

DEPARTEMENT ORTHOPHONIE
FACULTE DE MEDECINE
Pôle Formation
59045 LILLE CEDEX
Tél : 03 20 62 76 18
departement-orthophonie@univ-lille.fr



 Université
de Lille

 ufr35
faculté
de médecine

MEMOIRE

En vue de l'obtention du
Certificat de Capacité d'Orthophoniste
présenté par

Alice CARDON

soutenu publiquement en juin 2022

**Validation transculturelle d'une batterie
d'évaluation des troubles neurologiques acquis
du langage et de la communication (i-MEL)
Analyse des épreuves de langage écrit**

MEMOIRE dirigé par

Yves MARTIN, Orthophoniste et Neuropsychologue, C.R.R.F. L'Espoir, Hellemmes-Lille

Lille – 2022

Remerciements

Je tiens à remercier mon directeur de mémoire, Monsieur Yves Martin, d'avoir accepté de m'encadrer. Merci pour votre confiance et votre écoute tout au long du projet.

Je remercie ma lectrice, Madame Lucile Thuet, pour le temps consacré aux relectures et les suggestions de rédaction. Merci également de m'avoir formée pendant cette dernière année à la clinique en service de rééducation neurologique cérébro-lésion à l'hôpital Swynghedauw.

Merci au Docteur Etienne Allart et à Madame Tran pour votre implication et vos précieux conseils au cours de ce projet.

Merci à l'ensemble de l'équipe de recherche québécoise pour la disponibilité dont vous avez fait preuve, et plus particulièrement à Mesdames Anaïs Deleuze et Perrine Ferré.

Je remercie également Gwennaëlle Delcroix, Salomé Jupille et Ninon Weiss pour cette collaboration qui a permis que le projet aboutisse.

Mes remerciements vont naturellement vers tou.te.s les participant.es qui ont accepté de faire partie de cette étude. Merci pour votre disponibilité, patience et investissement.

Je souhaite remercier plus personnellement ma famille qui m'a accompagnée avec ce désir de me voir accomplie aussi bien sur le plan professionnel que personnel.

Un grand merci à mes amies futures orthophonistes Anouk, Marie, Lucile, Isaure, Léa et Sandrine dont la présence a été essentielle ces dernières années.

J'adresse un grand merci à Arthur qui m'a soutenue et m'a tant apporté dès le début.

Je remercie infiniment Luna et Juliette avec qui je me construis depuis toujours et qui ont été présentes malgré la distance.

Merci à toutes les personnes avec qui j'ai partagé des souvenirs lillois et qui ont rendu ces cinq années si chaleureuses.

Résumé :

Les troubles neurologiques acquis représentent un enjeu de santé publique en raison des conséquences linguistiques et fonctionnelles qu'ils peuvent entraîner. La prise en soin des patients cérébro-lésés est toutefois dépendante d'une évaluation fiable et complète. C'est dans cette optique que l'i-MEL a été conçu au Québec en réponse aux besoins cliniques et aux dernières avancées scientifiques. Pour envisager une diffusion de cet outil en France, une validation transculturelle s'avère essentielle. Celle-ci permet de déterminer si d'éventuelles différences langagières et culturelles sont présentes entre les deux pays, auquel cas une adaptation de l'i-MEL sera nécessaire. Ce projet, poursuivant l'étude de 2020-2021, a été réalisé auprès de cent sujets témoins. Les passations et le traitement des données ont été répartis entre les quatre étudiantes travaillant sur le projet ; ce mémoire s'intéresse aux modalités écrites du langage. Une première analyse statistique (ANOVA) a permis d'étudier l'influence des effets socio-démographiques sur les deux populations puis nous avons comparé les performances des sujets français et québécois avec un test-t de student. Parmi les huit épreuves étudiées en langage écrit, les analyses suggèrent un effet des caractéristiques socio-démographiques (sexe, âge, niveau socio-culturel) en France sur cinq épreuves, contrairement au Québec où une seule épreuve est sensible à un tel effet. La comparaison entre les performances françaises et québécoises a quant à elle mis en avant une différence significative pour cinq épreuves en langage écrit. L'étude menée sur l'ensemble des 51 épreuves indique que l'outil est partiellement validé selon une méthodologie transculturelle.

Mots-clés :

Validation transculturelle – Aphasie – Evaluation – Langage – Communication

Abstract :

Acquired neurological disorders represent a public health issue because of the linguistic and functional consequences they can cause. However, the care of brain-injured patients is dependent on a reliable and complete assessment. It is in this perspective that the i-MEL was designed in Quebec in response to clinical needs and the latest scientific advances. In order to consider the diffusion of this tool in France, a cross-cultural validation is essential. This validation allows us to determine if there are any language and cultural differences between the two countries, in which case an adaptation of the i-MEL will be necessary. This project, pursuing the study of 2020-2021, was conducted with one hundred control subjects. The data collection and processing were divided among the four students working on the project ; this dissertation focuses on written language modalities. A first statistical analysis (ANOVA) allowed us to study the influence of socio-demographic effects on the two populations then we compared the performances of the French and Quebec subjects with a student t-test. Among the eight tests studied in written language, the analyses suggest an effect of socio-demographic characteristics (gender, age, socio-cultural level) in France on five tests, unlike in Quebec where only one test is sensitive to such an effect. The comparison between French and Quebec performances showed a significant difference for five written language tests. The study conducted on all 51 tests indicates that the tool is partially validated according to a cross-cultural methodology.

Keywords :

Cross-cultural validation – Aphasia – Assessment – Language – Communication

Table des matières

Introduction	1
Contexte théorique, buts et hypothèses	2
1. L'aphasiologie et la démarche d'évaluation.....	2
1.1 Définition de l'aphasie	2
1.2 Conséquences fonctionnelles	2
1.3 Enjeux d'une évaluation pertinente	2
2. L'outil diagnostique i-MEL	3
2.1 Présentation de la batterie	3
2.2 Intérêts de la batterie	3
2.3 Les compétences évaluées en langage écrit dans l'i-MEL et les troubles diagnostiqués.....	4
2.4 Validation transculturelle.....	7
3. Buts et hypothèses	7
Méthodologie	7
1. Définition de la population cible	7
1.1 Recrutement des sujets.....	7
1.2 Critères d'inclusion et d'exclusion	8
1.3 Répartition et description de la population étudiée.....	8
2. Matériel requis.....	10
2.1 Outil informatisé	10
2.2 Lieux de passation.....	10
3. Procédure adoptée	10
3.1 Entretien préliminaire	10
3.2 Passations.....	11
3.3 Traitement des données.....	11
3.4 Analyses statistiques	11
Résultats	12
1. Etude de l'influence des variables par type d'épreuve.....	12
1.1 Epreuve de compréhension écrite	13
1.2 Epreuves d'expression écrite.....	14
1.3 Epreuves sollicitant les transpositions.....	15

2. Comparaison des populations française et québécoise	18
Discussion	19
1. Rappel des objectifs de l'étude.....	19
2. Influence des caractéristiques socio-démographiques selon les contextes culturels.....	19
2.1 Effet de l'âge	19
2.2 Effet du sexe.....	21
2.3 Effet du niveau scolaire.....	21
2.4 Ouverture des analyses socio-démographiques à d'autres épreuves de l'i-MEL..	22
3. Analyse des différences franco-québécoises	23
3.1 Comparaison à l'échelle des épreuves de langage écrit	23
3.2 Comparaison élargie à d'autres épreuves de la batterie	25
4. Intérêts et limites	26
4.1 Intérêts et limites de la présente étude.....	26
4.2 Intérêts et limites de l'outil.....	27
5. Issue de l'étude et perspectives futures	28
Conclusion.....	29
Bibliographie.....	31
Liste des annexes.....	35

Introduction

Les troubles neurologiques acquis peuvent faire suite à différentes étiologies : vasculaire, traumatique, tumorale, épileptique, infectieuse ou neurodégénérative. La cause la plus fréquente reste l'accident vasculaire cérébral (AVC) dont l'incidence est estimée à 140 000 cas chaque année en France. Des troubles du langage sont retrouvés chez 20 à 55% des patients en phase aiguë de leur hospitalisation (Flamand-Roze et al., 2012). Ainsi, l'AVC représente un réel enjeu de santé publique (Mazaux, 2006). Ces troubles du langage sont la principale cause de handicap car ils détériorent la qualité de vie des patients. Ils influencent négativement les situations de communication, altèrent les relations interpersonnelles, provoquent une perte d'autonomie et restreignent les activités (Le Dorze & Brassard, 1995). Les troubles phasiques s'accompagnent également de changements émotionnels et psychosociaux puisque les difficultés de communication entraînent potentiellement un découragement social et un isolement (Lee et al., 2015). Pour améliorer la qualité de vie des patients aphasiques et de leur entourage, qui connaît également des bouleversements, l'élaboration d'un projet de soins adapté est nécessaire. Toutefois, celui-ci est dépendant d'une évaluation fiable et pertinente des capacités et des troubles que présente le patient.

Depuis les années 1970, une réelle évolution des connaissances a eu lieu dans le domaine des neurosciences. Le langage est désormais considéré comme dépendant des autres fonctions cognitives, c'est pourquoi il paraît nécessaire de concevoir différemment l'évaluation en aphasiologie (Chomel-Guillaume et al., 2010). Pourtant, les tests diagnostiques actuellement utilisés sont anciens et ne prennent donc pas en compte les références théoriques récentes (Monetta, 2016). Dans le but de répondre aux différents cliniques et scientifiques, une nouvelle batterie d'évaluation des troubles acquis du langage et de la communication (i-MEL) a été élaborée au Québec par l'équipe d'Yves Joanette et d'Ana Inés Ansaldo. Les auteurs visent de réelles qualités psychométriques argumentées et fondées à partir de la littérature scientifique (Deleuze et al., 2016).

Cette batterie d'évaluation pourrait être mise à disposition des orthophonistes cliniciens en France. Pour cela, une validation transculturelle est essentielle pour s'assurer que les normes obtenues au Québec sont applicables en France. En effet, la batterie pourra voir le jour en France si les données de la normalisation québécoise sont transposables à la population française en termes d'âge, de sexe et de niveau socio-culturel. Si ce n'est le cas, des adaptations seront à prévoir spécifiquement pour les épreuves présentant des résultats statistiques différents. Cette validation transculturelle a été effectuée grâce au recrutement de sujets sains volontaires par les quatre étudiantes travaillant sur le projet. Ensuite, chacune s'est intéressée à un axe spécifique de la batterie : épreuves préalables (profil communicationnel et processus cognitifs) ; épreuves sémantico-phonologiques ; épreuves syntaxico-discursives et épreuves de langage écrit. Ce mémoire s'axe sur l'analyse des modalités écrites du langage.

Contexte théorique, buts et hypothèses

Après avoir défini l'aphasie et les répercussions fonctionnelles engendrées sur la qualité de vie des patients, nous nous intéresserons à l'importance d'une évaluation pertinente des troubles du langage et de la communication. Ensuite, nous présenterons le Protocole informatisé Montréal d'Évaluation du Langage (i-MEL). Les épreuves spécifiques au langage écrit feront enfin l'objet d'une présentation détaillée.

1. L'aphasiologie et la démarche d'évaluation

1.1 Définition de l'aphasie

Mazaux (2008) définit l'aphasie comme étant « la conséquence clinique d'une désorganisation du langage humain par une lésion cérébrale circonscrite ». Il s'agit donc d'une atteinte du système nerveux central qui provoque un ensemble de troubles chez un individu qui maîtrisait auparavant le langage (De Partz & Pillon, 2014). Les profils sémiologiques des patients sont divers car les manifestations sur les différents domaines et versants du langage restent variées. Elles peuvent aussi bien concerner les versants expressif et réceptif que les modalités écrite et orale. Ces symptômes aphasiques entraînent par la suite des limitations d'activités de communication et donc une restriction de la participation sociale (Mazaux, 2006).

1.2 Conséquences fonctionnelles

L'AVC constitue une source de changements brusques de l'état de santé et des capacités fonctionnelles du patient aphasique, impliquant ensuite une dégradation de sa qualité de vie (Lee et al., 2015) : impact sur l'indépendance, les relations sociales et professionnelles. En effet, l'aphasie engendre une probabilité réduite de retour au travail comparativement aux patients non aphasiques victimes d'un AVC (Engelter et al., 2006). Les retentissements sur la vie du patient aphasique ne sont pas négligeables en raison de la perte des fonctions du langage : une détresse émotionnelle, un changement de l'image de soi, une dépression et un isolement social sont possibles (Spaccavento et al., 2013). L'importance d'une prise en soins des patients aphasiques n'est plus à prouver puisque, dans l'ensemble, la communication de ces patients prédit leur bien-être psychologique et leur santé sociale (Cruice et al., 2003).

1.3 Enjeux d'une évaluation pertinente

Une démarche diagnostique précoce des troubles de ces patients semble primordiale. Elle devra également être complète pour apprécier le fonctionnement et la participation sociale des personnes ayant subi un AVC (Lee et al., 2015). Cette évaluation permettra de débiter la prise en charge efficacement et ainsi améliorer les perspectives de récupération en limitant les facteurs de comorbidité (Flamand-Roze et al., 2012). Le suivi thérapeutique du patient lui offrira ensuite de meilleures capacités de communication traduisant une participation sociale et une qualité de vie plus élevées. En effet, la prise en charge vise une réadaptation qui permettra

une meilleure autonomie fonctionnelle dans différents environnements sociaux (Spaccavento et al., 2013).

2. L’outil diagnostique i-MEL

2.1 Présentation de la batterie

2.1.1 Equipe de recherche

La batterie a été réfléchié et conçue par les équipes d’Yves Joanette (Université de Montréal) et d’Ana Inés Ansaldo (Centre de Recherche de l’Institut Universitaire Gériatrique de Montréal). Ces équipes de chercheurs sont reconnues dans le domaine de l’aphasiologie pour leur participation à l’élaboration d’outils diagnostiques fréquemment utilisés par les cliniciens : MT 86 (Nespoulous et al., 1986, révisé en 1992) et MEC (Joanette et al., 2004). Leur collaboration a récemment abouti à la création de cette nouvelle batterie d’évaluation des troubles neurologiques du langage et de la communication, appelée protocole informatisé Montréal d’Évaluation du Langage (i-MEL).

2.1.2 Structure de l’i-MEL

Cet outil diagnostique est constitué de 8 modules (portrait communicationnel ; processus cognitifs ; composantes sémantiques ; phonologiques ; discursives, prosodiques ; compréhension syntaxique ; lecture/écriture) et comprend 51 tâches dans l’ensemble. De cette façon, la batterie propose un portrait communicationnel qui se veut fonctionnel pour identifier, entre autres, les partenaires de communication du sujet. Au-delà de l’évaluation purement linguistique, il paraît essentiel d’apprécier l’usage du langage dans des situations dites écologiques (Martin, 2018). Une évaluation des différents niveaux de traitement du langage est également prévue dans l’i-MEL. Enfin, la batterie offre une appréciation des fonctions cognitives, notamment capacités attentionnelles, mnésiques et exécutives. La présence d’éventuels troubles cognitifs est à considérer lors de la démarche diagnostique car ceux-ci contraignent les capacités de récupération des patients aphasiques (Trauchesse, 2018). Ainsi, l’i-MEL fournit un large choix d’épreuves avec celles dites élémentaires car classiquement admises et utilisées en aphasologie mais aussi des épreuves complémentaires ayant pour but d’affiner l’analyse diagnostique.

2.2 Intérêts de la batterie

2.2.1 Réponse à un besoin clinique

La construction d’un projet thérapeutique adapté au patient aphasique résulte d’une démarche diagnostique pertinente et cohérente. Mais force est de constater que les tests francophones actuels présentent des limites psychométriques. Il n’existe pas d’outil diagnostique des troubles neurologiques acquis à la fois valide, sensible, récent et complet en France (Monetta et al., 2016). L’i-MEL vient ainsi répondre aux besoins des cliniciens francophones en permettant l’évaluation de l’ensemble des compétences langagières et de communication. Cette pertinence diagnostique doit cependant être justifiée par les concepteurs du test, notamment en fournissant les données psychométriques probantes (Bauer et al., 2012).

2.2.2 Qualités psychométriques

L'i-MEL répond ainsi à plusieurs qualités psychométriques indispensables. En premier lieu, la batterie est construite à partir de modèles théoriques récents qui prennent en compte les approches cognitive, pragmatique et psycho-sociale. Les modules portrait communicationnel et processus cognitifs élémentaires de l'i-MEL en sont une bonne illustration. La conception de l'outil, en accord avec les cadres de référence actuels, atteste par conséquent d'une validité de construit élevée (Deleuze et al., 2016). En effet, la consultation de la littérature en aphasiologie et l'état des lieux des tests francophones actuels ont permis d'aboutir à l'élaboration de cette batterie. Les concepteurs de l'i-MEL ont également veillé à pertinence fonctionnelle puisque l'évaluation prend en compte, entre autres, le profil communicationnel du sujet et ses partenaires de communication. De ce fait, la batterie considère le fonctionnement des sujets en situation écologique. Cela permet d'apprécier les capacités de communication du sujet aphasique, ce qui est primordial car les troubles du langage ne sont pas systématiquement proportionnels aux perturbations de communication (Mazaux et al., 2006). Enfin, l'i-MEL offre une approche fine des troubles des patients grâce au contrôle des stimuli présentés. L'équipe québécoise mentionne l'analyse qualitative approfondie des items notamment par le contrôle des facteurs confondants. Parmi eux, le support visuel (couleur, photos et dessins réalistes), l'influence du hasard et la familiarité des concepts participent à la spécificité de la tâche proposée (Deleuze et al., 2016).

2.3 Les compétences évaluées en langage écrit dans l'i-MEL et les troubles diagnostiqués

L'évaluation de l'ensemble des compétences langagières écrites est préférable afin de comparer les performances du sujet en modalités d'entrée et de sortie (Reiff Cherney, 2004). Dans cette perspective, des épreuves de lecture et d'orthographe doivent être proposées.

2.3.1 Epreuve de compréhension écrite

Vérification lexicale écrite

Le sujet est invité à dire si le mot qu'il lit correspond ou non à ce qu'il voit. Apparaissent simultanément à l'écran le mot écrit et l'image (lorsqu'il s'agit d'un nom) ou la vidéo (quand ce sont des actions). Une même image est présentée trois fois : une fois avec le mot exact, une fois avec un mot sémantiquement proche (distracteur sémantique) et une fois avec un mot orthographiquement proche (distracteur phonologique). L'épreuve se compose de cette façon de soixante stimuli dont vingt sont concordants et quarante sont considérés comme des distracteurs.

Dans une perspective fonctionnelle du langage, d'autres tâches de l'i-MEL font intervenir la compréhension écrite, notamment la compréhension de paragraphes écrits et de texte long. Ces épreuves procurent des informations sur la capacité du sujet à accéder au sens. Ainsi, toutes les batteries en aphasiologie proposent une variété de tâches de difficulté croissante évaluant la compréhension en lecture (Reiff Cherney, 2004).

2.3.2 Epreuves d'expression écrite

Dénomination écrite

Une épreuve de dénomination écrite doit systématiquement être proposée pour déterminer si une agraphie est associée aux difficultés de lecture ou s'il s'agit d'une alexie pure (Reiff Cherney, 2004). À travers cette épreuve, il est question d'évaluer la production lexicale écrite de noms et de verbes. Pour cela, le sujet est invité à écrire le mot correspondant à l'image (pour les noms) ou à la vidéo (pour les actions). La tâche est composée de trente-six stimuli présentés successivement. Qualitativement, l'administrateur pourra observer des erreurs de lexèmes et des erreurs orthographiques de la part du sujet. L'orthographe grammaticale ne sera toutefois pas prise en considération.

Production écrite

Dans cette tâche de production écrite, il convient d'écrire quelques phrases sur un thème libre. Le sujet est incité à écrire un minimum de trente-huit mots.

Seule l'orthographe sera évaluée ici, les éléments de cohérence, cohésion, informativité feront l'objet d'une analyse approfondie dans un autre module de l'i-MEL. Les erreurs possiblement produites par le sujet seront phonologiques, lexicales ou morphosyntaxiques.

2.3.3 Épreuves sollicitant les transpositions

Lecture de lettres et de chiffres

Ce type d'épreuve est généralement proposé pour déterminer si une dissociation est présente entre la reconnaissance de lettres et la lecture de chiffres (Starrfelt & Behrmann, 2011). Cette tâche évalue ainsi les habiletés à décoder et à lire à voix haute des lettres et des chiffres présentés successivement. La tâche comprend vingt-deux stimuli dont dix lettres minuscules et douze chiffres. Les processus périphériques de la lecture sont alors objectivés puisque l'identification des lettres implique une analyse préalable au traitement lexical. Si l'épreuve est échouée, l'acuité visuelle du sujet sera à vérifier afin d'éliminer l'éventuel diagnostic différentiel de la dyslexie par négligence.

Lecture de mots et de pseudo-mots

Les troubles acquis de la lecture, dus à une lésion cérébrale, peuvent être un symptôme isolé ou faire partie d'un tableau clinique associant différents troubles (Starrfelt et al., 2013). La démarche diagnostique est alors importante pour cibler des objectifs thérapeutiques appropriés. La lecture de mots et de pseudo-mots participe ainsi au diagnostic d'une alexie pure ou d'une atteinte plus globale des compétences en langage écrit. Cette tâche vise à évaluer les capacités à décoder et à lire à voix haute des mots réguliers, irréguliers et des pseudo-mots présentés successivement. L'épreuve est composée de vingt-quatre items qui répondent à des critères de lexicalité, régularité, fréquence, complexité et longueur. Les variables psycholinguistiques des stimuli ont donc été contrôlées par les concepteurs de la batterie. Ici, ce sont les processus centraux de la lecture qui sont appréciés avec l'analyse des procédures lexicale et phonologique.

Plusieurs types d'erreurs peuvent être relevés lors de l'analyse qualitative. Les transformations telles que les inversions, omissions, substitutions ou ajouts de phonèmes font

partie des erreurs possiblement observées sur tous les types de mots. Les paralexies, c'est-à-dire la substitution d'un mot par un autre, peuvent être constatées lors de la lecture des mots réguliers et irréguliers. Ensuite, des régularisations de mots irréguliers, traduisant une voie lexicale atteinte, constituent un autre type d'erreur plausible. Il s'agit d'une utilisation compensatoire de la voie phonologique, qui est typique des profils de dyslexie de surface (De Partz, 2014). Enfin, des lexicalisations de pseudo-mots peuvent être rencontrées dans cette épreuve, autrement dit le sujet les lit comme des mots existants. Cela témoigne cette fois d'une atteinte de la voie phonologique (De Partz, 2018). Notons que plusieurs types de fautes sont possibles au sein d'un même mot.

Dictée de lettres et de chiffres

Dans cette épreuve, le sujet est invité à écrire les vingt-deux stimuli. Les chiffres devront être écrits de façon numérique et non en lettres. Cette dictée évalue ainsi la capacité de transcription du participant. Les résultats seront ensuite à mettre en parallèle avec ceux obtenus en lecture de chiffres et de lettres.

Dictée de mots et de pseudo-mots

Ici, l'épreuve d'écriture sous dictée de mots et pseudo-mots détermine la manière dont une information transmise à l'oral est produite à l'écrit. Il est demandé au sujet d'écrire vingt-quatre stimuli : des mots réguliers, irréguliers et des pseudo-mots. Les concepteurs de l'i-MEL ont décidé de prendre en compte la réforme de l'orthographe de 1990 pour la cotation.

De manière concomitante aux erreurs de lecture de mots et pseudo-mots, on peut observer des transformations, régularisations et lexicalisations sur le versant transcription. Des paraphrasies phonémiques (traduisant une altération de la voie phonologique) et des paraphrasies littérales (reflétant une atteinte du buffer graphémique) peuvent être produites par les patients aphasiques. Un éventuel effet de longueur permettra de préciser le trouble puisqu'il s'observe en cas d'altération du buffer graphémique mais non en cas d'atteinte de la voie phonologique (De Partz, 2018).

Dictée de phrases

La dictée de phrases évalue la manière dont les informations lexicales et morphosyntaxiques transmises à l'oral sont reproduites à l'écrit par le sujet. Cette épreuve de l'i-MEL comprend six phrases de longueur et de complexité croissantes.

Les erreurs que l'administrateur pourra relever de manière qualitative sont d'ordre phonologique, lexicale ou morphosyntaxique. Le sujet peut donc produire une erreur non phonologiquement plausible qui se traduit par une production déviante de l'item cible avec un non-respect de sa forme orthographique ou phonologique (Lambert, 2008). Il peut également commettre une erreur portant sur l'orthographe canonique du mot. Enfin, l'erreur peut porter sur l'accord des groupes verbaux et nominaux ou sur les homophones. Plusieurs types d'erreurs peuvent être comptabilisés au sein d'un même mot.

2.4 Validation transculturelle

En créant l’outil i-MEL, les auteurs visaient une utilisation dans les différents pays francophones. C’est la raison pour laquelle le choix des stimuli a été réfléchi en conséquence, en contrôlant par exemple les effets psycholinguistiques (Deleuze et al., 2016). Toutefois, pour que la batterie puisse profiter aux cliniciens en France, une validation transculturelle est nécessaire. Celle-ci permettra d’éviter des erreurs d’interprétation clinique. Le processus comprend généralement une phase de traduction, de vérification puis l’adaptation des scores au contexte culturel (Caron, 1999). Ici, il s’agira d’effectuer seulement la dernière étape puisque le contenu est déjà francophone. Les résultats obtenus permettront de statuer si les différences culturelles sont significatives et nécessitent alors une normalisation française spécifique ou non.

3. Buts et hypothèses

Ce mémoire s’inscrit dans la continuité d’un travail de recherche québécois et français 2020-2021. L’objectif est de répondre à un besoin à la fois scientifique et clinique. La conception d’un outil diagnostique innovant tel que l’i-MEL permettra une meilleure identification des troubles acquis du langage et de la communication des patients. Cette évaluation pertinente et en accord avec les données actuelles de la recherche offrira ensuite une approche thérapeutique adaptée. La pratique clinique sera ainsi meilleure.

Pour que l’i-MEL puisse être administré en France, cela nécessite de pouvoir généraliser les données de la population québécoise à la population française. Afin d’étudier de possibles divergences transculturelles, une validation sur des sujets sains en France est essentielle. Dans l’optique de cette validation transculturelle, l’objectif est double : le recrutement d’un nombre suffisant de sujets et l’analyse des données récoltées.

Ce projet étudie le transfert des résultats de la population québécoise à la population française et donc la question de la similarité entre les résultats. Si les résultats statistiques sont semblables entre les deux populations, ils pourront être combinés. Toutefois, s’ils ne le sont pas, une normalisation française et non franco-québécoise sera indispensable pour les épreuves dont les résultats statistiques ne sont pas identiques. Une adaptation sera donc à envisager pour une diffusion de la batterie en France.

Méthodologie

1. Définition de la population cible

1.1 Recrutement des sujets

Dans le cadre de la validation transculturelle de l’i-MEL, a été déterminé un certain nombre de sujets sains nécessaires à l’obtention d’une puissance statistique acceptable. La taille de l’échantillon doit être suffisamment importante pour garantir des résultats significatifs (Ember & Ember, 2011). Classiquement, on retient un nombre d’au moins 100 participants pour

déterminer des scores seuils et donc une performance dite dans la norme (Ivanova & Hallowell, 2013). Ainsi, le recrutement de la population témoin de cette étude s'est fait auprès de 80 adultes volontaires répartis entre les quatre étudiantes, auxquels s'ajoutent les 20 participants de l'étude de 2020-2021.

1.2 Critères d'inclusion et d'exclusion

Lors de la constitution de l'échantillon, les mêmes critères d'inclusion et d'exclusion préalablement définis par l'équipe de recherche québécoise ont été retenus (cf. Annexe n°1). Pour intégrer l'étude, les participants devaient donc remplir les caractéristiques suivantes : être francophone (en cas de bilinguisme, le français devait être la langue principale), être majeur (sans limite d'âge supérieure) et avoir donné son accord et son consentement après avoir été informé des modalités de l'étude.

Les critères d'exclusion retenus étaient les troubles visuels ou auditifs (en cas de trouble perceptif, celui-ci devait être corrigé), les troubles du langage acquis ou développemental, les troubles des apprentissages, les troubles cognitifs ou la présence d'une atteinte neurologique à la suite d'un accident vasculaire cérébral (AVC) ou traumatisme crânien. Ne pouvaient faire partie de l'étude les personnes sous tutelle ou curatelle et les sujets présentant des antécédents psychiatriques ou antécédents de toxicomanie et d'alcoolisme. Enfin, les orthophonistes ou étudiants en orthophonie ne pouvaient être inclus dans l'échantillon afin de ne pas entraver l'interprétation des résultats.

1.3 Répartition et description de la population étudiée

Pour une comparaison optimale des performances franco-québécoises, les sujets québécois et français ont été appariés, formant ainsi des paires de participants. Pour cela, les sujets français ont été classés selon les mêmes critères que les québécois, à savoir leur âge (18-29 ans, 30-59 ans, 60-79 ans et 80 ans et plus), leur sexe et enfin par leur niveau socio-éducatif (scolarité inférieure ou supérieure à 14 ans). Au-delà du respect de l'appariement, l'enjeu était également d'obtenir une répartition la plus homogène possible au sein de l'échantillon français.

Tableau 1 : Répartition de l'échantillon français

	FEMMES		HOMMES		TOTAL
	14 ans de scolarité et moins	Plus de 14 ans de scolarité	14 ans de scolarité et moins	Plus de 14 ans de scolarité	
18/29 ans	6	10	5	11	32
30/59 ans	6	13	5	9	33
60/79 ans	4	9	9	5	27
80 ans et +	5	1	2	0	8
TOTAL	54		46		100

La population française se compose donc de 100 participants classés selon trois variables différentes : le sexe, la classe d'âge et le niveau de scolarité (cf. Tableau 1). Concernant le genre, les femmes représentent 54 % de l'échantillon et les hommes 46 % (cf. Figure 1). La répartition en termes d'âge est davantage hétérogène et s'étend de 18 ans à 96 ans avec une moyenne se situant à 47,5 ans : 32 % de 18/29 ans, 33 % de 30/59 ans, 27 % de 60/79 ans et 8 % de 80 ans et plus (cf. Figure 2). Enfin, le niveau de scolarité « supérieur à 14 ans de scolarité » est légèrement sur-représenté (58%) par rapport aux personnes ayant un niveau inférieur ou égal à 14 ans de scolarité (42 %). La durée de scolarité des participants à l'étude française varie entre 8 et 23 ans de scolarité avec un temps de scolarité moyen établi à 14,8 ans (cf. Figure 3). A propos de la répartition géographique, le recrutement des participants sains a été réalisé dans l'entourage des étudiantes collaborant sur cette étude. Ainsi, la majorité des sujets français ont été recrutés dans les Hauts-de-France (42 %).

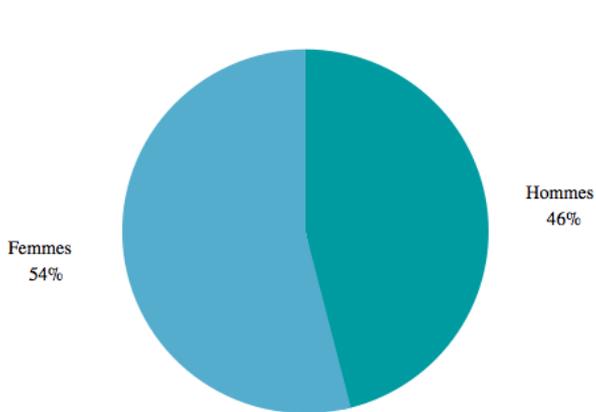


Figure 1 : Répartition du sexe dans l'échantillon français

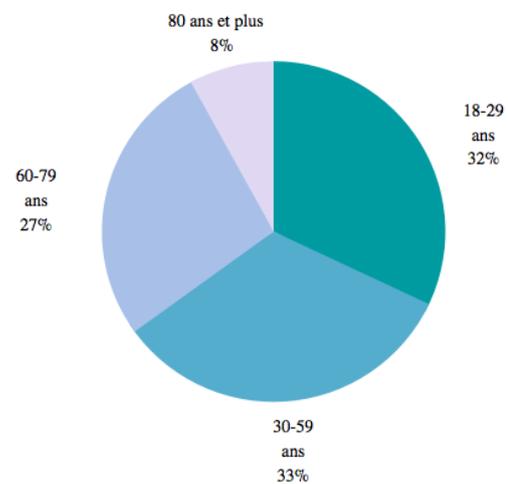


Figure 2 : Répartition des classes d'âge dans l'échantillon français

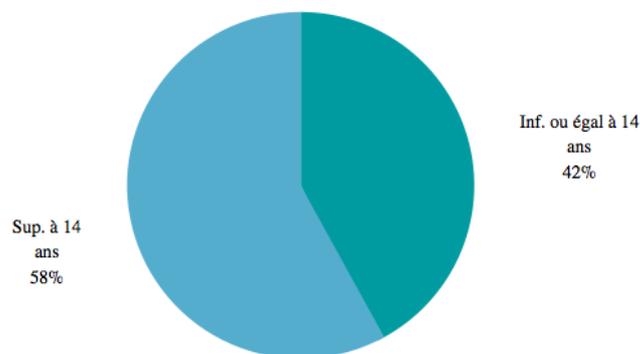


Figure 3 : Répartition du niveau de scolarité dans l'échantillon français

2. Matériel requis

2.1 Outil informatisé

La passation étant informatisée (Deleuze et al., 2016), une tablette exclusivement dédiée au projet de recherche a été fournie. Celle-ci servait tant pour la passation des tests de dépistage que pour les épreuves de l'i-MEL. L'évaluateur proposait donc les épreuves successivement et veillait à enregistrer les réponses des sujets selon le type d'épreuve (item désigné, enregistrement de la réponse orale ou photographie de la production écrite).

2.2 Lieux de passation

Les passations devaient se faire dans un endroit calme, sans source de distraction et suffisamment éclairé. Il pouvait s'agir des locaux du Département d'orthophonie de Lille (Faculté de Médecine Henri Warembourg), d'un lieu de travail ou encore d'un domicile. La passation ne pouvait se faire en distanciel puisque certaines épreuves nécessitent de valider les réponses du participant directement sur la tablette.

3. Procédure adoptée

3.1 Entretien préliminaire

Une première rencontre était prévue afin d'informer le sujet sur le projet de recherche et les modalités de participation. Il était indiqué au participant sa possibilité de refuser ou de se retirer de l'étude à tout moment. L'examineur répondait également aux éventuels questionnements du sujet et lui laissait un temps de réflexion si celui-ci en exprimait le besoin. Conformément aux règles éthiques et déontologiques, le sujet recruté communiquait son consentement après avoir reçu les informations à propos du projet. Enfin, l'examineur veillait à ce que le sujet réponde aux critères d'inclusion et d'exclusion préalablement définis.

Un dépistage visuel et auditif était administré pour tous les sujets. Les participants âgés de 50 ans ou plus recevaient un dépistage langagier et cognitif supplémentaire. Il s'agit du MoCA (Montreal Cognitive Assessment) et de la DTLA (Détection des Troubles du Langage chez l'Adulte et la personne âgée). Le choix du MoCA n'est pas anodin : ce test satisfait davantage les critères nécessaires pour détecter les troubles cognitifs légers que le MMSE (Mini-mental State Examination), tant en termes de sensibilité que de spécificité (Ciesielska et al., 2016). Quant à la DTLA, test de dépistage des troubles langagiers dans les pathologies neurodégénératives, il fait consensus en ce qui concerne sa validité, fiabilité test-retest et sa cohérence interne (Macoir et al., 2017). Au terme de ces tests de dépistage, si des troubles perceptifs et/ou cognitifs venaient à être suspectés, le sujet était retiré de l'étude conformément aux critères d'exclusion retenus. Il était par ailleurs invité à consulter son médecin traitant en vue d'exams complémentaires.

3.2 Passations

La passation des épreuves s'étalait sur une ou plusieurs sessions selon la fatigabilité des sujets. Il fallait prévoir approximativement cinq heures par participant. Si plusieurs sessions étaient nécessaires, une durée minimale de trois heures entre les deux était recommandée, et à l'inverse un maximum de quinze jours. Deux ordres de randomisation des épreuves, préalablement définis par l'équipe française, étaient proposés de manière aléatoire aux participants de l'étude. L'examineur veillait au maintien de l'investissement des sujets afin de limiter les arrêts de participation et était attentif à leur motivation. En outre, les résultats obtenus aux tests sont jugés valides si la motivation et la coopération des sujets sont convenables (Bauer et al., 2012).

3.3 Traitement des données

A l'issue des passations, les données des participants étaient anonymisées puis traitées. Chaque participant se voyait attribuer un numéro de dossier par souci de confidentialité. La cotation se faisait de manière automatisée pour certaines épreuves de la batterie mais d'autres nécessitaient une correction a posteriori. Enfin, plusieurs documents ont été créés afin de rassembler les données de passation de chaque étudiante : un tableau informant de l'avancement des passations avec toutes les caractéristiques des participants, un autre reprenant l'ensemble des binômes franco-qubécois et un espace de partage où toutes les sauvegardes (enregistrements audio et scans) étaient déposées.

3.4 Analyses statistiques

L'ensemble des données ont été extraites en fichiers CSV puis traitées à l'aide de la base de données ACCESS. L'exportation en fichier Excel a enfin permis une manipulation de ces données sur le logiciel statistique JAMOVI (Jamovi, 2021). Les quatre étudiantes travaillant sur le projet ont chacune analysé un domaine spécifique. Ce mémoire s'axe sur les modalités de langage écrit.

Les données de l'étude menée en 2020-2021 auprès de 20 participants avaient été analysées à partir de tests non paramétriques : test de Mann-Whitney pour l'influence du sexe, coefficient de corrélation de Spearman pour l'influence de l'âge et du niveau de scolarité et test de Wilcoxon pour la comparaison des résultats français et québécois. Ici, l'échantillon témoin ayant été élargi et comprenant désormais 100 sujets, des tests paramétriques ont pu être réalisés :

- pour étudier les effets sur les résultats obtenus, nous avons réalisé des ANOVA avec comme variables dépendantes chaque score aux épreuves, et comme facteurs le sexe (H ; F), la tranche d'âge (18-29 ; 30-59 ; 60-79 ; 80 et plus) et le niveau de scolarité (inférieur ou égal à 14 ans ; supérieur à 14 ans). Une transformation des données en rang était nécessaire avant de réaliser cette analyse. Les effets principaux ont donc pu être étudiés ainsi que les interactions doubles. En revanche, nous avons veillé à supprimer les interactions complexes. Enfin, les tests post-hoc ont utilisé une correction de Bonferroni pour prendre en compte le caractère répété

des mesures. Cette analyse a été effectuée à la fois sur la population française et québécoise après avoir récupéré les données brutes transmises par l'équipe de Montréal. Ainsi, nous avons pu comparer si l'influence de ces trois variables était identique dans les deux pays.

- nous avons ensuite étudié la comparaison entre performances des sujets français et québécois avec un test-t de student pour échantillons appariés.

Les données quantitatives continues ont été présentées par les moyennes et écarts-types et les données qualitatives non continues en effectifs et pourcentages. Le seuil de significativité a été fixé à $p < 0.05$ (bilatéral).

Résultats

Une étude de l'effet des variables socio-démographiques dans chacun des deux pays a été effectuée ainsi qu'une comparaison entre les performances françaises et québécoises. En effet, il s'agit d'analyser les variations à la fois individuelles mais aussi celles liées au contexte culturel (Marquer, 2014). Pour répondre à cette seconde question, les sujets français ont été appariés à des sujets québécois répondant aux mêmes critères d'âge, sexe et niveau scolaire. Ainsi, parmi les 189 personnes ayant participé à la normalisation québécoise, 100 ont été retenues et appariées aux sujets français pour observer d'éventuelles différences culturelles entre les performances de participants des deux pays.

1. Etude de l'influence des variables par type d'épreuve

Un contrôle des effets de variables propres aux sujets est essentiel puisque les performances obtenues aux épreuves de langage sont souvent influencées par ces variables (Lezak et al. ; Mitrushina et al., cités par Ivanova & Hallowell, 2013).

Pour une meilleure lisibilité des résultats, ces derniers ont été présentés sous forme de tableaux : pour chaque épreuve, on retrouve les scores et le temps quand celui-ci était comptabilisé. Une première partie relate les effets socio-démographiques retrouvés dans chaque pays avec les scores significatifs représentés en gras. La deuxième section rapporte les moyennes et écart-types pour les différents sous-groupes obtenus à la fois au Québec et en France. Enfin, les principales variables de la population française ont été présentées à l'aide de moyennes, écart-types et percentiles (le percentile 5 étant classiquement admis en clinique comme le seuil pathologique). Toutefois, il s'agira d'être prudents quant à l'analyse des résultats et de faire la distinction entre les scores et les temps obtenus. En effet, les temps sont inversés par rapport aux scores, ainsi un score élevé et un temps bas représentent tous deux une performance meilleure.

1.1 Epreuve de compréhension écrite

Tableau 2 : Epreuve de vérification lexicale écrite

		FR	QU		
		score /20	score /20		
Significativité des facteurs explicatifs (p-value)					
sexe		0.009	0.620		
tranche age		0.009	0.132		
classe_scolarité		0.319	0.684		
sexe * tranche age		0.428	0.262		
sexe * classe_scolarité		0.002	0.937		
tranche_age * classe_scolarité		0.540	0.083		
Moyennes (μ) et écart-types (σ) dans les différents sous-groupes					
		μ	σ	μ	σ
Sexe	F	19.3	0.851	18.7	1.231
	H	18.7	1.133	18.7	1.140
Tranche d'âge	18-29 ans	19.3	0.896	19.1	0.942
	30-59 ans	19.1	0.893	19.0	0.984
	60-79 ans	18.8	1.111	18.1	1.121
	80 ans et plus	17.9	0.991	17.6	1.685
Classe de scolarité	inf. ou égal à 14 ans	19.3	0.805	18.8	1.048
	sup. à 14 ans	18.8	1.120	18.6	1.280
Principales variables de la population générale française					
Moyenne		19.0			
Ecart-type		1.02			
Percentile	5	17.0			
	95	20.0			

A l'épreuve de vérification lexicale écrite, au sein de la population française, on observe un effet principal du sexe ($F(1,87) = 7.238$) avec des performances supérieures chez les femmes que chez les hommes ($p=0.009$). Un effet principal de l'âge est également retrouvé ($F(1,87) = 4.122$) où les performances de la classe d'âge la plus jeune sont supérieures à celles des deux classes d'âge plus élevées ($p=0.009$). On relève enfin une interaction entre le sexe et la scolarité ($F(1,87) = 10.358$) où les femmes ayant un niveau de scolarité supérieur à 14 ans obtiennent des performances meilleures comparativement aux hommes de scolarité inférieure ou égale à 14 ans. On observe aussi qu'au sein du groupe féminin, les performances de celles ayant un niveau supérieur à 14 ans sont plus élevées que celles possédant un niveau de scolarité inférieur à 14 ans. Enfin, à un niveau d'étude équivalent (14 ans et plus), les femmes obtiennent des scores supérieurs aux hommes ($p=0.002$). Ces effets n'ont pas été retrouvés dans l'échantillon québécois apparié.

1.2 Epreuves d'expression écrite

Tableau 3 : Epreuve de dénomination écrite

		FR				QU			
		score /36		temps (sec)		score /36		temps (sec)	
Significativité des facteurs explicatifs (p-value)									
sexe		0.773		0.392		0.358		0.341	
tranche age		0.321		0.613		0.155		0.216	
classe_scolarité		0.403		0.840		0.888		0.509	
sexe * tranche age		0.027		0.262		0.806		0.204	
sexe * classe_scolarité		0.214		0.642		0.248		0.966	
tranche_age * classe_scolarité		0.658		0.329		0.753		0.354	
Moyennes (μ) et écart-types (σ) dans les différents sous-groupes									
Sexe	F	μ	σ	μ	σ	μ	σ	μ	σ
	H	34.3	1.53	112.8	36.97	34.7	1.81	175.7	54.65
Tranche d'âge	18-29 ans	33.8	1.78	125.5	45.31	35.0	1.47	169.7	54.95
	30-59 ans	34.3	1.51	120.8	43.05	35.3	1.00	178.7	54.22
	60-79 ans	34.3	1.63	119.5	45.63	35.2	1.36	159.1	52.70
	80 ans et plus	33.9	1.69	113.0	39.77	34.4	1.67	181.3	60.24
Classe de scolarité	inf. ou égal à 14 ans	33.0	2.07	125.4	17.78	32.8	2.82	178.9	39.91
	sup. à 14 ans	34.0	1.76	137.2	51.65	34.9	1.54	170.7	52.75
Principales variables de la population générale française									
Moyenne		34.1		119					
Ecart-type		1.66		41.3					
Percentile	5	31.0		77.8					
	95	36.0		189					

Concernant les scores obtenus en dénomination écrite dans l'échantillon français, on remarque une interaction entre le sexe et la classe d'âge ($F(1,87) = 3.1980$; $p=0.027$). Cette observation n'a pas été retrouvée au sein de la population québécoise appariée.

Tableau 4 : Epreuve de production écrite

		FR		QU	
		score /38		score /38	
Significativité des facteurs explicatifs (p-value)					
sexe		0.325		0.875	
tranche age		0.659		0.604	
classe_scolarité		0.302		0.239	
sexe * tranche age		0.926		0.899	
sexe * classe_scolarité		0.427		0.634	
tranche_age * classe_scolarité		0.895		0.049	
Moyennes (μ) et écart-types (σ) dans les différents sous-groupes					
Sexe	F	μ	σ	μ	σ
	H	37.5	0.840	37.4	0.857
Tranche d'âge	18-29 ans	37.2	1.403	36.3	1.980
	30-59 ans	37.4	0.948	37.0	1.867
	60-79 ans	37.4	1.410	37.1	1.071
	80 ans et plus	37.2	1.145	36.7	1.408
Classe de scolarité	inf. ou égal à 14 ans	37.8	0.463	36.3	2.375
	sup. à 14 ans	37.4	1.21	37.0	1.17
Principales variables de la population générale française					
Moyenne		37.4			
Ecart-type		1.14			
Percentile	5	34.0			
	95	38.0			

A propos de l'épreuve de production écrite, aucun effet n'a été retrouvé dans les analyses statistiques dans le groupe français. En revanche, au sein des sujets québécois, il existe une interaction entre l'âge et le niveau de scolarité ($F(1,86) = 2.7215; p=0.049$).

1.2 Epreuves sollicitant les transpositions

Tableau 5 : Epreuve de lecture de lettres et de chiffres

		FR		QU					
		score /22	temps (sec)	score /22	temps (sec)				
Significativité des facteurs explicatifs (p-value)									
sexe		0.036	0.440	0.717	0.488				
tranche age		0.838	0.122	0.769	0.699				
classe scolarite		0.314	0.698	0.670	0.104				
sexe * tranche age		0.192	0.075	0.116	0.705				
sexe * classe scolarite		0.707	0.283	0.852	0.847				
tranche_age * classe_scolarite		0.215	0.234	0.537	0.207				
Moyennes (μ) et écart-types (σ) dans les différents sous-groupes									
		μ	σ	μ	σ	μ	σ	μ	σ
Sexe	F	21.9	0.351	38.5	8.827	22.0	0.191	35.4	7.410
	H	22.0	0.147	38.5	7.343	21.9	0.285	34.6	6.763
Tranche d'âge	18-29 ans	21.9	0.246	37.2	5.615	22.0	0.000	34.8	8.722
	30-59 ans	22.0	0.000	39.3	6.683	22.0	0.174	34.3	6.526
	60-79 ans	22.0	0.192	39.0	11.295	21.9	0.362	36.5	6.440
	80 ans et plus	21.6	0.744	38.7	10.348	21.9	0.354	34.2	3.830
Classe de scolarité	inf. ou égal à 14 ans	22.0	0.000	42.1	9.559	21.9	0.261	35.0	6.308
	sup. à 14 ans	21.9	0.360	35.9	5.747	21.9	0.223	35.1	7.670
Principales variables de la population générale française									
Moyenne		21.9		38.5					
Ecart-type		0.278		8.14					
Percentile	5	22.0		30.7					
	95	22.0		54.0					

Au sujet du score de la lecture de lettres et de chiffres, un effet principal du sexe a été relevé ($F(1,87) = 4.518$) parmi les participants français. Les hommes obtiennent des performances supérieures aux femmes ($p=0.036$). Cet effet n'est en outre pas retrouvé dans l'échantillon québécois apparié.

Tableau 6 : Epreuve de lecture de mots et de pseudo-mots

		FR		QU	
		score /24	temps (sec)	score /24	temps (sec)
Significativité des facteurs explicatifs (p-value)					
sexe		0.616	0.178	0.574	0.984
tranche age		0.311	0.824	0.767	0.800
classe scolarite		0.669	0.359	0.378	0.983
sexe * tranche age		0.622	0.194	0.710	0.428
sexe * classe scolarite		0.372	0.782	0.252	0.583
tranche_age * classe_scolarité		0.259	0.251	0.268	0.557
Moyennes (μ) et écart-types (σ) dans les différents sous-groupes					
		μ	σ	μ	σ
Sexe	F	23.7	0.521	37.3	6.705
	H	23.7	0.553	39.4	8.495
Tranche d'âge	18-29 ans	23.7	0.602	39.3	7.406
	30-59 ans	23.8	0.442	35.8	6.664
	60-79 ans	23.7	0.465	39.5	8.922
	80 ans et plus	23.5	0.756	40.1	5.912
Classe de scolarité	inf. ou égal à 14 ans	23.8	0.490	40.6	9.437
	sup. à 14 ans	23.6	0.552	36.6	5.463
Principales variables de la population générale française					
Moyenne		23.7		38.2	
Ecart-type		0.533		7.62	
Percentile	5	23.0		27.5	
	95	24.0		51.3	

Aucun effet n'a été observé à la lecture de mots et de pseudo-mots, tant dans la population française que québécoise.

Tableau 7 : Epreuve de dictée de lettres et de chiffres

		FR	QU
		score /22	score /22
Significativité des facteurs explicatifs (p-value)			
sexe		0.308	0.940
tranche age		0.229	0.666
classe scolarite		0.550	0.895
sexe * tranche age		0.113	0.827
sexe * classe scolarite		0.092	0.142
tranche_age * classe_scolarité		0.080	0.769
Moyennes (μ) et écart-types (σ) dans les différents sous-groupes			
		μ	σ
Sexe	F	22.0	0.136
	H	22.0	0.000
Tranche d'âge	18-29 ans	22.0	0.000
	30-59 ans	22.0	0.000
	60-79 ans	22.0	0.192
	80 ans et plus	22.0	0.000
Classe de scolarité	inf. ou égal à 14 ans	22.0	0.000
	sup. à 14 ans	22.0	0.131
Principales variables de la population générale française			
Moyenne		22.0	
Ecart-type		0.100	
Percentile	5	22.0	
	95	22.0	

Concernant la dictée de lettres et de chiffres, on ne retrouve pas d'effet non plus, ni au sein des participants français, ni chez les québécois.

Tableau 8 : Epreuve de dictée de mots et pseudo-mots

		FR		QU	
		score /24	temps (sec)	score /24	temps (sec)
Significativité des facteurs explicatifs (p-value)					
sexe		0.358	0.683	0.969	0.825
tranche age		0.042	0.088	0.823	0.213
classe_scolarité		0.814	0.111	0.285	0.401
sexe * tranche age		0.782	0.800	0.459	0.907
sexe * classe_scolarité		0.962	0.511	0.330	0.287
tranche_age * classe_scolarité		0.644	0.771	0.676	0.720
Moyennes (μ) et écart-types (σ) dans les différents sous-groupes					
		μ	σ	μ	σ
Sexe	F	22.4	1.38	128.8	31.28
	H	22.5	1.22	129.4	37.37
Tranche d'âge	18-29 ans	22.8	0.998	126.9	28.742
	30-59 ans	22.6	1.223	116.7	31.699
	60-79 ans	22.2	1.075	130.2	29.616
	80 ans et plus	21.0	2.268	177.8	32.220
Classe de scolarité	inf. ou égal à 14 ans	22.3	1.22	122.7	32.37
	sup. à 14 ans	22.5	1.37	132.5	34.52
Principales variables de la population générale française					
Moyenne		22.4	129		
Ecart-type		1.30	33.9		
Percentile	5	20.0	96.4		
	95	24.0	201		

Quant aux scores obtenus à la dictée de mots et de pseudo-mots, on retrouve un effet principal de l'âge ($F(1,87) = 2.8573$; $p=0.042$) au sein des participants français. Cet effet n'est toutefois pas observé au Québec.

Tableau 9 : Epreuve de dictée de phrases

		FR	QU
		score /38	score /38
Significativité des facteurs explicatifs (p-value)			
sexe		0.541	0.237
tranche age		0.660	0.706
classe_scolarité		0.023	0.311
sexe * tranche age		0.141	0.809
sexe * classe_scolarité		0.587	0.966
tranche_age * classe_scolarité		0.388	0.608
Moyennes (μ) et écart-types (σ) dans les différents sous-groupes			
		μ	σ
Sexe	F	37.6	0.570
	H	37.3	0.861
Tranche d'âge	18-29 ans	37.4	0.615
	30-59 ans	37.4	0.867
	60-79 ans	37.5	0.700
	80 ans et plus	37.5	0.756
Classe de scolarité	inf. ou égal à 14 ans	37.3	0.835
	sup. à 14 ans	37.6	0.626
Principales variables de la population générale française			
Moyenne		37.4	
Ecart-type		0.729	
Percentile	5	36.0	
	95	38.0	

Enfin, pour les scores de la dictée de phrases, on note un effet principal de scolarité en France ($F(1,87) = 5.368$). En effet, les sujets ayant un niveau scolarité supérieur à 14 ans ont des meilleures performances que ceux qui ont un niveau scolarité inférieur ou égal à 14 ans ($p=0.023$). Cet effet n'a au contraire pas été retrouvé chez les participants québécois.

2. Comparaison des populations française et québécoise

Tableau 10 : Comparaison des populations française et québécoise

Epreuves de langage écrit			Moyenne et écart-type		P	% de différence
			FR	QU		
Compréhension écrite	Vérification lexicale écrite (/20)		19 (1,025)	18,70 (1,185)	0,056	
Expression écrite	Dénomination écrite	Score (/36)	34,09 (1,664)	<u>34,83 (1,658)</u>	0,002	2,12 %
		Temps	<u>118,63 (41,295)</u>	172,95 (54,596)	< 0,001	31,41 %
	Production écriture (/38)		<u>37,36 (1,147)</u>	36,89 (1,564)	0,019	1,26 %
Transpositions	Lecture de lettres et de chiffres	Score (/22)	21,94 (0,278)	21,94 (0,239)	1,000	
		Temps	38,49 (8,138)	<u>35,07 (7,096)</u>	0,001	8,89 %
	Lecture de mots et pseudo-mots	Score (/24)	<u>23,72 (0,533)</u>	23,34 (0,913)	< 0,001	1,60 %
		Temps	<u>38,25 (7,619)</u>	41,59 (7,992)	0,003	8,03 %
	Dictée de mots et pseudo-mots	Score (/24)	22,43 (1,305)	22,45 (1,445)	0,917	
		Temps	<u>129,05 (33,921)</u>	150,78 (35,016)	< 0,001	14,41 %
	Dictée de lettres et de chiffres (/22)		21,99 (0,100)	21,92 (0,339)	0,052	
	Dictée de phrases (/38)		37,44 (0,729)	37,28 (0,944)	0,178	

Note : différences significatives indiquées **en gras** avec une différence inférieure à 5 % **en vert**, différence comprise entre 5 % et 10 % **en orange** et différence au-delà de 10 % **en rouge**. Lorsqu'une différence a été retrouvée, la meilleure performance entre les deux groupes (score supérieur ou temps inférieur) a été soulignée.

Parmi les huit épreuves étudiées dans ce mémoire, cinq présentent une différence significative entre les sujets français et québécois ($p < 0,05$). Les sujets français obtiennent de meilleurs résultats en dénomination écrite (temps), en production écrite (score), en lecture de mots et pseudo-mots (score et temps) et en dictée de mots et pseudo-mots (temps). En revanche, les participants québécois obtiennent des performances supérieures en dénomination écrite (score) et sont plus rapides en lecture de lettres et de chiffres. Une différence supérieure à 5 % a été retrouvée entre les deux populations pour la vitesse de lecture de lettres et de chiffres, de mots et pseudo-mots, en dénomination écrite et en dictée de mots et pseudo-mots. Ainsi, les écarts importants entre les deux groupes se situent sur la composante temporelle. Les résultats détaillés obtenus au test-t de student pour les épreuves de langage écrit sont présentés en annexes (cf. Annexe n°2).

Discussion

1. Rappel des objectifs de l'étude

Ce travail s'intéresse donc à la validation transculturelle de la version québécoise de l'i-MEL dans le contexte français. L'objectif principal est de déterminer si la diffusion de l'outil est possible à un contexte culturel différent et s'il peut profiter à d'autres cliniciens francophones. Pour cela, la seule traduction littérale des items ne s'avère pas suffisante, il convient aussi de constater les différences langagières et culturelles entre les deux pays, et le cas échéant de statuer si elles remettent en cause l'utilisation de cet outil en France. Bien que les deux pays soient francophones et partagent une culture occidentale, Joannette et al. (2014) précisent que « le langage et la communication prennent sens dans un contexte donné et les variables culturelles ne peuvent être ignorées ». En effet, la langue commune ne garantit pas la sensibilité, validité ou fidélité de l'i-MEL dans le contexte français. Ainsi, une validation transculturelle permet de mesurer les mêmes concepts dans des contextes culturels différents, en intégrant les populations ciblées dans les schémas d'évaluation (Marquer, 2014).

Dans cette optique, l'i-MEL a été administré à 100 sujets français pour obtenir des résultats en adéquation avec le nouveau contexte culturel. A ce titre, les données présentées ici sont issues d'une étude incluant deux pays, impliquant une large équipe et un travail sur plusieurs années. En outre, ce processus permet de ne pas reproduire un même travail scientifique déjà mené mais d'en tirer parti pour répondre à des besoins cliniques actuels.

2. Influence des caractéristiques socio-démographiques selon les contextes culturels

Gadet (2021) indique que « les linguistes s'accordent aisément sur le fait qu'une propriété majeure des langues est la variabilité intralinguistique des productions, qui complète la diversité interlinguistique de langue à langue ». Ainsi, au-delà des aspects culturels pouvant varier d'une langue à l'autre, la prise en compte des paramètres caractérisant les sujets de façon macro-sociologique, tels que l'âge, le sexe et le niveau d'étude, paraît essentielle. En effet, ces caractéristiques peuvent influencer la performance des sujets. Il s'agit également d'observer si ces effets sont semblables entre les sujets français et québécois. C'est dans cette perspective que nous avons étudié les variables (sexe, âge, niveau de scolarité) et leurs effets significatifs retrouvés dans les épreuves impliquant le langage écrit. Nous les comparerons avec les données présentées dans la littérature puis nous élargirons aux autres épreuves de la batterie partageant des similitudes avec celles traitées dans ce mémoire.

2.1 Effet de l'âge

Concernant l'âge, quatre tranches ont été définies : 18-29 ans, 30-59 ans, 60-79 ans et 80 ans et plus. On aurait tendance à penser que l'âge a une influence sur les résultats obtenus : les capacités de lecture et d'orthographe se détérioreraient chez les sujets les plus âgés (80 ans

et plus) en raison du vieillissement cognitif normal. Un effet principal de l'âge a été retrouvé sur les scores de seulement deux épreuves impliquant la modalité écrite en France : la vérification lexicale écrite et la dictée de mots et de pseudo-mots.

A l'épreuve de vérification lexicale écrite, les sujets français les plus jeunes (18-29 ans) obtiennent des performances supérieures à celles des sujets appartenant aux classes d'âge les plus âgées (60-79 ans et 80 ans et plus). Ainsi, pour cette épreuve, les performances des sujets décroissent avec l'âge. De la même manière, la passation auprès de participants français a montré un effet principal de l'âge à l'épreuve de dictée de mots et de pseudo-mots. Les scores obtenus diminuent progressivement avec l'âge, montrant un écart de 1,8 point entre les 18-29 ans ($M=22,8$) et les 80 ans et plus ($M=21,0$), l'épreuve étant notée sur 24 points. Nos analyses rapportent donc des corrélations significatives, corroborant ainsi les travaux de Béland et Lecours (1990). Lors de leur analyse sur les données normatives de la batterie d'aphasiologie MT-86 β , les auteurs démontrent que les sujets jeunes (19-49 ans) obtiennent de meilleurs résultats que les sujets âgés (70-87 ans) à l'association mot-image et en dictée. On pourrait penser que la perception auditive des sujets âgés ait pu affecter la dictée de pseudo-mots car aucune compensation sémantique n'est possible. Pourtant la prise en compte de la presbyacousie à travers le dépistage perceptif a permis de préciser l'interprétation des résultats obtenus : la différence entre les scores des sujets plus jeunes et plus âgés ne peut s'expliquer par une baisse auditive.

Au-delà de l'effet principal de l'âge, une interaction double a été retrouvée dans les scores des sujets français en dénomination écrite impliquant les variables du sexe et de l'âge. Très peu de données sont disponibles dans la littérature concernant ces effets en dénomination écrite. Cependant nous pouvons suggérer que ces derniers sont corrélés à ceux obtenus en dénomination orale. En effet, les performances en dénomination orale baissent avec l'âge. Le Dorze et Durocher (1992) rapportent que l'âge a un impact important sur l'exactitude des réponses. Cet effet a également été identifié par Tran et Godefroy (2011) qui mentionnent qu'en tâche de dénomination orale, les performances des 20-64 ans sont meilleures que celles des 65-79 ans, ces derniers étant eux-mêmes meilleurs que les plus de 80 ans. Enfin, Verhaegen et Poncelet (2013) confirment cet effet de l'âge où une diminution des réponses correctes s'observe chez les participants âgés de 60 ans. Ils notent également que ce phénomène s'accroît chez les sujets de plus de 70 ans. Ainsi, les auteurs décrivent une corrélation entre une diminution du niveau de précision en dénomination orale et l'âge. Notons que de tels effets ne sont pas retrouvés dans l'échantillon québécois apparié. Bien que la littérature rapporte une augmentation des temps de latence en dénomination orale chez les sujets plus âgés, cet effet n'a pas été retrouvé en épreuve de dénomination écrite de l'i-MEL, aussi bien chez les sujets français que québécois.

Une seconde interaction double a été retrouvée lors des analyses : l'âge et le niveau de scolarité influencent les scores des sujets québécois en épreuve de production écrite. Cela peut s'expliquer par une baisse des performances à l'écrit dû au vieillissement normal cognitif. Quant à l'influence du niveau scolaire, nous en discuterons ultérieurement dans la partie dédiée à l'analyse de la scolarité.

2.2 Effet du sexe

Seules deux épreuves se sont montrées sensibles à un effet principal du sexe parmi les huit étudiées en langage écrit. Ce facteur est significatif pour les scores des sujets français aux épreuves de vérification lexicale écrite et de lecture de lettres et de chiffres. En vérification lexicale écrite, les femmes obtiennent des scores plus élevés ($M=19,3$) que les hommes ($M=18,7$). Toutefois, lorsque l'on s'intéresse à la différence entre les scores des deux groupes, ceux-ci restent relativement proches, sachant que cette épreuve est notée sur 20. De la même façon, les scores obtenus par les sujets français sont marqués significativement par une influence du sexe en lecture de lettres et de chiffres. Pourtant, ce résultat reste à nuancer au regard des scores obtenus chez les femmes ($M=21,9$) et chez les hommes ($M=22$). On remarque que les données recueillies ont été marquées par un effet plafond important, le score maximal de l'épreuve se situant à 22. Cet effet plafond rend une faible différence de 0,1 significative étant donné les scores très élevés, resserrés et donc homogènes (écart-types très faibles). De cette façon, un score inférieur à 22 sera mis en avant de manière plus évidente. Cependant, en clinique, nous ne relevons pas de distinction comme le démontrent les moyennes proches obtenues dans les deux groupes. Notons que cet effet plafond était attendu dans ce type d'épreuve au regard du niveau de complexité pour des sujets témoins. En effet, la tâche reste relativement simple pour des sujets sains, ce qui explique que les scores des individus se regroupent dans les valeurs élevées. Ainsi, cette épreuve semble pertinente pour discriminer des performances pathologiques des performances normales. Il s'avère que ces effets principaux du sexe n'ont pas été retrouvés dans la population québécoise, ce qui peut s'expliquer par l'effet plafond susmentionné.

On relève également une interaction double du sexe avec le niveau de scolarité dans l'épreuve de vérification lexicale écrite chez les sujets français. En effet, les femmes de scolarité supérieure à 14 ans ont des performances meilleures que les hommes ayant le même niveau de scolarité. Ainsi, avec le même niveau scolaire, les femmes obtiennent des performances supérieures aux hommes. Une autre interaction double a été observée : les facteurs sexe et âge influencent de manière significative les scores des participants français à l'épreuve de dénomination écrite d'images. Ces mêmes effets ne sont pas présents dans l'échantillon québécois apparié.

2.3 Effet du niveau scolaire

Deux niveaux d'étude ont été définis, selon les mêmes critères qu'au Québec : inférieur ou égal à 14 ans et supérieur à 14 ans de scolarité. On pourrait supposer qu'un plus grand usage de l'écrit dans la vie professionnelle impliquerait de meilleures performances aux épreuves de langage écrit. Les analyses de variance ont par ailleurs montré que le test était significatif pour l'épreuve de dictée de phrases chez les sujets français. En effet, les sujets possédant un niveau scolaire plus élevé ont de meilleurs scores que ceux possédant un moindre niveau scolaire. Cet effet de scolarité a également été décrit, entre autres, par Béland et Lecours (1990) en épreuve de dictée du MT-86 β .

On note aussi des interactions doubles comprenant le facteur scolarité, notamment à l'épreuve de production écrite chez les sujets québécois où une influence de l'âge et de la

scolarité est présente. Cette interaction ne se retrouve pas chez les sujets français. On observe également une interaction des facteurs sexe et niveau scolaire à l'épreuve de vérification lexicale écrite chez les participants français. Ainsi, au sein même du groupe féminin, les participantes possédant un niveau supérieur à 14 ans obtiennent de meilleurs scores que celles ayant un niveau de scolarité inférieur à 14 ans. On note aussi que les femmes détenant un meilleur niveau d'étude ont des scores supérieurs aux hommes de scolarité inférieure ou égale à 14 ans. Toutefois, l'inverse ne se vérifie pas : la différence entre les scores des hommes de niveau scolaire supérieur à 14 ans et les femmes de niveau scolaire inférieur ou égal à 14 ans n'est pas significative. Par conséquent, le niveau d'étude influence les scores aux épreuves impliquant des processus plus élaborés en langage écrit. En effet, les épreuves de dictée de phrases et de production écrite font appel à une maîtrise de l'orthographe grammaticale, contrairement, par exemple, à la dictée de mots nécessitant seulement un usage de l'orthographe lexicale.

En ce qui concerne les épreuves impliquant la modalité écrite, les différences significatives ont été montrées seulement sur les scores. En effet, les ANOVA réalisées ne montrent pas d'influence des caractéristiques socio-démographiques sur le temps de réalisation des épreuves en langage écrit. Pourtant, on aurait pu supposer que l'âge influence le temps de réponse des sujets plus âgés, notamment dans la classe d'âge 80 ans et plus, qui présentent classiquement un ralentissement de la vitesse de traitement. Il en est de même pour l'influence du niveau scolaire sur le temps de réponse. On aurait pu s'attendre à ce qu'une plus grande utilisation de l'écrit et donc une plus grande familiarité avec le code écrit permette une meilleure vitesse de lecture et d'écriture.

2.4 Ouverture des analyses socio-démographiques à d'autres épreuves de l'i-MEL

Les résultats statistiques obtenus aux autres épreuves de la batterie i-MEL ont fait l'objet d'une analyse spécifique au sein des trois autres mémoires de cette étude : épreuves préalables, sémantico-phonologiques et syntaxico-discursives. L'ensemble des effets socio-démographiques sont présentés dans un tableau récapitulatif (cf. Annexe n°3). Nous discuterons ici des résultats aux épreuves ayant des caractères communs avec celles analysées dans ce mémoire.

Par analogie à l'épreuve de dénomination écrite, nous nous intéressons aux effets socio-démographiques en dénomination orale. Une interaction double (âge et scolarité) a été retrouvée sur le temps de réponse des sujets français. Cette observation concorde avec les travaux menés par Verhaegen et Poncelet (2013) et Tran et Godefroy (2011) qui relatent une augmentation des temps de latence de dénomination orale avec l'âge. De plus, les scores obtenus augmentent avec le niveau scolaire à cette même tâche (Tran & Godefroy, 2011). Cette observation ne se retrouve pas en dénomination écrite.

En vérification lexicale orale, aucun effet n'a été retrouvé, que ce soit en population française ou québécoise. Néanmoins, un effet principal de l'âge et du sexe ainsi qu'une interaction double (sexe et scolarité) ont été observés en vérification lexicale écrite, comme décrits plus haut. Ces effets semblent donc être propres à la modalité écrite du langage.

Trois autres épreuves faisant intervenir la compréhension écrite du langage ont été analysées : la compréhension morphosyntaxique écrite, la compréhension de paragraphes écrits et la compréhension écrite de texte long. Concernant la compréhension morphosyntaxique écrite, un effet principal de la scolarité a été relevé sur les scores des sujets français : ceux possédant un meilleur niveau de scolarité obtiennent des scores plus élevés. Un effet de l'âge a également été observé sur les scores des participants français à l'épreuve de compréhension de paragraphes écrits où les 18-29 ans obtiennent de meilleurs scores que les 30-59 ans et les 60-79 ans. On ne retrouve pas de différence significative avec la tranche d'âge la plus élevée (80 ans et plus), cependant nous rappelons que l'effectif de ce sous-groupe étant peu élevé, cette analyse mériterait d'être poursuivie sur un échantillon davantage représentatif et statistiquement plus fiable. Enfin, un effet principal de l'âge est présent sur le temps de réponse des sujets québécois en compréhension écrite de texte long. Les sujets québécois âgés d'au moins 80 ans sont significativement plus lents que les 18-29 ans et les 30-59 ans. Ainsi, dans les trois épreuves impliquant des processus de compréhension écrite plus élaborée, des effets d'âge et de scolarité sont présents. Cela rejoint des observations décrites dans ce mémoire. Toutefois, les effets entre les deux pays ne sont pas semblables. Une comparaison des performances franco-québécoises semble donc pertinente dans cette étude.

3. Analyse des différences franco-québécoises

L'enjeu de cette validation transculturelle était de récolter des données spécifiques à la population française compte tenu de l'influence de la culture sur les fonctions cognitives. Ce recueil de scores auprès des sujets français permet alors d'étudier l'impact des possibles biais culturels sur les fonctions langagières, notamment le vocabulaire ou la familiarité des concepts (Macoir et al., 2016). Par ailleurs, l'évaluation interculturelle est dépendante des aspects méthodologiques puisqu'ils garantissent des comparaisons valides entre les deux cultures étudiées. Ainsi, il s'agit d'éviter qu'un facteur culturel affecte les résultats aux épreuves de façon différentielle entre les deux pays (Van de Vijver & Poortinga, 1997).

3.1 Comparaison à l'échelle des épreuves de langage écrit

Dans cette optique, nous avons comparé les données franco-québécoises. Parmi les épreuves qui font l'objet d'une analyse dans ce mémoire, sept différences significatives ont été rapportées entre les sujets français et leurs homologues québécois appariés.

A l'épreuve de dénomination écrite, où le score maximal est de 36, les sujets québécois obtiennent de meilleurs scores ($M=34,93$) que les français ($M=34,09$). Les écart-types étant sensiblement similaires entre les deux pays, la dispersion équivalente des résultats ne peut expliquer la différence des moyennes entre les deux populations. Ce faible écart entre les scores peut se justifier par les consignes données à l'examineur dans le manuel : ce dernier peut accepter un synonyme s'il juge que la réponse du participant correspond à l'image présentée. Ainsi, la liberté offerte à l'examineur de juger de l'exactitude de la réponse du participant nuit

à la fidélité inter-juge et peut ainsi expliquer les différences entre les scores français et québécois.

Une autre différence entre les deux populations a été retrouvée concernant le temps à l'épreuve de dénomination écrite. Les sujets français ($M=118,63$) sont plus rapides que les québécois ($M=172,85$) à cette épreuve. Cette différence importante de 54,32 secondes semble s'expliquer par un biais méthodologique. En effet, la particularité de cette épreuve est de ne pas prendre en compte le temps d'écriture des mots. Les concepteurs de l'i-MEL précisent que l'examineur doit appuyer sur le bouton « réponse donnée » dès que le participant commence à écrire la réponse. En effet cette épreuve a pour but de mesurer l'accès lexical à l'écrit et non la vitesse du geste graphique. De ce fait, il semble possible que les examinateurs français et québécois n'aient pas respecté la consigne avec la même rigueur d'application.

Les sujets français ont de meilleurs scores à l'épreuve de production écrite. En effet, les participants français obtiennent en moyenne 37,36/38 comparativement aux québécois dont la moyenne se situe à 36,89/38. Notons que l'écart-type est plus important au sein de la population québécoise ($ET=1,564$) que dans la population française ($ET=1,147$). Cette dispersion québécoise plus élevée reflète des scores davantage hétérogènes et peut justifier la différence de moyennes entre les deux pays. En effet, les écart-types permettent d'examiner et de comparer le niveau de dispersion des sujets entre les deux pays et devraient être d'étendue modérée (Vallerand, 1989).

A l'épreuve de lecture de lettres et de chiffres, une différence en termes de vitesse a été relevée entre les deux groupes. Les participants québécois mettent en moyenne 35,07 secondes pour lire les 22 items contrairement aux français qui y parviennent en moyenne en 38,49 secondes. Cette faible différence de vitesse peut s'expliquer par le fait que de nombreux sujets ont confondu l'item [l] qui représente un [L] minuscule et non un [l] majuscule. Si tel était le cas, l'examineur devait indiquer au sujet qu'il s'agissait de lettres minuscules et accepter l'auto-correction. Ce délai de précision a donc pu jouer dans la prise en compte du temps selon les réponses apportées par les sujets.

Concernant l'épreuve de lecture de mots et de pseudo-mots notée sur 24, les sujets français obtiennent des scores meilleurs ($M=23,72$) que les québécois ($M=23,34$). Il convient néanmoins de nuancer cette différence au vu des moyennes très rapprochées entre les deux groupes qui sont presque similaires. Vallerand (1989) précise par ailleurs que les moyennes et écart-types sont nécessaires pour apprécier la distinction entre les deux groupes et que si tel est le cas, les moyennes n'infirmant pas nécessairement la version originale : il est probable que les différences soient liées à l'échantillon.

Une différence a été observée à propos du temps de lecture de mots et de pseudo-mots entre les sujets français qui mettent en moyenne 38,25 secondes à lire les 24 items, contrairement aux québécois qui parviennent à lire l'ensemble des stimuli en 41,59 secondes. La comptabilisation du temps de réponse se faisant automatiquement avec la tablette, l'écart ne peut s'expliquer par un biais méthodologique. Cependant, les moyennes sont relativement proches et nous relevons un écart-type légèrement plus important au sein de la population québécoise ($ET=7,992$) à la différence de la population française ($ET=7,619$). De la même manière, les scores plus disparates peuvent expliquer cet écart de performance entre les deux groupes.

Enfin, une dernière différence franco-québécoise a été retrouvée sur le temps d'écriture en dictée de mots et de pseudo-mots. Les sujets français mettant en moyenne 129,05 secondes pour réaliser l'épreuve, contrairement aux participants québécois qui eux mettaient en moyenne 150,78 secondes. Cet écart important entre les deux groupes peut s'expliquer par un défaut du programme informatique en France. En effet, la tablette n'a pas enregistré le temps de passation de 10 sujets français. Les analyses statistiques sur le temps de cette épreuve ont donc été faites sur 90 sujets au lieu de 100 en France.

3.2 Comparaison élargie à d'autres épreuves de la batterie

Tout comme les analyses socio-démographiques, seuls les résultats des épreuves en langage écrit ont été discutés ici. Les résultats aux autres épreuves de la batterie sont consultables en annexes (cf. Annexe n°3). Toutefois, il semble pertinent de s'intéresser aux résultats des épreuves présentant des similitudes avec celles traitées dans ce mémoire.

En ce sens, on retrouve à l'épreuve de dénomination orale des scores meilleurs chez les sujets français ($M=35,62$) par rapport aux sujets québécois ($M=34,67$). Cette observation va à l'encontre de celle faite en dénomination écrite où la population québécoise était meilleure en exactitude que la population française. En revanche, la même tendance est observée en vitesse de dénomination, que ce soit à l'écrit ou à l'oral, les sujets français sont en moyenne plus rapides que les sujets québécois. Ces différences peuvent être imputées aux erreurs méthodologiques de passation au même titre que l'épreuve de dénomination écrite. En effet, à l'épreuve de dénomination orale, il est aussi demandé à l'examineur d'appuyer sur le bouton « réponse donnée » dès que le sujet commence à répondre, ce qui permet d'arrêter le chronomètre. Le respect strict de la consigne de passation n'a peut-être pas été appliqué de la même façon entre les examinateurs des deux pays. Enfin, aux épreuves de vérification lexicale orale, compréhension morphosyntaxique écrite, compréhension de paragraphes écrits et compréhension écrite de texte long, aucune différence n'est ressortie significative lors de la comparaison des performances franco-québécoises.

Ainsi, la plupart des différences entre les sujets français et québécois se relativisent par des moyennes proches ou des biais méthodologiques dont nous ne pouvons sous-estimer leur influence. En effet, les conditions de passation ne se montrent pas neutres quant au recueil des résultats. Grégoire (2004) décrit que les modalités d'administration sont elles-mêmes imprégnées par la culture. Par ailleurs, Flaherty et al. (1988) exposent la problématique de l'équivalence technique où il convient de s'interroger sur les conditions de passation pouvant affecter de façon différente les résultats dans les deux cultures. Cependant, nous pouvons nuancer ces arguments dans le cadre de cette étude puisque les deux populations étudiées ici partagent une culture occidentale commune. Tenant compte de ces différentes remarques, l'équipe de recherche franco-québécoise devra statuer si les données pourront ou non être fusionnées.

4. Intérêts et limites

4.1 Intérêts et limites de la présente étude

La validation de l'i-MEL en France permet d'élargir l'utilisation de l'outil à un autre territoire où les besoins sont identiques (Joanette et al., 2014). Dans cette perspective, une validation transculturelle avait été entamée en 2020-2021 mais n'avait pu aboutir en raison de la crise sanitaire liée au Covid. La poursuite de cette recherche a donc permis une étude plus approfondie en termes de répartition des sujets et plus élargie en effectifs. En effet, l'étude préliminaire traduisait une sur-représentation des certains sous-groupes tels que les participants de 18-29 ans ou ceux possédant un niveau de scolarité supérieur à 14 ans. A l'inverse, certains sous-groupes n'étaient pas du tout représentés comme les participants de 80 ans et plus (n=0). La suite et finalisation du projet a ainsi permis d'obtenir une répartition respectant davantage celle de la population générale avec des effectifs assurant une puissance statistique plus fiable. Enfin, les échanges pendant la validation transculturelle entre les étudiantes en France et concepteurs de l'i-MEL ont donné lieu à des améliorations et précisions dans le manuel. L'insertion d'observations qualitatives, comme l'enrichissement des consignes de passation, assure par conséquent la fonctionnalité de l'outil (Bourgeois et al., 2019).

Concernant les difficultés liées à l'étude, deux principales problématiques ont été rencontrées. La première a concerné le recrutement long et fastidieux des sujets puisqu'il fallait compter environ cinq heures de passation par participant. Les contraintes organisationnelles étaient donc importantes. Par ailleurs, l'exigence en termes de recrutement s'est avérée élevée pour obtenir une population témoin homogène. Malgré un souci du respect d'uniformité des sous-groupes (par exemple autant de participants par classe d'âge), certains demeurent toujours non équivalents en termes d'effectif. La répartition par variables socio-démographiques de notre population n'est donc pas parfaitement homogène mais s'attache tout de même à être représentative de la population française. La catégorie des 80 ans et plus en est une bonne illustration : elle représente 8 % des sujets dans cette étude mais reste représentative de la population générale française selon les données de l'INSEE. Au regard de la validation future, il serait intéressant de recruter davantage de sujets âgés puisque la fréquence des accidents vasculaires cérébraux augmente avec l'âge (Bohic, 2012). Enfin, il est important de rappeler que la répartition géographique des participants constitue une autre limite à cette étude puisque la majorité de la recherche s'est déroulée dans les Hauts-de-France.

La seconde difficulté s'est traduite par des biais méthodologiques, inhérents à la passation de la batterie. Nous pouvons, entre autres, relever les différences décrites plus haut quant à la rigueur d'application des consignes de passation. En effet, la question de la fidélité inter-juge se pose aussi bien entre les examinateurs français et québécois qu'entre les examinateurs d'une même équipe de recherche. Pourtant une fiabilité inter-juge est essentielle pour que les scores ne fluctuent pas lorsque différents examinateurs administrent les épreuves (Friberg, 2010). Notons que les concepteurs de l'i-MEL prévoient une bonne fidélité inter-juge pour les épreuves dont la cotation est dichotomique (0/1 point) ou lorsque le participant est directement invité à répondre sur la tablette. Pour d'autres épreuves mettant en jeu le jugement perceptif du clinicien (par exemple épreuves prosodiques), il est conseillé de réaliser une

cotation a posteriori. Ainsi, pour ces dernières, même en présence du manuel, la multiplicité des expérimentateurs peut constituer un biais. De plus, les étudiantes ayant mené l'étude préliminaire l'année dernière, ont manqué de précisions sur certaines conditions de passation puisque le manuel n'était que partiellement rédigé à ce moment-là. La maîtrise de l'outil s'est également améliorée au fil de l'étude et des différences entre les premières et dernières passations sont envisageables. En effet, une plus grande habileté à faire passer le test s'est développée bien que les étudiantes aient été formées et se soient assurées de maîtriser l'outil avant le début du projet. Enfin, la plupart des sujets recrutés en France font plus ou moins partie de l'entourage des étudiantes, limite qu'il s'agit de considérer ici. La relation entre l'expérimentateur et le sujet risque d'affecter les résultats puisque les caractéristiques propres à l'examineur peuvent s'avérer perturbatrices dans la situation d'examen (Grégoire, 2004).

4.2 Intérêts et limites de l'outil

Le travail mené au Québec et en France constitue une réponse aux attentes cliniques des orthophonistes dans les deux pays. Les cliniciens disposeront ainsi d'un outil d'évaluation fin et normalisé permettant d'identifier les troubles de manière rigoureuse et de mieux préciser les objectifs thérapeutiques (Trauchessec, 2018). Les concepteurs de l'i-MEL ont répondu à des exigences psychométriques élevées, notamment par l'informatisation de la batterie. Les avantages d'un tel format sont divers : contrôle du temps et de l'ordre de présentation du stimulus ; recueil des réponses et calcul automatique des scores et temps ; automatisation de la cotation. Ces facteurs participent à la bonne standardisation de la batterie en réduisant les éventuelles erreurs manuelles du clinicien et en mesurant finement les critères tels que les temps de réponses, qui sont, comme on l'a vu pour la dénomination orale, essentiels pour mesurer les temps de latence.

De nombreux tests existent en aphasiologie mais peu d'entre eux permettent d'évaluer de manière complète les troubles acquis du langage écrit. Monetta et ses collègues (2016) ont recensé et décrit tous les tests franco-québécois validés et/ou normés de 1980 à 2014, parmi eux nous retiendrons ceux évaluant les composantes du langage écrit : le BAT Bilingual Aphasia Test (Golblum & Paradis, 1989) ; la BECLA Batterie d'Evaluation Cognitive du Langage chez l'Adulte (Macoir et al., 2005) ; l'épreuve de compréhension écrite (Fontaine & Joubert, 2010) ; le MT-86 protocole d'examen linguistique de l'aphasie Montréal-Toulouse (Nespoulous et al., 1986, révisé en 1992) ; le « Pyramids and Palm trees Test » (Howard & Patterson, 1992) ; le MEC-P protocole Montréal d'Evaluation de la Communication de poche (Ferré et al., 2011) ; le TCC protocole d'évaluation de la communication (Bisson, 2012) et le TCT test de compréhension de textes (Chesneau, 2012). D'autres outils, évaluant en partie le langage écrit, n'ont pas été cités par Monetta et ses collègues (2016) tels que le BDAE Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1972) ; le BIA Bilan Informatisé d'Aphasie (Gatignol et al., 2012) ; l'EDA Examen des Dyslexies Acquisées (Lemay, 1990) et la BETL Batterie d'Evaluation des Troubles Lexicaux (Tran & Godefroy, 2015). L'ensemble de ces tests présentent des limites quant à l'évaluation des troubles du langage écrit, soit par leur ancienneté, leur évaluation partielle des composantes écrites du langage ou par leurs qualités psychométriques trop faibles ou non étudiées. De fait, la validité et la fidélité ne sont pas disponibles pour la BECLA, l'épreuve de compréhension écrite, le MT-86, le Pyramids and

Palm trees Test et le TCC. D'autres présentent des défauts de normalisation : le TCC ayant été normalisé sur seulement 80 sujets ou le TCT n'ayant pas inclus dans sa normalisation les sujets de 40-60 ans ayant une scolarité supérieure ou égale à 12 ans. Les épreuves évaluant le langage écrit sont également souvent au second plan de l'évaluation et donc incomplètes. L'i-MEL est organisé de façon à avoir un module spécifiquement dédié aux épreuves de langage écrit et se veut plus représentatif de l'ensemble des composantes du langage écrit.

Enfin, il semble pertinent de discuter des limites identifiées à propos de la batterie i-MEL. La première concerne le coût financier que représente l'acquisition du test et de son support, à savoir un Ipad. L'informatisation de l'évaluation peut également composer un frein pour certains cliniciens qui se refuseraient à utiliser un tel format. Il est important de rappeler que les modalités de passations informatisées peuvent aussi avoir un impact sur les performances des patients étant donné les « effets positifs de la familiarité avec la technologie sur les tests informatisés » (Parsey & Schmitter-Edgecombe, 2013). Un autre frein porte sur la composante écrite du langage qui reste incomplète dans l'i-MEL. Certains processus ne peuvent être évalués puisque des épreuves d'écriture automatique ou de copie ne sont pas proposées dans la batterie. Pour finir, une dernière limite concerne la reprise de certaines épreuves du MEC-P (Ferré et al., 2011) sans en avoir actualisé le contenu. Certaines formulations restent désuètes et ne prônent pas une égalité entre les hommes et les femmes. C'est le cas, par exemple, de l'épreuve d'interprétation d'actes de langage indirects où quelques situations présentées ont été perçues comme malvenues par de nombreux participants à l'étude.

5. Issue de l'étude et perspectives futures

Le travail collaboratif entre les étudiantes en France a permis de récolter suffisamment de données auprès de 100 sujets pour obtenir des résultats statistiquement puissants. Sur l'ensemble de la batterie comprenant 51 épreuves, nous retiendrons que 33 sont sensibles à un effet des caractéristiques socio-démographiques en France mais seulement 8 le sont au Québec. La seconde analyse a montré que 25 épreuves présentent une différence statistiquement significative entre les sujets français et québécois :

- Dix-neuf d'entre elles présentent une différence sur les scores : automatismes verbaux, répétition de mots et pseudo-mots, répétition de phrases, questions oui/non, production écrite, discours narratif et descriptif (film) version A, vigilance auditive, métaphonologie, questionnaire sémantique, interprétation de métaphores, prosodie linguistique compréhension, prosodie émotionnelle répétition, prosodie émotionnelle production, décision lexicale, dessin, identification auditive non verbale, fluence sémantique, fluence libre, fluence alternée. Notons que parmi ces épreuves, 5 obtiennent un pourcentage de différence supérieur à 5 % : le discours narratif et descriptif (film) version A, le dessin et les fluences sémantique, libre et alternée.
- Trois d'entre elles présentent des différences sur le temps de réponse : lecture de lettres et de chiffres, dictée de mots et de pseudo-mots version A, vigilance visuelle.

- Trois présentent une différence aussi bien sur les scores que sur les temps de réponse : dénomination orale version A, lecture de mots et de pseudo-mots version A, dénomination écrite version A.

L'objectif était donc d'obtenir un outil à la fois valide et d'universalité culturelle suffisante pour que son utilisation soit possible en France. Pour cela, peu d'adaptations semblent nécessaires. La majorité des épreuves présentant une différence franco-québécoise obtiennent des résultats à nuancer au vu du pourcentage de différence inférieur à 5 %. Ainsi, certaines ressortent statistiquement différentes mais cette différence n'est pas retenue comme pertinente en clinique. La batterie i-MEL semble donc en grande partie adaptée, excepté pour le discours narratif et descriptif (film) version A, le dessin et les fluences sémantique, libre et alternée. Deux options sont alors envisageables pour ces dernières : elles pourraient faire l'objet d'une note dans le manuel à destination des cliniciens exerçant en France ou d'une normalisation spécifiquement française.

Il serait également intéressant de poursuivre la validation transculturelle sur la version B de l'outil, c'est-à-dire la version retest. En effet, par manque de temps et d'expérimentateurs, l'étude n'a été effectuée que sur la version A des 14 épreuves possédant une version B. A l'échelle de ce mémoire traitant les modalités écrites du langage, la validation transculturelle de la version B de la lecture et dictée de mots et de pseudo-mots et de la dénomination écrite permettrait de statuer si les différences entre la France et le Québec observées en version A sont aussi présentes dans la seconde version.

Conclusion

Ce mémoire a ainsi permis l'aboutissement d'un projet débuté en 2020-2021. L'objectif de ce dernier était la validation transculturelle d'un nouvel outil d'évaluation des troubles acquis du langage et de la communication : l'i-MEL. Ce processus a compris plusieurs étapes, la première étant le recrutement et la passation de la batterie auprès de cent sujets témoins français. Nous les avons répartis, le plus homogènement possible, selon trois caractéristiques socio-démographiques (âge, sexe, niveau socio-culturel) pour étudier leur influence sur les performances linguistiques des participants. Nous avons réalisé cette même analyse avec les données québécoises pour voir si des effets similaires étaient retrouvés dans les deux pays. Ensuite, les sujets français ont été appariés avec leurs homologues québécois permettant ainsi de comparer les résultats obtenus au Québec et en France. Ce travail a été possible grâce à la collaboration avec trois autres étudiantes puisque chacune analysait une composante spécifique de la batterie, il s'agissait ici du langage écrit. Si les résultats étaient semblables entre les deux populations et que l'aspect culturel ne semblait pas intervenir sur les performances des sujets, la diffusion de l'i-MEL était alors possible en France. Dans le cas contraire, une adaptation serait à envisager pour les épreuves présentant une différence significative.

L'analyse statistique des données a mis en avant un effet des caractéristiques socio-démographiques en France sur cinq épreuves parmi les huit étudiées dans ce mémoire. Nous n'avons retrouvé un tel effet que pour une épreuve au Québec. Concernant l'ensemble de la batterie, 33 épreuves sur les 51 sont sensibles à un effet du sexe, de l'âge et/ou du niveau socio-

culturel en France, et seulement 8 au Québec. En parallèle, la comparaison entre les performances des sujets français et québécois a mis en avant une différence significative pour cinq épreuves en langage écrit. Sur l'ensemble de batterie, nous retiendrons que 25 le sont. Toutefois, la plupart de ces épreuves présentent une différence statistique significativement faible (pourcentage de différence inférieur à 5 %). Cinq épreuves présentent une différence très significative (supérieure à 5 %) sur les scores : le discours narratif et descriptif (film) version A, le dessin et les fluences sémantique, libre et alternée. Plusieurs perspectives sont alors possibles pour ces épreuves : le stipuler dans le manuel pour les cliniciens en France afin qu'ils en tiennent compte de manière qualitative dans leur cotation ou effectuer une normalisation spécifiquement française.

Ainsi, l'i-MEL est un outil récent qui présente des qualités psychométriques reconnues et en partie validé selon une méthodologie transculturelle. La création de cette batterie d'évaluation pour personnes cérébro-lésées et sa validation transculturelle apportent une réponse aux besoins cliniques des orthophonistes francophones.

Bibliographie

- Bauer, R. M., Iverson, G. L., Cernich, A. N., Binder, L. M., Ruff, R. M., & Naugle, R. I. (2012). Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Archives of Clinical Neuropsychology*, 27(3), 362-373.
- Béland, R., & Lecours, A. R. (1990). The MT-86 β aphasia battery: A subset of normative data in relation to age and level of school education. *Aphasiology*, 4(5), 439-462.
- Bisson, J. (2012). Protocole d'évaluation de la communication chez la clientèle adulte traumatisée craniocérébrale. Institut de réadaptation en déficience physique de Québec (IRDPPQ).
- Bohic, N. (2012). Prévention des accidents vasculaires cérébraux et vieillissement. *Gerontologie et société*, 35(5), 217-227.
- Bourgeois, M. È., Bergeron, A., Perron, M., & Martel-Sauvageau, V. (2019). Développement, validation et normalisation de la Batterie d'évaluation de la compréhension syntaxique : Une collaboration Québec-Suisse. *Journal of Speech-Language Pathology and Audiology*, 43(2), 109-120.
- Caron, J. (1999). Un guide de validation transculturelle des instruments de mesure en santé mentale. <https://doi.org/10.13140/RG.2.1.2157.7041>
- Chesneau, S. (2012). T.C.T : Test de Compréhension de Textes - Évaluation de la compréhension de textes chez les adultes de 16 à 80 ans. Paris : Mot à mot.
- Chomel-Guillaume, S., Leloup, G., & Bernard, I. (2010). Évaluation et rééducation des aphasies. Paris : Masson.
- Ciesielska, N., Sokołowski, R., Mazur, E., Podhorecka, M., Polak-Szabela, A., & Kędziora-Kornatowska, K. (2016). Is the Montreal Cognitive Assessment (MoCA) test better suited than the Mini-Mental State Examination (MMSE) in mild cognitive impairment (MCI) detection among people aged over 60 ? Meta-analysis. *Psychiatr. Pol.*, 50(5), 1039-1052.
- Cruice, M., Worrall, L., Hickson, L., & Murison, R. (2003). Finding a focus for quality of life with aphasia: Social and emotional health, and psychological well-being. *Aphasiology*, 17(4), 333-353.
- Deleuze, A., Ferré, P., Ansaldo, A.I., & Joannette, Y., (2016, 8-9 décembre). Evaluation de la communication de l'adulte cérébrolésé et tablette numérique : quels apports pour la pratique clinique ? Rencontres d'orthophonie, Congrès de l'Unadreo, chap. 4.
- De Partz, M. P. (2014). L'évaluation des troubles du langage écrit. Dans X. Seron et M. Van der Linden (dir.), *Traité de neuropsychologie clinique de l'adulte* (2e éd., vol. 1, p.297-319).
- De Partz, M. P. (2018). Les troubles du buffer graphémique : de la théorie cognitive à l'évaluation clinique. *Rééducation Orthophonique*, 274, 147-168.

- Ember, C. R., & Ember, M. (2011). A basic guide to cross-cultural research. User's Guide : HRAF Collection of Ethnography. <https://hraf.yale.edu/cross-cultural-research/basic-guide-to-cross-cultural-research/>
- Engelter, S. T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., Gutzwiller, F., & Lyrer, P. A. (2006). Epidemiology of aphasia attributable to first ischemic stroke : incidence, severity, fluency, etiology, and thrombolysis. *Stroke*, 37(6), 1379-1384.
- Ferré, P., Lamelin, F., Côté, H., Ska, B. & Joannette, Y. (2011). Protocole MEC-P : Protocole Montréal d'Évaluation de la Communication de Poche. Isbergues : Ortho Édition.
- Flaherty, J. A., Gaviria, F. M., Pathak, D., Mitchell, T., Wintrob, R., Richman, J. A., & Birz, S. (1988). Developing instruments for cross-cultural psychiatric research. *Journal of Nervous and Mental Disease*, 176(5), 257-263.
- Flamand-Roze, C., Roze, E., & Denier, C. (2012). Troubles du langage et de la déglutition à la phase aiguë des accidents vasculaires cérébraux : Outils d'évaluation et intérêt d'une prise en charge précoce. *Revue Neurologique*, 168(5), 415-424. <https://doi.org/10.1016/j.neurol.2011.10.009>
- Fontaine, F. & Joubert, S. (2010). Épreuve de Compréhension écrite –14 questions : Données normatives pour une population francophone âgée vivant au Québec depuis au moins 40 ans. Montréal : Institut universitaire de gériatrie de Montréal, Centre de recherche.
- Friberg, J. C. (2010). Considerations for test selection: How do validity and reliability impact diagnostic decisions?. *Child Language Teaching and Therapy*, 26(1), 77-92.
- Gadet, F. (2021). Variation. *Langage et société*, (HS1), 331-336.
- Gatignol P., Jutteau S., Oudry M., & Weill-Chounlamountry A. (2012). BIA : Bilan Informatisé Aphasie. Isbergues : Ortho Édition.
- Golblum, M.C. & Paradis, M. (1989). Bilingual Aphasia Test (French Version). Québec : McGill University.
- Goodglass, H. & Kaplan, E. (1972). BDAE : Boston Diagnostic Aphasia Examination, échelle française : Mazaux J.M., Orgogozo J.M. Paris : EAP.
- Grégoire, J. (2004). L'examen clinique de l'intelligence de l'adulte. Editions Mardaga.
- Howard, D. & Patterson, K. (1992). Pyramids and Palm Trees: A test of semantic access from pictures and words. Bury St. Edmunds : Thames Valley Test Company.
- 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. <https://doi.org/10.1080/02687038.2013.805728>
- Jamovi project (2021). jamovi (Version 1.6) [Computer Software]. Retrieved from <https://www.jamovi.org>

- Joanette, Y., Ska, B., & Côté, H. (2004). Protocole MEC : Protocole Montréal d'Évaluation de la Communication. Ortho Edition.
- Joanette, Y., Ferré, P., Côté, H. (2014). Recherche et pratique clinique en orthophonie : du transfert à l'échange de connaissances. *Rééducation orthophonique*, 257, 131-141.
- Lambert, J. (2008). Chapitre 24. Langage et parole. Dans F. Eustache, B. Lechevalier et F. Viader (dir.), *Traité de neuropsychologie clinique* (p. 439-541). De Boeck Supérieur. <http://www.cairn.info/traite-de-neuropsychologie-clinique--9782804156787-page-439.htm>
- Le Dorze, G., & Durocher, J. (1992). The effects of age, educational level, and stimulus length on naming in normal subjects. *Journal of Speech-Language Pathology and Audiology*, 16(1), 21-29.
- Le Dorze, G. L., & Brassard, C. (1995). A description of the consequences of aphasia on aphasic persons and their relatives and friends, based on the WHO model of chronic diseases. *Aphasiology*, 9(3), 239-255.
- Lee, H., Lee, Y., Choi, H. & Pyun, S.B. (2015). Community Integration and Quality of Life in Aphasia after Stroke. *Yonsei Medical Journal*, 56(6), 1694-1702.
- Lemay, M.-A. (1990). EDA : Examen des Dyslexies Acquis. Montréal : Éditions PointCarré.
- Macoir, J., Gauthier, C., & Jean, C. (2005). BECLA : Batterie d'Évaluation Cognitive du Langage chez l'Adulte. Québec : Université Laval.
- Macoir, J., Gauthier, C., Jean, C., & Potvin, O. (2016). BECLA, a new assessment battery for acquired deficits of language: Normative data from Quebec-French healthy younger and older adults. *Journal of the neurological sciences*, 361, 220-228.
- Macoir, J., Fossard, M., Lefebvre, L., Monetta, L., Renard, A., Tran, T. M. & Wilson, M. A. (2017). Detection Test for Language Impairments in Adults and the Aged. A New Screening Test for Language Impairment Associated With Neurodegenerative Diseases: Validation and Normative Data. *American Journal of Alzheimer's Disease & Other Dementias*, 32 (7), 382-392.
- Marquer, C. (2014). Validation transculturelle d'une échelle en contexte humanitaire : dépistage des difficultés psychologiques du jeune enfant haoussa au Niger (Maradi) [thèse de doctorat, Université Paris Descartes]. <https://tel.archives-ouvertes.fr/tel-01195996>
- Martin, Y. (2018). Les perturbations de la communication chez la personne aphasique. *Rééducation Orthophonique*, 274, 169-184.
- Mazaux, J.-M., (2008). Aphasie, Evolution des concepts, évaluation et rééducation. DES médecine Physique et Réadaptation, 2008, module Neuropsychologie, Cofemer.
- Mazaux, J.-M., Daviet, J.-C., Darrigrand, B., Stuit, A., Muller, F., Dutheil, S., Joseph, P.-A., & Barat, M. (2006). Difficultés de communication des personnes aphasiques. Dans *Évaluation des troubles neuropsychologiques en vie quotidienne* (p. 73 82). Springer-Verlag. https://doi.org/10.1007/2-287-34365-2_8

- Monetta, L. (2016). Fiches descriptives des outils validés et/ou normés en franco-québécois pour l'évaluation du langage et de la parole, de 1980 à 2014. Repéré à <http://www.repar.veille.qc.ca/fichier.php/92/Fiches%20descriptives%20orthophonie%20.pdf>
- Monetta, L., Desmarais, C., MacLeod, A. A., St-Pierre, M. C., Bourgeois-Marcotte, J., & Perron, M. (2016). Recension des outils franco-québécois pour l'évaluation des troubles du langage et de la parole. *Revue canadienne d'orthophonie et d'audiologie*, 40, 165-175.
- Nespoulous, J.-L., Lecours, A.R., Lafond, D., Lemay, A., Puel, M., Joannette, Y., Cot, F., & Rascol, A. (1992). Protocole Montréal-Toulouse d'examen linguistique de l'aphasie MT 86. Isbergues, France : Ortho Édition.
- Parsey, C. M., & Schmitter-Edgecombe, M. (2013). Applications of technology in neuropsychological assessment. *The Clinical Neuropsychologist*, 27(8), 1328-1361.
- Pillon, A., & De Partz, M. P. (2014). Sémiologie, syndromes aphasiques et examen clinique des aphasies. *Traité de neuropsychologie clinique de l'adulte*, 1, 249-265.
- Reiff Cherney, L. (2004). Aphasia, alexia, and oral reading. *Topics in stroke rehabilitation*, 11(1), 22-36.
- Spaccavento, S., Craca, A., Del Prete, M., Falcone, R., Colucci, A., Di Palma, A., & Loverre, A. (2013). Quality of life measurement and outcome in aphasia. *Neuropsychiatric Disease and Treatment*, 10, 27-37. <https://doi.org/10.2147/NDT.S52357>
- Starrfelt, R., & Behrmann, M. (2011). Number reading in pure alexia - A review. *Neuropsychologia*, 49(9), 2283-2298.
- Starrfelt, R., Olafsdóttir, R. R., & Arendt, I. M. (2013). Rehabilitation of pure alexia: A review. *Neuropsychological rehabilitation*, 23(5), 755-779.
- Tran, T. M., & Godefroy, O. (2011). La Batterie d'Évaluation des Troubles Lexicaux : effet des variables démographiques et linguistiques, reproductibilité et seuils préliminaires. *Revue de neuropsychologie*, 3(1), 52-69.
- Tran, T.-M., & Godefroy, O. (2015). BETL : Batterie d'évaluation des troubles lexicaux. Isbergues : Ortho Édition.
- Trauchessec, J. (2018). Aphasie et troubles cognitifs : des concepts à l'évaluation. *Rééducation orthophonique*, 274, 295-320.
- Vallerand, R. J. (1989). Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques : Implications pour la recherche en langue française. *Canadian Psychology*, 30(4), 662.
- Van de Vijver, F. J., & Poortinga, Y. H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment*, 13(1), 29-37.
- Verhaegen, C., & Poncelet, M. (2013). Changes in naming and semantic abilities with aging from 50 to 90 years. *Journal of the International Neuropsychological Society*, 19(2), 119-126.

Liste des annexes

Annexe n°1 : Fiche d'inclusion à la validation transculturelle de l'i-MEL

Annexe n°2 : Résultats détaillés du test-t de student aux épreuves de langage écrit

Annexe n°3 : Synthèse des résultats de la validation transculturelle de l'i-MEL