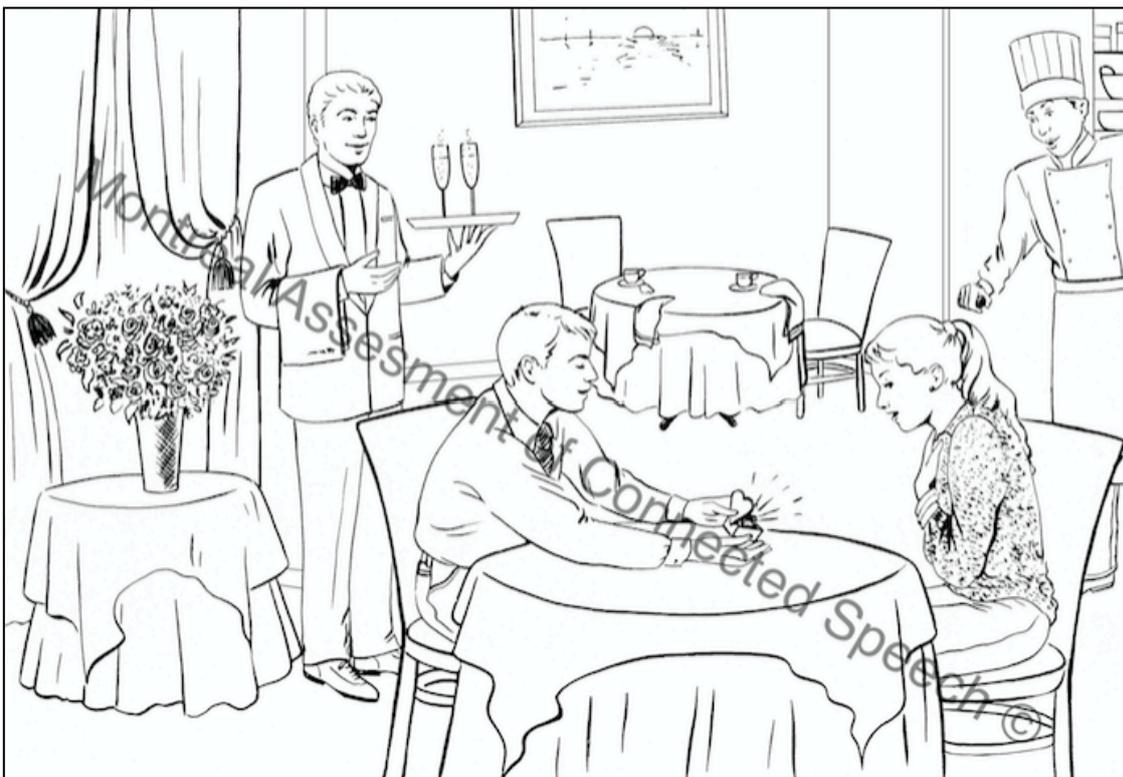
Annexe J - Image 10 Pêche



**Annexe K - Image 11 Plage**



**Annexe L - Image 12 Restaurant**



Annexe M - Image 13 Rue



Annexe N - Image 14 Salon



Annexe O - Image 15 Zoo



## **RÉSUMÉ**

Le MACS (Montreal Assessment of Connected Speech) est un test orthophonique en cours d'élaboration palliant le manque d'outils pour évaluer le discours naturel de patients aphasiques et quantifier leur progression au cours du traitement orthophonique. Nous présentons les normes pour la tranche d'âge 31-50 ans et finalisons ainsi l'analyse de données normatives de 101 participants francophones québécois.

## **ABSTRACT**

The Montreal Assessment of Connected Speech (MACS), an evaluation tool for speech and language pathologists, is currently being developed to deal with the lack of available tools for the assessment of natural connected speech of people with aphasia and the quantification of improvements during therapy. We present the reference norms of healthy participants aged 31 to 50, hence completing the data analysis of 101 Quebec French speaking controls.

## **TITRE**

« Établissement de normes francophones d'un outil d'évaluation quantitative du langage naturel. »

## **MOTS CLÉS**

Aphasie

Discours naturel

Informativité

Grammaticalité

Évaluation quantitative

**NOMBRE DE PAGES : 46**

**NOMBRE D'ANNEXES : 8**

**NOMBRE DE RÉFÉRENCES BIBLIOGRAPHIQUES : 31**