

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

Anouk MALEWSKI

soutenu publiquement en juin 2022

**Evaluation du langage en péri-opératoire de la
chirurgie éveillée des gliomes
Normalisation d'un test de dénomination orale**

MEMOIRE dirigé par

Isabelle POISSON, Orthophoniste, Hôpital Lariboisière, Paris

Corinne HENGUELLE, Orthophoniste, Hôpital Salengro, Lille

Lille – 2022

Remerciements

Je tiens à remercier Isabelle Poisson et Marion Barberis pour leur encadrement chaleureux, leurs conseils pertinents et la confiance qu'elles m'ont accordée pour ce projet. Un grand merci Isabelle de m'avoir initiée et formée à tes côtés à la neurologie : ta passion de transmettre et ton exigence bienveillante m'ont énormément apporté. Merci également à Marion, Cécile et Sophie pour leur disponibilité et leur partage de savoirs. Merci à Emmanuel Mandonnet : je mesure la chance d'avoir pu assister à des chirurgies éveillées et cela m'a permis d'être en plein cœur de mon sujet d'étude. Merci à Corinne Henguette d'avoir accepté de me co-encadrer : merci pour votre disponibilité et vos relectures avisées.

Mes remerciements les plus sincères vont aux étudiantes de Master 1 Aude Banchet, Lucie Boutié, Marie Chauvel, Marion Lalanne, Fanny Lecornu et Emmanuelle Sissou qui ont donné de leur énergie et de leur temps pour m'aider dans les passations, ainsi qu'à tous les participants sans qui cette étude n'aurait pu aboutir.

Toute ma reconnaissance à ma famille, proche et plus éloignée (et particulièrement aux Stéphanois Marion, Raymond, Nicolette et Karima pour leur accueil, leur gentillesse et leur aide précieuse au recrutement de participants). Un grand merci à mes parents pour leur soutien moral et financier durant toute la durée de mes longues études, merci de m'avoir encouragée et fait confiance jusqu'à ce que je trouve ma voie ; merci également à ma sœur, mon frère et mes neveux chéris.

Ces cinq années auraient été bien différentes sans la présence essentielle de mes chères amies futures orthophonistes à qui je souhaite le meilleur pour la suite : Sandrine, Emmanuelle, Isaure, Marie et ma chère colocataire et compagne de galères Alice.

Je me sens très chanceuse d'avoir pu compter toutes ces années sur un entourage sans faille : une pensée à Agathe, Marilou, Jaya, Juliette, Antoinette, Adèle. Un merci tout particulier à ma chère Hannah pour son amitié, son soutien et son aide au recrutement de participants. Un grand merci à celle qui depuis seize ans m'a toujours soutenue, avec qui j'ai tant partagé, et sans qui je n'en serais sans doute pas là, Faustine ; vivement nos futures aventures lilloises ! Difficile d'imaginer cette dernière année sans toi Thomas, merci pour tout...

Enfin, un grand merci à toutes celles et ceux que je n'ai pu citer, mais qui m'ont supportée, soutenue, encouragée tout au long de mes études.

Résumé :

Les tests de dénomination orale actuellement utilisés ne sont pas spécifiques à la chirurgie éveillée, ne disposent pas de normes solides et récentes et n'évaluent que les noms. L'hétérogénéité des outils d'évaluation limite les études à grande échelle. L'objectif de cette étude est de finaliser et de normaliser en français le Multimap, un test de dénomination orale d'images destiné à l'évaluation des noms et des verbes en péri-opératoire de la chirurgie éveillée des gliomes. Le test a été administré à 416 sujets sains, recrutés parmi la population française selon 3 variables démographiques ; leurs performances ont fait l'objet d'analyses qualitatives et quantitatives. Seuls les 80 items obtenant un fort taux de consensus ont été conservés. Les effets des variables démographiques ont été étudiés afin d'étalonner les normes pour chaque épreuve. Une analyse comparative et corrélative des résultats a également été menée. Les résultats mettent en évidence un effet de l'âge pour les deux épreuves, mais aucun effet du sexe, ainsi qu'un effet du niveau d'études uniquement pour l'épreuve de dénomination d'actions. Les résultats concordent en grande partie avec ceux retrouvés dans la littérature. Les normes ont été stratifiées sur trois groupes significatifs pour l'épreuve de dénomination d'objets, et sur six groupes pour l'épreuve de dénomination d'actions. Afin d'être utilisé en clinique, le Multimap doit désormais faire l'objet d'une validation auprès de patients. Le Multimap est un test sensible et pertinent, disposant de normes françaises solides, et dont la portée internationale pourrait en faire un nouvel outil de référence.

Mots-clés :

Chirurgie éveillée – gliomes – évaluation – dénomination orale – normalisation

Abstract :

Currently used oral naming tests are not specific to awake surgery, lack strong and recent norms, and only assess nouns. The heterogeneity of assessment tools limits large-scale studies. The aim of this study is to finalize and normalize in French the Multimap, an oral picture naming test for the assessment of nouns and verbs in peri-operative awake glioma surgery. The test was administered to 416 healthy subjects, recruited in the French population according to 3 demographic variables, and their performances were analyzed qualitatively and quantitatively. The tests were completed by keeping only the 80 items with a high consensus rate. The effects of demographic variables were studied in order to calibrate the norms for each test. A comparative and correlative analysis of the results was also conducted. The results show an effect of age for both tests, but no effect of gender. The level of education has an effect on the performance only for the action naming task. The results are largely consistent with those found in the literature. The norms were stratified on three significant groups for the object naming test and on six groups for the action naming test. In order to be used in the clinic, the Multimap must now be validated with patients. The Multimap is a sensitive and relevant test, with solid French standards, and its international scope could make it a new reference tool.

Keywords :

Awake surgery – gliomas – assessment – picture-naming task – normalization

Table des matières

Introduction	1
Contexte théorique, buts et hypothèses	2
1. L'évaluation du langage en chirurgie éveillée	2
1.1 Principes généraux de la neurochirurgie éveillée	2
1.2 Nécessité d'une évaluation langagière complète en plusieurs étapes	3
1.3 Manque de tests spécifiques, standardisés et validés en français	4
2. La tâche de dénomination orale : un « gold standard » pour le « brain mapping »	5
2.1 Modèles psycholinguistiques sous-tendant l'accès lexical	5
2.2 Facteurs influençant la dénomination orale	6
2.3 Type d'erreurs rencontrées et implications neuroanatomiques.....	6
3. Multiples intérêts de l'épreuve de dénomination orale Multimap	7
3.1 Un test spécifique, standardisé, dont les items ont été contrôlés... ..	7
3.2 ...qui offre une banque d'images en couleurs libres de droits.....	8
3.3 ...et qui permet d'évaluer à la fois les verbes et les noms, conformément à la double dissociation retrouvée dans la littérature	8
4. Buts et hypothèses	8
Méthode.....	9
1. Participants à l'étude	9
1.1 Normalisation du test auprès d'une population contrôle	9
1.2 Critères d'inclusion et d'exclusion	9
1.3 Description de l'échantillon	9
1.4 Protection des données.....	10
2. Matériel	10
2.1 Elaboration de l'épreuve française à partir des données basques	10
2.2 Digitalisation de l'épreuve	10
3. Procédure.....	10
3.1 Modalités des passations.....	10
3.2 Entretien préalable et passation de l'épreuve.....	11
4. Traitement des données	12
4.1 Exportation des données	12
4.2 Analyses préliminaires.....	12
4.3 Finalisation de l'épreuve française du Multimap.....	12
4.4 Analyses statistiques	13

Résultats	13
1. Analyses préliminaires et finalisation de l'épreuve française de dénomination orale Multimap (substantifs et verbes)	13
1.1 Analyse du pourcentage de réussite par item.....	14
1.2 Analyse du temps moyen par item.....	15
1.3 Analyse des réponses alternatives et choix des différentes acceptions.....	16
1.4 Corrélation et différences significatives de performances entre les épreuves	17
2. Analyses de l'effet des variables démographiques sur les performances des sujets témoins	18
2.1 Analyses descriptives et distribution des données	18
2.2 Effets des variables démographiques	18
3. Stratification des normes pour le test de dénomination orale du Multimap version française	19
3.1 Définition des sous-groupes significatifs.....	19
3.2 Etablissement des normes	19
Discussion	20
1. Rappel des principaux objectifs et intérêts de l'étude.....	20
1.1 Objectifs de l'étude et hypothèses	20
1.2 Forces de l'étude	20
2. Finalisation du test du Multimap à partir des résultats de la normalisation	20
2.1 Analyse qualitative des réponses non attendues	21
2.2 Sélection des items composant l'épreuve finale et de leurs acceptions.....	23
2.3 Analyse de la pertinence et de la sensibilité d'une épreuve évaluant à la fois les noms et les verbes	24
3. Analyse des principaux effets des variables démographiques et étalonnage du test.....	25
3.1 Principaux effets des variables démographiques sur les performances des sujets témoins	26
3.2 Etalonnage du test.....	27
4. Limites et directions futures	28
4.1 Limites et pistes d'amélioration du test	28
4.2 Limites de l'étude : biais méthodologiques	28
4.3 Directions futures	29
Conclusion.....	29
Bibliographie.....	31
Liste des annexes.....	35

Introduction

Les gliomes diffus de bas grade sont des tumeurs primitives cérébrales rares, de croissance lente mais constante, touchant en général les patients jeunes (Duffau, 2014). Souvent situées dans les zones fonctionnelles du cerveau contrôlant le langage ou la motricité, elles ont longtemps été considérées comme des tumeurs bénignes inopérables (Mandonnet, 2014). Toutefois, ce dogme a été largement remis en cause : en effet, ces tumeurs infiltrantes de bas grade (grade II) migrent le long des faisceaux de substance blanche et se transforment inéluctablement en tumeurs malignes (de grade III ou IV), mettant en jeu le pronostic vital des patients (Mandonnet et al., 2017). Sans traitement adéquat, la médiane de survie est de l'ordre de six ans (Jakola et al., 2012).

La chirurgie éveillée avec électrostimulation corticale est actuellement le traitement de référence et de première intention des gliomes cérébraux (De Witt Hamer et al., 2012). Cette technique consiste à utiliser un système de stimulations électriques cortico-sous-corticales pour créer des « lésions virtuelles transitoires » (Duffau, 2010) tandis que le patient, éveillé, effectue différentes tâches motrices et/ou langagières sous supervision d'un professionnel spécialisé (orthophoniste ou neuropsychologue). Cela permet d'établir une cartographie dynamique des « structures éloquentes » à préserver afin d'optimiser la balance onco-fonctionnelle, c'est-à-dire de maximiser la résection tumorale sans induire de déficit neurologique permanent et ainsi préserver, voire améliorer la qualité de vie des patients (Lemaitre et al., 2022) et allonger l'espérance de vie.

Une évaluation complète du langage et des autres fonctions cognitives est donc nécessaire avant et après l'opération afin d'identifier d'éventuels troubles tandis qu'une évaluation plus ciblée durant l'opération permet une cartographie fonctionnelle propre à chaque individu (Herbet et al., 2017). Toutefois, il existe entre les centres une grande hétérogénéité dans les tâches proposées ainsi qu'un manque de précision quant à la définition des erreurs provoquées par les stimulations (Mandonnet et al., 2016). En effet, à ce jour, les centres qui pratiquent la chirurgie éveillée ne disposent pas de protocole commun, spécifique et validé qui permette d'une part, d'évaluer finement les troubles du langage chez les patients porteurs de gliomes, et, d'autre part, de favoriser les comparaisons inter-centres (Martín-Monzón et al., 2020; Rofes et al., 2017).

La normalisation en français du Dutch Instrument for Mild Aphasia, (DIMA-fr), batterie hollandaise d'évaluation du langage oral en péri-opératoire de la chirurgie éveillée des gliomes, a fait l'objet d'un mémoire d'orthophonie soutenu en 2020 (Clément & Perez, 2020). Cependant, le DIMA ne comporte pas d'épreuve spécifique de dénomination orale, pourtant essentielle en chirurgie éveillée. En effet, la tâche de dénomination est une tâche unanimement reconnue et utilisée dans tous les centres européens (Rofes et al., 2017) : considérée comme un « gold standard » pour tester les fonctions langagières en temps limité en conditions chirurgicales, elle a en outre l'avantage d'être très sensible (Gobbo et al., 2021). Le test de DO 80 (Deloche & Hannequin, 1997), actuellement largement utilisé, n'est pas spécifique à l'évaluation en chirurgie éveillée ; il ne dispose pas de normes solides et récentes et n'est pas libre de droits.

L'objectif de ce mémoire est de normaliser en français un test de dénomination orale d'images, le MULTIMAP (Gisbert-Muñoz et al., 2021). Ce test offre une banque d'images en couleurs, standardisées, libres de droits, permettant d'évaluer en condition de chirurgie éveillée la dénomination de noms (objets) mais également de verbes (actions), et dont les variables psycholinguistiques ont été contrôlées. Cette étude a pour buts d'obtenir des normes plus précises par tranche d'âge et par niveau socio-culturel sur un échantillon plus large de sujets contrôlés et de définir des temps de réalisation par item et par liste, grâce à une présentation des images et un enregistrement des temps de réponse informatisés.

Contexte théorique, buts et hypothèses

Après avoir redéfini les modalités et intérêts de la chirurgie éveillée et de l'évaluation langagière dans ce contexte, nous nous attacherons à décrire les principes de l'épreuve de dénomination orale, tâche per-opératoire de référence pour le « brain mapping ». Enfin, nous présenterons l'épreuve de dénomination orale Multimap que nous nous proposons de normaliser, et ses multiples intérêts cliniques.

1. L'évaluation du langage en chirurgie éveillée

L'évaluation du langage et des autres fonctions cognitives fait partie intégrante de la chirurgie éveillée, à toutes ses étapes. Malgré l'importance d'une évaluation complète et standardisée, il y a peu de protocoles valides et spécifiques, qui faciliteraient les comparaisons inter-centres. Cela limite la connaissance scientifique quant aux apports désormais unanimement reconnus de cette technique novatrice (Mandonnet et al., 2016).

1.1 Principes généraux de la neurochirurgie éveillée

1.1.1 Indications de la chirurgie éveillée : les gliomes diffus de bas grade

Les gliomes sont les tumeurs primitives les plus fréquentes du cerveau. Avec une croissance diamétrale de l'ordre de 4 mm/an, les gliomes diffus de bas grade mettent en jeu le pronostic fonctionnel puis vital des patients (Mandonnet et al., 2017). Le principal problème réside dans le caractère diffus et infiltrant de ces tumeurs, qui conduit à une absence de délimitation nette entre le processus pathologique néoplasique et le parenchyme cérébral (Duffau, 2018). La croissance très lente de ces tumeurs favorise la réorganisation fonctionnelle des aires cérébrales par des mécanismes de plasticité cérébrale. Cela explique pourquoi elles sont le plus souvent découvertes tardivement, généralement suite à une crise d'épilepsie, après une longue phase asymptomatique (Mandonnet et al., 2017).

1.1.2 La chirurgie en condition éveillée comme traitement de référence

La chirurgie d'exérèse précoce supra-maximale est désormais le traitement de référence des gliomes de bas grade (De Witt Hamer et al., 2012). Elle permet un allongement significatif des médianes de survie : de 6-7 ans lorsqu'une simple biopsie est réalisée, elle augmente à 13-15 ans lorsqu'il y a résection (Capelle et al., 2013, cité par Mandonnet et al., 2017). Elle consiste en la réalisation par le neurochirurgien d'une cartographie fonctionnelle per-opératoire, grâce à

la collaboration du patient éveillé et d'une équipe multidisciplinaire (anesthésiste, orthophoniste...). Les aires pour lesquelles la stimulation électrique provoque une perturbation transitoire du fonctionnement cérébral et un déficit fonctionnel sont dites « éloquentes », et identifiées pour être préservées afin d'éviter toute séquelle post-opératoire. Cette technique permet ainsi d'optimiser la balance onco-fonctionnelle en maximisant l'étendue de l'ablation tumorale tout en minimisant le risque de déficits post-opératoires (Duffau, 2018).

1.1.3 Le « brain mapping » : d'une vision localisationniste à une vision hodotopique du fonctionnement cérébral

Parallèlement aux progrès de l'imagerie fonctionnelle et de la tractographie, la chirurgie éveillée a permis de faire évoluer les connaissances sur le fonctionnement cérébral : loin d'une vision traditionnelle localisationniste, les processus cérébraux seraient supportés par des réseaux cortico-sous-corticaux parallèles largement distribués, capables de se réorganiser et de compenser certaines aires lorsqu'elles sont lésées (Duffau, 2014). La très grande variabilité inter-individuelle, due entre autres aux phénomènes de neuroplasticité, nécessite de prendre en compte non seulement les repères anatomiques, mais également les repères fonctionnels rendus visibles par la cartographie per-opératoire (Duffau, 2018).

1.1.4 Multiples apports cliniques et scientifiques

Ainsi, cette technique qui prévient les complications fonctionnelles per-opératoires permet à la fois d'améliorer les médianes de survie, mais également de préserver et même d'améliorer la qualité de vie de patients jeunes, souvent actifs professionnellement (Duffau, 2018). Elle contribue également au renouvellement des connaissances scientifiques sur la compréhension du fonctionnement cérébral. Le principe de cette technique chirurgicale repose en grande partie sur une évaluation cognitive, et notamment langagière, approfondie.

1.2 Nécessité d'une évaluation langagière complète en plusieurs étapes

1.2.1 Principes généraux de l'évaluation orthophonique

L'évaluation du langage et des autres fonctions cognitives dans le contexte de la chirurgie éveillée doit répondre à plusieurs critères. Elle doit être adaptée à chaque patient et tenir compte des spécificités individuelles (âge, latéralité, activité professionnelle, loisirs, localisation de la tumeur). Elle doit également couvrir l'ensemble des processus cognitifs (Rofes et al., 2017) tout en respectant un nombre limité de tâches d'une durée raisonnable (Rofes & Miceli, 2014) puisqu'il faut prendre en considération la durée d'éveil, habituellement limitée à deux heures (Duffau, 2010), ainsi que la fatigabilité du patient, afin de ne pas induire de faux positifs (erreurs indépendantes de la stimulation).

1.2.2 L'évaluation pré-opératoire

Même lorsque le gliome est découvert fortuitement après une phase dite asymptomatique, les patients présentent très souvent de légers déficits cognitifs (Herbet et al., 2017) qui peuvent être mis en évidence grâce à une évaluation approfondie. Celle-ci permet d'établir un bilan de référence afin d'aider l'équipe chirurgicale à choisir les tâches les plus pertinentes en per-opératoire (Mandonnet & Herbet, 2021). Elle permet également d'écarter les stimuli que le patient ne réussirait pas afin d'éviter tout risque de faux positif lors de la chirurgie

(Lubrano et al., 2012; Rofes et al., 2015). Enfin, cette évaluation sert de base de comparaison avec le bilan post-opératoire à distance.

1.2.3 L'évaluation per-opératoire

Elle permet au neurochirurgien d'effectuer la cartographie fonctionnelle des zones éloquentes. La sélection des tâches est adaptée à chaque patient et décidée en amont de l'opération. Elle peut être constituée de tâches de dénomination orale chronométrée, de langage automatique, de fluences (phonologique et sémantique), d'appariement sémantique, de lecture ainsi que d'une évaluation de l'informativité en langage spontané (Herbet et al., 2017). Le patient est également parfois évalué en double tâche (motrice et langagière).

1.2.4 L'évaluation post-opératoire

Selon les centres, cette évaluation peut s'effectuer en une ou deux étapes : immédiatement en post-opératoire (dans les jours qui suivent l'opération) et à distance de la chirurgie (dans un délai variable de quelques semaines à plusieurs mois) (Herbet et al., 2017). Elle permet d'identifier les éventuels déficits post-opératoires et de cibler la rééducation.

1.3 Manque de tests spécifiques, standardisés et validés en français

1.3.1 Hétérogénéité des évaluations proposées

Malgré l'importance et la spécificité de l'évaluation langagière en chirurgie éveillée, les nombreuses études ayant recensé les protocoles et tâches utilisés dans les différents centres européens font le même constat d'une grande hétérogénéité, rendant ainsi impossibles les revues systématiques ou les méta-analyses (Mandonnet et al., 2016; Martín-Monzón et al., 2020; Rofes et al., 2017). Plus encore, de nombreux centres utilisent un mélange de tests standardisés et de tests « maison », qui manquent souvent d'objectifs clairement définis, de cohérence, de fiabilité ainsi que de données normatives (Rofes et al., 2017).

1.3.2 Absence de protocole commun

Il n'existe pas de protocole d'évaluation inter-centres, ni même de publication concernant un noyau d'outils d'évaluation utilisés de manière routinière dans tous les centres de chirurgie éveillée (Martín-Monzón et al., 2020). La majeure partie des tests utilisés n'ont pas été spécialement conçus pour la chirurgie en condition éveillée, mais servent pour d'autres pathologies, comme les tests d'aphasie post-AVC par exemple (Papatzalas et al., 2021).

1.3.3 La normalisation du DIMA

C'est dans ce contexte que le Dutch Instrument for Mild Aphasia (DIMA) a été adapté en français et standardisé (Satoer et al., 2019). Il a également fait l'objet d'une normalisation auprès de 391 sujets-contrôles recrutés dans la population française (Clément & Perez, 2020). Ce test, fondé sur les épreuves les plus complexes de la batterie d'évaluation du Dutch Linguistic Intraoperative Protocol (DuLIP), évalue la réception et la production du langage oral en sémantique, phonologie et syntaxe grâce à 14 épreuves. Malheureusement, il ne comporte pas de tâche de dénomination orale, épreuve considérée comme le gold standard pour l'évaluation des fonctions langagières en chirurgie éveillée.

2. La tâche de dénomination orale : un « gold standard » pour le « brain mapping »

La tâche de dénomination est « l'épreuve per-opératoire de référence » (Lubrano et al., 2012, p. 99). Utilisée classiquement par les pionniers de la chirurgie éveillée (Rofes et al., 2015) et dans toutes les études ayant fait l'objet de revues systématiques (Martín-Monzón et al., 2020), son rôle essentiel dans l'évaluation neurolinguistique n'est plus à prouver. Après avoir rappelé les modèles psycholinguistiques qui sous-tendent cette épreuve, nous décrirons les facteurs influençant la dénomination qu'il convient de contrôler, puis les types d'erreurs que l'on peut rencontrer et les corrélations anatomo-fonctionnelles sous-jacentes.

2.1 Modèles psycholinguistiques sous-tendant l'accès lexical

2.1.1 Les différents niveaux de la production verbale orale

La majorité des auteurs s'accordent pour distinguer différents sous-systèmes dans le processus de production verbale orale et donc de dénomination orale : un niveau conceptuel, un niveau sémantico-syntaxique (traits sémantiques et propriétés syntaxiques : lemmes), un niveau phonologique (lexèmes) et un niveau articulatoire. Notre étude s'appuiera sur le modèle de Levelt et al. (1999) qui est un modèle sériel par étapes successives : préparation conceptuelle du message (concept lexical) → sélection lexicale (lemme : traits syntaxiques) → encodage morphologique → encodage phonologique → encodage phonétique (patrons gestuels syllabiques) → articulation.

2.1.2 Les modèles de dénomination orale

Les modèles de dénomination reprennent les modèles de production verbale orale, mais proposent en premier lieu une étape d'analyse visuelle. Selon le modèle d'Ellis et al. (1992), la première étape (input) consiste en une analyse perceptive élémentaire de l'objet (différenciation fond-forme, taille, couleur). L'objet est ensuite reconnu parmi les représentations structurales visuelles stockées en mémoire et les informations conceptuelles sont activées dans le système sémantique. Les informations phonologiques du mot sont ensuite activées (lexique phonologique de sortie) et maintenues dans le buffer phonologique (sélection, sériation et planification des phonèmes) jusqu'à la réalisation articulatoire du mot attendu (output).

2.1.3 Proposition d'un modèle hodotopique dynamique grâce aux données issues de la stimulation cérébrale en chirurgie éveillée (Duffau et al., 2014)

Dans un article de 2014, Duffau et al. proposent un nouveau modèle dynamique et hodotopique de la production du langage (voir annexe n°1). Fondé sur le modèle de Levelt et al. (1999), ce nouveau modèle s'oppose à la vision classique localisationniste ainsi qu'au caractère sériel et modulaire du modèle initial. Grâce à une analyse des erreurs produites en dénomination orale suite à la stimulation de certaines zones cérébrales, ils proposent un modèle selon lequel le fonctionnement langagier est organisé en larges réseaux cortico-sous-corticaux parallèles, distribués et interconnectés qui sous-tendent les processus sémantiques, phonologiques et syntaxiques. Ce modèle offre de nombreux avantages : il explique plusieurs doubles dissociations survenues durant la stimulation cérébrale, il tient compte des contraintes

anatomiques cortico-sous-corticales, il explique la récupération possible d'une aphasie suite à une lésion survenue dans les aires classiques du langage, et établit un lien avec les fonctions exécutives (Duffau et al., 2014).

2.2 Facteurs influençant la dénomination orale

Plusieurs facteurs influencent les performances au test de dénomination orale et sont à prendre en considération lors de la construction d'une telle épreuve (Luzzatti et al., 2002).

2.2.1 Effet de fréquence

Les mots de haute fréquence sont dénommés plus rapidement que les mots de basse fréquence.

2.2.2 Effet de familiarité

Il s'agit de la fréquence d'exposition du concept : si le système sémantique est touché, les mots désignant des concepts peu familiers pourraient entraîner plus d'erreurs que les mots désignant des concepts familiers.

2.2.3 Effet de longueur

Les mots longs sont plus difficiles à dénommer que les mots courts, notamment si le buffer phonologique est atteint. Les derniers phonèmes du mot sont ceux qui sont le plus souvent déformés.

2.2.4 Effet de la catégorie sémantique

Certains patients présentent une dissociation avec atteinte sélective de certaines catégories sémantiques (objets manufacturés, parties du corps, vêtements...). Cet effet peut être le signe d'une atteinte des représentations sémantiques ou pré-sémantiques.

2.2.5 Effet de la catégorie grammaticale

Cet effet a été beaucoup étudié dans la littérature et suggère une potentielle dissociation entre les processus de production des noms et des verbes (Rofes et al., 2019) : en effet, la dénomination d'actions est plus lente que la dénomination d'objets (Rofes et al., 2015).

2.2.6 Effet du niveau d'abstraction

Dans la plupart des cas, les mots abstraits sont plus difficiles à dénommer que les mots concrets.

2.3 Type d'erreurs rencontrées et implications neuroanatomiques

L'analyse des erreurs engendrées par la stimulation électrique de certaines aires corticales et sous-corticales peut nous renseigner sur l'organisation fonctionnelle du cerveau (Corina et al., 2010).

2.3.1 Paraphasies sémantiques

Ce type d'erreur consiste à produire un mot non attendu ayant un lien conceptuel avec le mot attendu. Les paraphasies sémantiques peuvent être de plusieurs sortes : paraphasie sémantique catégorielle coordonnée (chat → chien), superordonnée (chat → animal), associative (chat → souris), attributive (chat → gris), subordonnée (main → doigt). Elles peuvent être provoquées par la stimulation du gyrus temporal supérieur, de la partie postérieure

du gyrus temporal moyen, du gyrus supramarginal, du cortex préfrontal dorsolatéral (gyrus frontal moyen), mais aussi du cortex orbitofrontal (pars orbitalis du gyrus frontal inférieur). Au niveau sous-cortical, la stimulation du faisceau occipito-frontal inférieur reliant les régions temporo-pariétales postérieures et le cortex frontal inférieur peut également induire des paraphasies sémantiques (Duffau, 2005).

2.3.2 Anomie

Le patient est capable de parler puisqu'il peut par exemple lire la phrase introductive « C'est... », mais est incapable de produire le mot-cible. Cela s'observe notamment suite à la stimulation de la pars opercularis/triangularis du gyrus frontal inférieur et du gyrus supra-marginal (Lubrano et al., 2012).

2.3.3 Paraphasies phonémiques

Ce sont des transformations en lien avec des perturbations de l'agencement des phonèmes : omissions, ajouts, inversions, déplacements (locomotive → colotomive). Elles s'observent suite à la stimulation du gyrus supra-marginal, de la partie moyenne et postérieure du gyrus temporal supérieur et du cortex frontal inférieur operculaire. Au niveau sous-cortical, les paraphasies phonémiques observées lors de la stimulation du faisceau longitudinal supérieur ou faisceau arqué suggèrent la connectivité entre les régions pariétale inférieure, temporale postéro-supérieure et le cortex frontal inférieur (Lubrano et al., 2012).

2.3.4 Anarthrie / blocage du discours

Le patient est incapable de s'exprimer (perturbations de la programmation motrice des phonèmes). Cela peut s'observer suite à la stimulation du cortex prémoteur ventral, qui joue un rôle dans la syllabation, et de l'insula, qui intervient dans la planification complexe de l'articulation, en lien avec le cortex prémoteur ventral (Lubrano et al., 2012).

3. Multiples intérêts de l'épreuve de dénomination orale Multimap

L'évaluation langagière en chirurgie éveillée est essentielle et s'appuie notamment sur la dénomination orale afin de cartographier les zones cérébrales éloquentes pour le langage. Il n'existe pas de test spécifique, standardisé et validé en français, et l'épreuve de dénomination orale Multimap (Gisbert-Muñoz et al., 2021) répond à ce besoin.

3.1 Un test spécifique, standardisé, dont les items ont été contrôlés...

Selon les recommandations de la littérature, un test utilisé en chirurgie éveillée doit avoir été conçu spécifiquement pour la cartographie fonctionnelle des aires langagières, reposer sur des modèles neuro-psycho-linguistiques actuels, être contrôlé, standardisé, normé sur une population saine, et validé (Papatzalas et al., 2021). Le test de dénomination orale Multimap répond en partie à ces critères : c'est un outil sensible, pour lequel les variables psycholinguistiques des items ont été contrôlées (longueur, fréquence, familiarité, concrétude, imageabilité). Sa composition (noms + verbes) en fait un outil complet pour affiner l'évaluation du langage pendant la phase per-opératoire de la chirurgie éveillée.

3.2 ...qui offre une banque d'images en couleurs libres de droits...

En comparaison des tests de dénomination les plus couramment utilisés, tels que la DO 80 (Metz-Lutz et al., 1991), le Boston Naming Test (Goodglass & Kaplan, 1983) ou la batterie d'images de Snodgrass et Vanderwart (1980), le test de dénomination Multimap offre de nombreux avantages : conçu spécifiquement pour l'évaluation en chirurgie éveillée, il est libre de droits et composé d'items en couleurs. En effet, il a été montré que la colorisation des items permettait de faciliter leur reconnaissance en tâche de dénomination (Rossion & Pourtois, 2004). En outre, il fournit les informations relatives à l'accord sur le nom des items et l'analyse des variables psycholinguistiques, à la différence de certains tests (Gisbert-Muñoz et al., 2021).

3.3 ...et qui permet d'évaluer à la fois les verbes et les noms, conformément à la double dissociation retrouvée dans la littérature

Tous les auteurs s'accordent sur le fait que la production de noms et de verbes implique des systèmes psycholinguistiques et neuroanatomiques au moins partiellement distincts. Cette dissociation a été retrouvée en contexte de chirurgie éveillée lors de tâches de dénomination (Gisbert-Muñoz et al., 2021; Lubrano et al., 2014; Martín-Monzón et al., 2020; Rofes & Miceli, 2014). Alors même que la tâche de dénomination d'objets est systématiquement utilisée, celle-ci n'est peut-être pas la plus sensible pour détecter des déficits langagiers plus subtils et il a été montré que les tâches de dénomination d'actions améliorent la sensibilité, la spécificité et les résultats des procédures de cartographie fonctionnelle (Rofes & Miceli, 2014). Le test de dénomination Multimap offre cet avantage puisqu'il présente à la fois des objets et des actions et permet ainsi une évaluation plus précise.

4. Buts et hypothèses

Le but de cette étude est donc de répondre à un besoin clinique et scientifique, en finalisant un test de dénomination orale conçu spécifiquement pour l'évaluation du langage en péri-opératoire de la chirurgie éveillée des gliomes de bas grade, et en établissant les normes françaises de celui-ci. Les scores des sujets témoins feront l'objet d'une analyse en fonction des variables démographiques susceptibles d'influencer les performances au test. Nous faisons l'hypothèse que l'âge et le niveau socioculturel ont une influence sur les scores, comme cela a été montré dans d'autres normalisations de tests de dénominations (Tran & Godefroy, 2011) : les scores sont susceptibles de diminuer avec l'âge et d'augmenter avec le niveau d'études, tandis que les temps de réponse pourraient augmenter avec l'âge et diminuer avec le niveau d'études. Nous faisons également l'hypothèse que le sexe n'a pas d'effet sur les résultats. En outre, nous faisons l'hypothèse que les performances aux deux épreuves de dénomination (substantifs et verbes) seront corrélées.

Méthode

1. Participants à l'étude

1.1 Normalisation du test auprès d'une population contrôle

Afin d'obtenir des données psychométriques fiables, il est primordial d'étalonner le test auprès d'une population de référence dont les scores servent de normes de comparaison (Ivanova & Hallowell, 2013). Les performances aux tests de langage sont influencées par différents facteurs tels que l'âge, le sexe, le niveau d'éducation ou encore le statut socio-économique (Mitrushina, 2005). Ainsi, les données normatives doivent être le plus représentatives possible afin que les scores des patients évalués puissent être comparés aux scores obtenus par les sujets contrôles (Turkstra et al., 2005).

1.2 Critères d'inclusion et d'exclusion

La population contrôle est composée de volontaires sains, consentants, âgés d'au moins dix-huit ans et de langue maternelle française ou ayant effectué leur scolarité en France. Ils ne présentent pas de trouble cognitif ni de trouble visuel non corrigé. Les sujets qui avaient des antécédents de pathologies cardiovasculaires, cognitives, neurologiques et psychiatriques, de dépendance à un toxique ou à l'alcool ou de troubles de développement du langage ont également été exclus.

1.3 Description de l'échantillon

La normalisation française du Multimap s'est effectuée auprès de 416 sujets contrôles, nombre qui garantit une puissance statistique satisfaisante (Ivanova & Hallowell, 2013). Ils ont été recrutés, parmi la population générale et sur l'ensemble du territoire français, selon trois variables démographiques guidées par la littérature, pouvant avoir une influence sur les performances : le sexe (masculin ; féminin), l'âge (18-29 ans ; 30-49 ans ; 50-69 ans ; 70 ans et plus) et le niveau d'études (sans le baccalauréat ; Bac à Bac+3 ; Bac+3 et plus). Les sujets ont été répartis de manière équitable afin d'analyser l'effet des variables démographiques sur les performances des sujets et permettre la généralisation des résultats. L'échantillon est donc composé de 416 sujets âgés de 18 à 95 ans (moyenne d'âge : 49.7 ans), dont 134 ont un niveau d'études inférieur au Bac (32.2%), 131 ont un niveau d'études compris entre Bac et Bac+3 (31.5%) et 151 ont un niveau d'études supérieur à Bac+3 (36.3%). Les caractéristiques de l'échantillon sont présentées dans le *Tableau 1* ci-dessous.

	Femmes			Hommes			N (%)
	< Bac	> Bac	> Bac+3	< Bac	> Bac	> Bac+3	
18-29 ans	16	18	23	15	16	16	104 (25%)
30-49 ans	17	18	21	14	15	18	103 (24.8%)
50-69 ans	22	22	20	18	11	16	109 (26.2%)
70 ans et +	20	20	21	12	11	16	100 (24%)
N	75	78	85	59	53	66	416
N (%)	238 (57.2%)			178 (42.8%)			

Tableau 1: Description de la population contrôle

1.4 Protection des données

Afin d'être en conformité avec le Règlement sur la Protection des Données (RGPD) et de garantir la protection des données à caractère personnel des sujets, celles-ci ont été anonymisées. Un dossier a été déposé auprès du DPO de l'Université de Lille et une déclaration de conformité (cf. annexe n°2) a été reçue le 17/06/2021 avant le début des passations.

2. Matériel

2.1 Elaboration de l'épreuve française à partir des données basques

Le test que nous présentons a été adapté du test original Multimap, créé par l'équipe basque de Gisbert-Muñoz et al. (2021). Dans leur étude initiale, les auteurs ont mis à disposition une banque de 218 images libres de droits. Plusieurs listes monolingues et bilingues ont fait l'objet d'une validation, dans le but de créer un outil international pour l'évaluation du langage en chirurgie éveillée. Leur protocole comporte cependant quelques limitations : la liste initiale d'items en espagnol a été normalisée auprès d'une population réduite, dont les variables démographiques n'ont pas été véritablement contrôlées. Ceux-ci ont répondu par écrit à un questionnaire en ligne, et l'accord sur la dénomination (supérieur à 80%) a été calculé à partir de ces résultats. La liste française de 80 items a été normalisée de la même façon auprès de 104 sujets (89 femmes et 15 hommes), dont la moyenne d'âge était de 24.29 ans. Au vu de ces limites méthodologiques et afin d'anticiper d'éventuels taux de réussite inférieurs à 80% pour certains items de la liste initiale, il a été décidé d'ajouter 10 items issus de cette même banque d'images à chaque épreuve (substantifs et verbes) pour remplacer les images problématiques le cas échéant.

2.2 Digitalisation de l'épreuve

L'épreuve a ensuite été digitalisée en collaboration avec la plateforme HappyNeuron. Elle peut être présentée sur n'importe quel support informatique (tablette, ordinateur) et propose une interface en trois temps (pré-opératoire, per-opératoire, post-opératoire) en concordance avec les conditions spécifiques de l'évaluation langagière dans le cadre de la chirurgie éveillée. La digitalisation de l'épreuve permet une présentation standardisée du test afin de minimiser les biais liés aux examinateurs, et un enregistrement précis des temps de réponse, par item et total. En effet, la mesure de la vitesse d'accès lexical devrait être systématiquement incluse dans les évaluations langagières, dans la mesure où celle-ci semble corrélée entre autres avec le retour au travail après une chirurgie éveillée (Moritz-Gasser et al., 2012).

3. Procédure

3.1 Modalités des passations

3.1.1 Participants au projet

Afin de maximiser le nombre de passations, six étudiantes de l'école d'orthophonie de Paris ont été recrutées dans le cadre d'un stage recherche pour contribuer à la normalisation du test.

3.1.2 Formation des étudiantes et suivi des passations

Pour garantir une standardisation optimale et un protocole similaire d'un sujet à l'autre et d'un examinateur à l'autre (Ivanova & Hallowell, 2013), les étudiantes ont été formées à la passation des épreuves de screening et du test de dénomination. Plusieurs réunions ont eu lieu entre juin et décembre 2021 pour suivre l'avancement des passations et cibler les éventuels problèmes rencontrés. Les étudiantes ont reçu des instructions spécifiques et concrètes quant à la présentation des items, à la cotation des réponses et au recueil des résultats (observations qualitatives, réponses non attendues). Une cotation stricte a été appliquée : chaque réponse non attendue a été retranscrite dans le logiciel et a fait l'objet d'une analyse détaillée.

3.1.3 Déroulement des passations

Les passations se sont déroulées dans un endroit calme, dépourvu de source de distractions et bien éclairé, entre juin 2021 et janvier 2022. Chaque étudiante a interrogé une soixantaine de sujets, recrutés dans toute la France, en tenant compte pour le recrutement de l'équilibre entre les variables démographiques. Un tableau collaboratif en ligne rempli au fur et à mesure a permis de suivre l'avancement des passations.

3.2 Entretien préalable et passation de l'épreuve

3.2.1 Note d'information et formulaire de consentement

Les participants ont été informés des intérêts et modalités de l'étude via une note d'information explicitée oralement lors de l'entretien préalable et ont signé un formulaire de consentement éclairé (voir annexe n°3).

3.2.2 Questionnaire d'inclusion

Un questionnaire d'inclusion a permis d'écartier les sujets qui ne pouvaient pas participer à l'étude en raison d'antécédents médicaux particuliers (voir annexe n°4).

3.2.3 Tests de screening

Afin de dépister d'éventuels troubles cognitifs légers non connus des participants, le test de screening de la MoCA (Montreal Cognitive Assessment, version française 8.1, Nasreddine et al., 2005) a été administré. En effet, selon la méta-analyse de Ciesielska et ses collègues (2016), celui-ci serait plus sensible et spécifique que le MMSE (Mini-Mental State Examination, 1975) pour détecter un trouble cognitif léger chez les individus âgés de plus de 60 ans. Pour être inclus, les sujets ne devaient pas obtenir un score inférieur à 26/30.

Pour s'assurer de l'acuité visuelle des sujets, l'échelle d'acuité visuelle de près de Parinaud, utilisée en pratique courante par les ophtalmologistes français, a été proposée. Pour être inclus, les sujets ne devaient pas obtenir un score inférieur à P2 (vision normale, avec ou sans correction).

3.2.4 Epreuve de dénomination

L'épreuve de dénomination orale, présentée aux sujets sains sur un support informatisé, est composée d'une liste de 2x50 items (substantifs / verbes) précédés de 2 items d'exemples pour chaque épreuve. La durée de passation de l'épreuve est d'environ dix minutes.



Figure 1: Exemple d'item (nom) → « un poisson »

Pour l'épreuve de dénomination d'objets (cf. *Figure 1* pour un exemple), la consigne est la suivante : « Vous allez voir défiler des objets. Vous devez dire ce que vous voyez en ajoutant un article (la, du, un...). Voici deux exemples pour commencer ».

Pour l'épreuve de dénomination d'actions (cf. *Figure 2* pour un exemple), la consigne est la suivante : « Vous allez voir défiler des actions. Vous devez dire « il, elle, ils... » suivi du verbe illustré par l'image. Voici deux exemples pour commencer ».



Figure 2: Exemple d'item (verbe) → « elle mange »

Si le sujet échoue aux exemples, la réponse attendue est donnée ; si le sujet donne une réponse avec expansion (nom suivi d'un adjectif ou verbe suivi d'un complément), la consigne est rappelée de ne dénommer strictement que l'objet ou l'action présentés, sans complément.

4. Traitement des données

4.1 Exportation des données

Les données anonymisées ont été exportées de la base de données du logiciel Multimap. Celui-ci étant encore en développement, l'exportation de données au format recherche s'est effectuée en différentes phases avec signalement et correction des éventuels bugs à l'équipe de développeurs. Plusieurs modifications et vérifications ont été réalisées avant d'obtenir le tableau de données final : suppression des sujets qui avaient été effacés du logiciel mais pas de la base de données, suppression des sujets pour lesquels il y avait eu des problèmes lors de la passation et pour qui certains items étaient en double ou manquants, vérification des données démographiques des sujets.

4.2 Analyses préliminaires

Afin d'éviter une trop grande variabilité des réponses lors de l'évaluation de patients, il est primordial de procéder à une analyse approfondie des réponses collectées chez les volontaires sains (Rofes & Miceli, 2014). C'est pourquoi une analyse quantitative et qualitative des réponses non attendues a ensuite été réalisée : le taux de réussite par item a été calculé et les trente items (quinze substantifs et quinze verbes) obtenant le taux de réussite le plus bas ont été identifiés pour une analyse plus approfondie. Le tableau résumant le pourcentage de réussite par item au début des analyses préliminaires et avant finalisation de l'épreuve est disponible en annexe n°5. Les items pour lesquels la cotation n'était pas évidente ont fait l'objet d'une décision de groupe. Le temps moyen de réponse a également été calculé pour chaque item.

4.3 Finalisation de l'épreuve française du Multimap

Cette analyse quantitative et qualitative a abouti à l'élimination des 20 items (10 substantifs et 10 verbes) qui n'atteignaient pas le seuil de 80% de réussite. Pour certains items, il a été décidé d'accepter plusieurs réponses alternatives lorsque celles-ci étaient retrouvées très fréquemment dans la population normale, selon des critères explicités ci-après.

4.4 Analyses statistiques

Les données ont été exportées et traitées sur Excel puis les analyses statistiques ont été effectuées à l'aide du logiciel Jamovi 2.2.5.

4.4.1 Finalisation du tableau de données

Après avoir corrigé le tableau en supprimant les items éliminés et en modifiant la cotation des items acceptant plusieurs réponses, les scores et les temps globaux de chaque sujet pour l'épreuve de dénomination de substantifs (score /40) et l'épreuve de dénomination de verbes (score /40) ont été calculés. Un nouveau tableau a été créé afin de visualiser les scores et les temps globaux en fonction des variables démographiques des sujets.

4.4.2 Test de la normalité des données

Le test de Shapiro-Wilk a été réalisé afin de déterminer si les données suivaient une loi normale.

4.4.3 Mesure de la corrélation entre les épreuves

Les corrélations entre les épreuves (pour les scores et les temps) ont été mesurées à l'aide du test non-paramétrique de Kendall.

4.4.4 Analyse de la différence significative de performance entre les épreuves

Afin de déterminer s'il existait une différence significative de performances entre les deux épreuves, le test des rangs signés de Wilcoxon pour échantillons appariés a été réalisé.

4.4.5 Contrôle de l'effet des variables démographiques sur les performances

Le test non-paramétrique de Kruskal-Wallis a été utilisé afin de déterminer l'effet des variables démographiques (sexe, âge, niveau d'étude) sur les performances des sujets. Des analyses post-hoc (comparaisons par paires du test non-paramétrique de Dwass-Steel-Critchlow-Fligner) ont été effectuées afin de déterminer quels sous-groupes obtenaient des performances significativement différentes.

4.4.6 Analyses descriptives et étalonnage

Pour chaque épreuve (substantifs et verbes), la moyenne, l'écart-type, le minimum, le maximum et la répartition des performances par centiles (5, 10, 25, 50, 75, 90, 95) des scores et des temps globaux ont été calculés pour l'ensemble de l'échantillon et par sous-groupe significatif.

Résultats

1. Analyses préliminaires et finalisation de l'épreuve française de dénomination orale Multimap (substantifs et verbes)

1.1 Analyse du pourcentage de réussite par item

1.1.1 Analyses préliminaires

Le tableau qui recense le taux de réussite pour chaque item, en tenant compte d'une cotation stricte et avant modifications, est disponible en annexe n°5 (*Tableau 6*). Les quinze items de chaque épreuve ayant le taux de concordance le plus bas ont fait l'objet d'une analyse approfondie. La cotation ayant été volontairement très stricte pour la normalisation, certains items (comme « ciseaux ») ont un taux de concordance bas, car les réponses non attendues (telles que « une paire de ciseaux » ou « un ciseau »), bien que très répandues en population générale, ont été cotées fausses dans un premier temps.

Une première analyse a donc consisté à calculer le taux de réussite pour tous les items, après analyse qualitative des réponses non attendues et en tenant compte des différentes réponses acceptables (détaillées plus bas).

1.1.2 Identification des items ayant le plus faible taux de concordance après analyses

Le tableau ci-dessous (*Tableau 2*) détaille le temps moyen par item ainsi que le taux de réussite pour chaque item en tenant compte des différentes réponses alternatives acceptées, détaillées ci-après. Les chiffres obtenus ont été arrondis au centième près. Le tableau a été trié dans l'ordre croissant en fonction du taux de réussite, ce qui permet de visualiser facilement les vingt items ayant le taux de réussite le plus bas (surlignés en jaune).

Substantifs	Taux de réussite (%)	Temps moyen par item (s)	Verbes	Taux de réussite (%)	Temps moyen par item (s)
serviette	55,53	3,24	imaginer	1,92	2,484
bière	59,86	2,729	confesser	22,84	4,123
désert	60,34	3,806	accélérer	29,09	8,665
mouche	71,88	2,876	livrer	42,79	4,143
berceau	76,2	2,388	traverser	44,47	3,184
cadeau	78,13	2,462	aider	49,76	3,809
béret	78,37	3,909	transpirer	50,24	4,515
plage	81,97	2,808	fermer	56,25	2,897
lune	82,21	3,103	attendre	59,13	5,102
dos	84,13	2,541	vendre	68,99	6,557
manteau	84,13	2,254	servir	74,04	3,466
fleur	88,22	2,281	gagner	75,72	3,497
beurre	89,9	2,347	surligner	76,68	3,352
bouquet	90,63	2,174	signer	80,53	2,552
cuisine	90,87	2,482	toucher	82,45	3,08
jambe	90,87	2,562	construire	82,93	4,274
épaule	91,59	2,55	creuser	83,17	3,015
glaçon	91,59	2,311	saluer	83,17	3,319
nœud	91,59	1,932	poster	83,41	2,818
oignon	92,31	2,53	étendre	83,65	2,86
tasse	92,79	2,097	ranger	84,62	3,074
bras	93,75	2,313	se moucher	85,82	2,607
queue	94,23	2,507	nettoyer	87,74	2,492
dent	95,43	1,864	chasser	88,7	3,008
cadenas	96,15	2,388	allumer	89,66	3,235
calendrier	96,15	2,261	casser	89,9	3,131
cochon	96,39	1,856	brûler	91,83	2,666
couteau	97,36	1,746	monter	92,31	2,245
corde	97,6	2,162	se marier	93,27	3,164

château	98,08	2,082	écouter	94,95	2,368
cheval	98,32	1,947	peindre	95,19	2,081
chemise	98,32	2,029	peser	95,91	3,48
journal	98,32	2,163	sauter	96,63	2,298
œuf	98,56	1,812	voler	97,12	2,541
serrure	98,56	2,24	prier	97,36	2,257
fauteuil	98,8	1,923	s'asseoir	97,36	2,159
poisson	98,8	2,039	chanter	97,6	2,055
chaise	99,04	1,758	jouer	97,6	2,432
ciseaux	99,04	2,074	regarder	98,32	2,874
écureuil	99,04	2,026	mesurer	98,56	2,417
lapin	99,04	1,872	fumer	98,8	2,107
os	99,28	1,826	découper	99,04	2,539
vélo	99,28	1,847	écrire	99,04	2,265
ceinture	99,52	1,777	pousser	99,04	2,123
timbre	99,52	1,819	boire	99,28	1,925
église	99,76	1,829	conduire	99,52	2,184
feuille	99,76	1,841	courir	99,76	1,96
papillon	99,76	1,801	lire	99,76	2,038
parapluie	99,76	1,89	manger	99,76	2,181
chapeau	100	1,712	pêcher	99,76	2,399

Tableau 2 : Taux de concordance et temps moyen par item de l'épreuve finalisée

1.1.3 Choix des items à supprimer

Pour chaque épreuve (substantifs et verbes), les dix items ayant le taux de concordance le plus bas ont donc été sélectionnés pour être supprimés de l'épreuve finale et du calcul des performances des sujets témoins. Pour les substantifs, deux items (« dos » et « manteau ») obtiennent le même taux de concordance (84,13%) : nous avons choisi de garder l'item « manteau » et de supprimer l'item « dos » pour plusieurs raisons détaillées dans la discussion.

1.1.4 Résultats du taux de concordance de l'épreuve finale

L'épreuve finale contient ainsi 40 substantifs et 40 verbes. Le taux de concordance moyen de l'épreuve de dénomination d'objets est de 96,1% (avec un écart-type de 3,99%) et celui de l'épreuve de dénomination d'actions est de 91,7% (avec un écart-type de 7,82%). Seuls trois items de l'épreuve de dénomination d'actions n'atteignent pas le seuil de 80% de réussite fixé au préalable, mais atteignent au minimum 74%.

1.2 Analyse du temps moyen par item

1.2.1 Résultats du temps moyen par item de l'épreuve finale

Le temps moyen pour chacun des items de l'épreuve finale et pour les items supprimés est détaillé dans le *Tableau 2* ci-dessus.

Le temps moyen de réponse pour les substantifs est de 2,07 s (avec un écart-type de 257 ms) et celui des verbes est de 2,66 s (avec un écart-type de 541 ms).

1.2.2 Analyse de la corrélation entre le pourcentage de réussite par item et le temps moyen de réponse

Afin d'observer si les deux variables (temps moyen et taux de réussite par item) étaient corrélées, le coefficient de Kendall, qui s'applique pour les valeurs non-paramétriques, a été calculé. Une corrélation significative négative est retrouvée dans les deux épreuves, ce qui signifie qu'en moyenne, plus un item est réussi, plus il est dénommé rapidement.

Pour les substantifs, le Tau de Kendall est égal à $-0,568$ ($p < 0.001$). Pour les verbes, le Tau de Kendall est égal à $-0,574$ ($p < 0.001$). La représentation graphique des droites (Figures 2 et 3) permet de rendre compte de cette forte association. Le coefficient de Kendall est encore plus élevé si l'on inclut les substantifs supprimés (Tau de Kendall = $-0,672$; $p < 0.001$) et les verbes supprimés (Tau de Kendall = $-0,608$; $p < 0.001$).

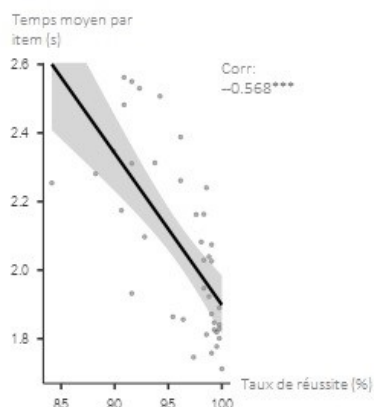


Figure 3 : Corrélation (Kendall) entre le taux de réussite et le temps moyen par item pour l'épreuve des substantifs

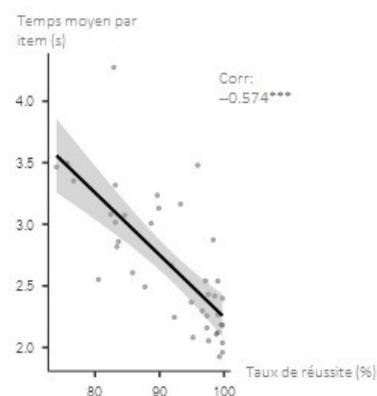


Figure 4 : Corrélation (Kendall) entre le taux de réussite et le temps moyen par item pour l'épreuve des verbes

1.3 Analyse des réponses alternatives et choix des différentes acceptations

Une analyse qualitative et quantitative des réponses non attendues a été réalisée. Lorsqu'une réponse non attendue était retrouvée très fréquemment dans la population normale, celle-ci a fait l'objet d'une décision de groupe pour déterminer si elle pouvait faire partie des réponses acceptables de l'item.

1.3.1 Pourcentages des réponses alternatives pour les substantifs

Le tableau ci-dessous (Tableau 3) recense les pourcentages de réponses alternatives pour huit des quarante substantifs de l'épreuve finale.

Items	% strict	Réponse alternative	% réponse alternative	Acceptations	% de réussite
bouquet	60,58	bouquet de fleurs	30,05	bouquet, bouquet de fleurs	90,63
château	79,33	château fort	18,75	château, château fort	98,08
ciseaux	70,67	un / une paire de...	28,37	un / une paire de / des ciseaux	99,04
queue	85,98	queue de chat	8,25	queue, queue de chat	94,23
serrure	94,23	trou de serrure	4,33	serrure, trou de serrure	98,56
tasse	80,77	mug	12,02	tasse, mug	92,79
timbre	98,32	timbre-poste	1,2	timbre, timbre-poste	99,52
vélo	85,34	bicyclette	13,94	vélo, bicyclette	99,28

Tableau 3 : Pourcentage des réponses alternatives acceptées pour les substantifs

1.3.2 Pourcentages des réponses alternatives pour les verbes

Le tableau ci-dessous (Tableau 4) recense les pourcentages de réponses alternatives pour sept des quarante verbes de l'épreuve finale.

Items	% strict	Réponses alternatives	% réponses alternatives	Acceptations	% de réussite
regarder	88,7	il/elle regarde + la tv	9,62	regarder, regarder la tv	98,32

étendre	80,53	il/elle étend, tend + le linge	3,12	étendre, tendre le linge	83,65
surligner	70,19	il/elle stabilote, stabilobosse	6,49	surligner, stabiloter, stabilobosser	76,68
découper	65,38	il/elle coupe	33,66	découper, couper	99,04
servir	61,06	il/elle sert + COD, verse + COD	12,98	servir, verser + COD	74,04
se moucher	83,65	il/elle mouche	2,17	se moucher, moucher	85,82
gagner	73,56	il/elle gagne, a gagné + COD	2,16	gagner, avoir gagné + COD	75,72

Tableau 4 : Pourcentage des réponses alternatives acceptées pour les verbes

1.4 Corrélation et différences significatives de performances entre les épreuves

1.4.1 Mesure de la corrélation entre les épreuves

Les mesures de corrélation ont été effectuées grâce au calcul du coefficient Tau de Kendall. Les scores et les temps aux épreuves de dénomination d'objets et d'actions sont tous significativement ($p < 0.001$) corrélés entre eux (détail des coefficients de corrélation dans le *Tableau 7* en annexe n°6). Une corrélation forte (Tau de Kendall = 0.586) est retrouvée entre les temps de réponse pour la dénomination d'objets et d'actions et une corrélation modérée (comprise entre $-0,1$ et $-0,5$ ou entre $0,1$ et $0,5$) est retrouvée entre toutes les autres performances.

1.4.2 Comparaison des performances des sujets sains (scores et temps) entre les deux épreuves

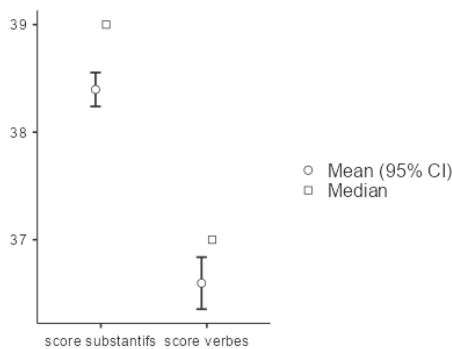


Figure 5 : Ecart entre la distribution moyenne des scores aux deux épreuves

Le test des rangs signés de Wilcoxon a été effectué pour observer si les scores obtenus par les sujets témoins aux deux épreuves étaient significativement différents.

Une différence significative est retrouvée ($p < 0.001$) : la moyenne du score à l'épreuve de dénomination d'objets est supérieure de 2 points à la moyenne du score à l'épreuve de dénomination d'actions.

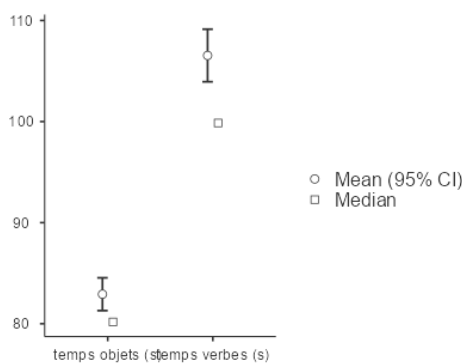


Figure 6 : Ecart entre la distribution moyenne des temps de réponse aux deux épreuves

Le même test a été effectué pour comparer les temps de réponse des sujets aux deux épreuves.

Une différence significative ($p < 0.001$) est retrouvée là encore : le temps de réponse à l'épreuve de dénomination d'objets est en moyenne plus court de 21,65 secondes que le temps de réponse à l'épreuve de dénomination d'actions.

Les objets sont en moyenne dénommés plus rapidement et avec moins d'erreurs que les actions.

2. Analyses de l'effet des variables démographiques sur les performances des sujets témoins

2.1 Analyses descriptives et distribution des données

Le tableau en annexe n°7 (*Tableau 8*) résume les analyses descriptives des scores et temps globaux pour l'épreuve de dénomination d'objets et d'actions, ainsi que le résultat du test de Shapiro-Wilk appliqué aux données afin de tester leur normalité. Pour rappel, si la p-valeur du Shapiro-Wilk est inférieure à 0.05, alors l'hypothèse nulle selon laquelle les données sont normalement distribuées est rejetée. La p-valeur est ici inférieure à 0.001 pour toutes les données, ce qui indique qu'elles ne sont pas gaussiennes. Cela implique qu'il faudra utiliser des tests non-paramétriques pour les analyses statistiques. En outre, les percentiles, non dépendants de la distribution des données, seront à privilégier pour l'établissement des normes (Aguert & Capel, 2018).

2.2 Effets des variables démographiques

Les p-valeurs obtenues aux tests statistiques réalisés sont résumées en annexe n°7.

2.2.1 Effet du sexe

Le sexe n'a pas d'effet significatif sur les performances des sujets témoins, tant pour les scores aux épreuves de dénomination de substantifs et de verbes que pour les temps de réponse. Les p-valeurs sont toutes supérieures à 0.05 et peuvent être retrouvées en annexe n°7.

2.2.2 Effet de l'âge

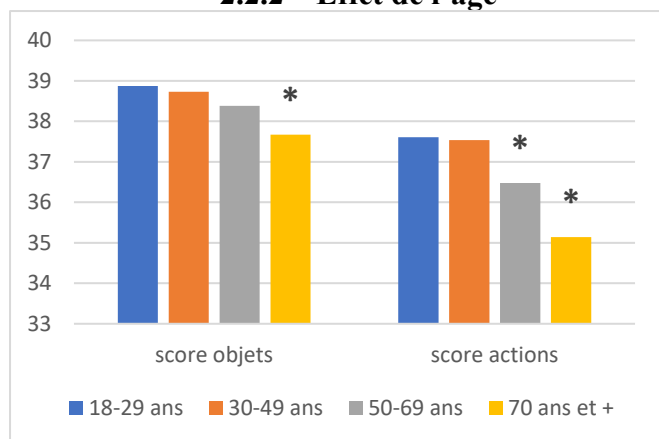


Figure 7 : Effet de l'âge sur les scores (* $p < 0.05$)

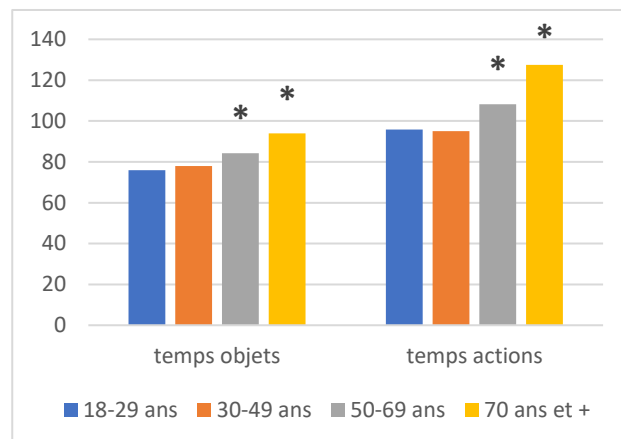


Figure 8 : Effet de l'âge sur les temps de réponse (* $p < 0.05$)

L'âge a un effet significatif sur les scores et les temps de réponse des sujets normaux pour les deux épreuves ($p < 0.001$). En revanche, une différence significative n'est pas retrouvée entre toutes les tranches d'âge : en effet, il n'existe pas de différence entre les sujets âgés de 18 à 29 ans et les sujets âgés de 30 à 49 ans dans aucune des épreuves. De plus, aucune différence significative n'est retrouvée entre les sujets âgés de 18 à 49 ans et les sujets âgés de 50 à 69 ans pour le score à l'épreuve des substantifs, mais le temps de réponse est significativement différent.

2.2.3 Effet du niveau d'études

Le niveau d'études a un effet significatif sur les performances et les temps de réponse uniquement pour l'épreuve de dénomination d'actions. On ne retrouve pas d'effet du niveau d'études pour l'épreuve de dénomination d'objets. En outre, il n'y a qu'un groupe sur les trois qui obtient des performances significativement différentes : pour l'épreuve de dénomination de verbes, les sujets qui n'ont pas le baccalauréat (Bac) ont en moyenne des scores plus faibles et des temps de réponse plus longs que les sujets qui ont le Bac ou plus, mais aucune différence significative n'est relevée entre les sujets qui ont le Bac et les sujets qui ont un niveau supérieur à Bac+3.

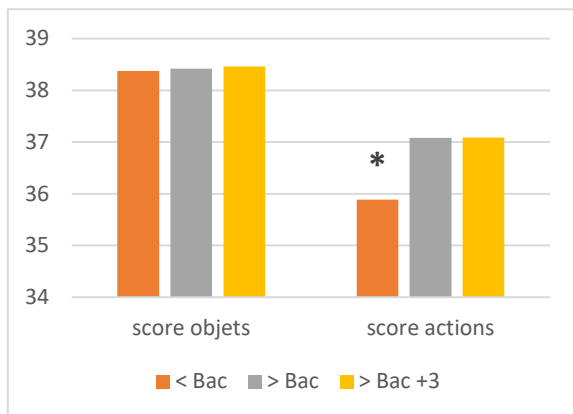


Figure 9 : Effet du niveau d'études sur les scores (* $p < 0.05$)

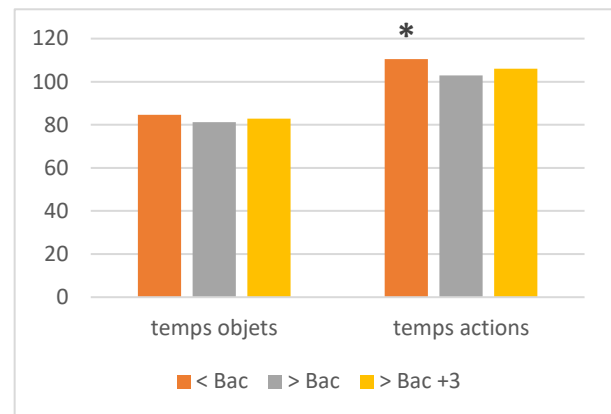


Figure 10 : Effet du niveau d'études sur les temps de réponse (* $p < 0.05$)

3. Stratification des normes pour le test de dénomination orale du Multimap version française

3.1 Définition des sous-groupes significatifs

L'analyse des effets des variables démographiques sur les performances des sujets a permis de définir les sous-groupes pertinents pour la stratification des normes : trois tranches d'âge (18-49 ans, 50-69 ans, 70 ans et plus) ainsi que deux niveaux d'étude (niveau inférieur au baccalauréat, niveau supérieur au baccalauréat) sont ainsi retenus. Le sexe n'ayant pas d'effet significatif, il n'est pas retenu comme un critère pertinent.

3.2 Etablissement des normes

Les normes ont été stratifiées en six groupes pour l'épreuve de dénomination d'actions : 18-49 ans sans le Bac ($n = 62$) ; 18-49 ans avec le Bac ($n = 145$) ; 50-69 ans sans le Bac ($n = 40$) ; 50-69 ans avec le Bac ($n = 69$) ; 70 ans et plus sans le Bac ($n = 32$) ; 70 ans et plus avec le Bac ($n = 68$). Aucun effet du niveau d'études n'ayant été relevé pour l'épreuve de dénomination d'objets, les normes ont été stratifiées en 3 groupes pour cette épreuve : 18-49 ans ($n = 207$) ; 50-69 ans ($n = 109$) ; 70 ans et plus ($n = 100$). Les données normatives de l'ensemble de l'échantillon ($n = 416$) figurent en annexe n°8. La moyenne, l'écart-type, le minimum, le maximum et la répartition des performances par centiles (5, 10, 25, 50, 75, 90, 95) sont indiqués par groupe significatif, pour le score et le temps de chaque épreuve (dénomination d'objets et dénomination d'actions). Par convention, le percentile 5 correspond aux moins bonnes performances pour les scores et les temps de réponse.

Discussion

1. Rappel des principaux objectifs et intérêts de l'étude

1.1 Objectifs de l'étude et hypothèses

Cette étude avait pour objectifs la finalisation et la normalisation du Multimap, un test de dénomination orale d'objets et d'actions (substantifs et verbes), spécifiquement conçu pour l'évaluation du langage en péri-opératoire de la chirurgie éveillée des gliomes. Ces deux épreuves pourraient à terme être incluses dans un protocole commun international d'évaluation du langage, qui manque cruellement actuellement, et favoriser ainsi les comparaisons inter-centres.

Les principaux objectifs devaient être conduits de la manière suivante : d'une part, l'analyse des résultats de la normalisation devait permettre de finaliser l'épreuve en ne conservant que les items ayant un fort taux de réussite auprès de la population contrôle ; d'autre part, l'analyse des effets des variables démographiques (sexe, âge et niveau d'études) sur les performances des sujets devait conduire à l'étalonnage des normes pour les deux épreuves.

Nous avons formulé les hypothèses suivantes : le sexe n'a pas d'effet sur les performances en dénomination des sujets, tandis qu'un effet des variables âge et niveau d'études est attendu ; enfin, les deux épreuves sont corrélées entre elles.

1.2 Forces de l'étude

Ce travail repose sur l'analyse des performances de 416 sujets dans le but d'étalonner

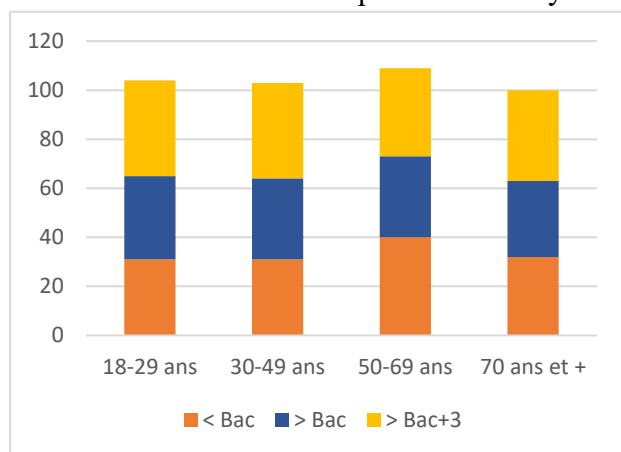


Figure 11 : Répartition de la population contrôle en fonction de l'âge et du niveau d'études

les normes pour les scores et les temps de réponse de chaque épreuve (dénomination d'objets et dénomination d'actions). Sept expérimentatrices ont recruté 238 femmes et 178 hommes, âgés de 18 à 95 ans, équitablement répartis entre quatre tranches d'âge et trois niveaux d'étude. Malgré une très légère surreprésentation de femmes, la population témoin qui compose notre étude est équilibrée pour les variables démographiques comme le souligne la *Figure 11* ci-joint, ce qui garantit la puissance statistique de nos résultats et la représentativité de notre échantillon.

2. Finalisation du test du Multimap à partir des résultats de la normalisation

Le premier objectif de notre étude était de finaliser les épreuves de dénomination orale à partir de l'analyse des résultats de la normalisation. Nous détaillons ici l'analyse qualitative réalisée, les critères de sélection des items ainsi que l'analyse de la pertinence du test finalisé.

2.1 Analyse qualitative des réponses non attendues

Une analyse qualitative des réponses non attendues a été réalisée. Cela présente un double intérêt pour notre étude : d'une part, cette analyse a permis la sélection des items de l'épreuve finale ; d'autre part, il pourrait être intéressant de constituer un inventaire des réponses non attendues chez les sujets sains, afin d'éventuellement comparer en clinique si les erreurs des sujets cérébro-lésés correspondent à celles-ci (Metz-Lutz et al., 1991). Les réponses non attendues peuvent être classées en plusieurs catégories. Nous en présentons quelques exemples de manière non exhaustive. Tous les pourcentages donnés sont arrondis à l'unité la plus proche.

2.1.1 Erreurs dues au phénomène du « mot sur le bout de la langue »

Plusieurs formules d'échec (« je ne sais plus », « ah comment ça s'appelle... ») ont été retrouvées parmi les réponses non attendues de la population contrôle. Les hyperonymes ou les périphrases explicatives retrouvés (« l'emplacement d'une clé » pour « serrure ») pourraient aussi être analysés comme des moyens de compenser un « mot sur le bout de la langue ». Ces erreurs restent cependant très minoritaires, ce qui est attendu auprès d'une population contrôle. Leur fréquence n'a pas été calculée, mais il serait intéressant d'analyser statistiquement s'il y a un effet de l'âge pour ce type d'erreurs, comme retrouvé dans la littérature. En effet, le manque du mot fait partie du langage ordinaire et il a tendance à augmenter avec le vieillissement normal, probablement du fait d'un affaiblissement des connexions du système sémantique avec le lexique phonologique de sortie (Burke & Shafto, 2004).

2.1.2 Erreurs sémantiques

Plusieurs erreurs sémantiques sont retrouvées parmi les réponses non attendues données par les sujets sains.

Certaines sont des hyponymes de l'item-cible : « molaire », « canine » ou « pré-molaire » (retrouvés à hauteur de 4%) pour « dent », « fémur » (1%) pour « os » ou encore « marguerite » (10%) pour « fleur ». Ces erreurs peuvent être liées à une volonté de précision de la part de sujets. Il faut également rappeler que toute situation d'évaluation n'est pas anodine : elle peut avoir un impact sur les phénomènes psychologiques du sujet testé (anxiété de performance par exemple) et donc sur les résultats à l'évaluation (Grégoire, 2004). Pour l'item « fleur », on peut également faire l'hypothèse que la réponse « marguerite », retrouvée assez fréquemment, est liée au test de screening de la MoCA proposé à tous les sujets avant l'épreuve de dénomination. En effet, la MoCA contient une épreuve de rétention de cinq mots, parmi lesquels figure le mot « marguerite ». Cet effet de contamination pourrait expliquer la fréquence importante de cet hyponyme auprès des sujets contrôles. Cette erreur sera très probablement retrouvée moins fréquemment lorsque le test sera utilisé auprès de patients.

Des hyperonymes de l'item-cible sont également produits : « édifice moyen-âgeux » pour « château », « une partie du corps » pour « épaule », « vêtement » pour « manteau », « insecte » pour « mouche ». Ces erreurs, plutôt rares, peuvent être aussi considérées comme des manifestations d'un « mot sur le bout de la langue » (cf. ci-dessus).

Des synonymes de l'item-cible sont retrouvés aussi de manière assez fréquente : « bicyclette » (14%) pour « vélo », « mug » (12%) pour « tasse », « laver » (3%) pour « nettoyer », « envoyer / expédier » (10%) pour poster, « suer » (16%) pour « transpirer ».

Lorsque le synonyme proche était retrouvé de manière très fréquente, celui-ci a fait l'objet d'une décision de groupe pour statuer s'il pouvait faire partie des acceptions de l'item-cible.

Enfin, les erreurs sémantiques peuvent être des co-hyponymes de l'item-cible : ils résultent le plus souvent d'un défaut d'analyse de l'image (cf. ci-après).

2.1.3 Erreurs liées à l'image

Les erreurs peuvent être liées à un défaut de reconnaissance de l'item-cible : pour l'item « serviette » qui n'a obtenu que 55,53% de réussite, les réponses majoritairement retrouvées sont « couverture », « drap », « portefeuille » ou encore « classeur ». L'item « béret » induit des réponses également très extravagantes : de « caillou » à « étui à lunettes » en passant par « foie de veau », « poubelle » ou encore « aubergine pourrie ». L'item « mouche » est reconnu par un quart des sujets comme une « guêpe » ou une « abeille ». Pour l'item « dos », une partie non négligeable de sujets (8%) répond « buste » ou « torse ».

Les erreurs peuvent également résulter d'un défaut d'analyse du fait que les images présentent trop d'éléments. Les sujets ne prennent peut-être pas le temps d'analyser ces éléments du fait de la rapidité de l'épreuve de dénomination. C'est le cas de l'image de l'item « fermer » qui n'est pas claire : beaucoup de sujets se fient au dessin et ne voient le sens de la flèche qu'après, c'est pourquoi on retrouve chez 41% d'entre eux le terme « ouvrir ». L'image de l'item « traverser » est également problématique et 14% des sujets identifient les bandes du passage piéton comme des escaliers. Pour l'item « accélérer » dont le temps moyen de réponse est particulièrement long (8,665 secondes), les erreurs sont à mettre en relation avec l'ambiguïté des nombreux éléments présents sur l'image : on retrouve alors des réponses comme « il roule », « il chronomètre » ou encore « une course de voitures ».

Sont présentes également des erreurs liées à une mauvaise interprétation de l'image. Les sujets ne se focalisent que sur un détail ou font des erreurs de type partie/tout : « parasol » ou « mer » pour l'item « plage », « dunes » (25%) pour « désert », « pichet, bock / pinte / verre de bière » (39%) pour « bière ». Dans l'épreuve de dénomination d'actions, plusieurs images sont ambiguës : de nombreux sujets répondent « il dort » ou « il regarde l'heure » pour l'item « attendre », ou encore « il grimpe » ou « il monte » pour « aider ».

Enfin, les erreurs peuvent résulter d'un défaut d'analyse de type agent/patient, notamment pour les verbes. En effet, certaines images présentent une flèche pour indiquer qui fait l'action, mais cette flèche n'est pas toujours bien analysée par les sujets de la population : pour l'item « confesser », les réponses « il écoute » (44%) ou « il se confesse » (21%) sont très majoritaires. Parfois encore, les images sont ambiguës et ne présentent pas de flèche pour indiquer qui est le protagoniste dont on doit dénommer l'action : c'est le cas pour l'item « livrer » où plus d'un quart des sujets répondent « elle reçoit / elle est livrée / elle réceptionne », ou l'item « servir » pour lequel un cinquième des réponses est centré sur le client qui se fait servir et non sur le serveur.

2.1.4 Erreurs liées aux biais culturels ou aux variantes régionales

Le test de dénomination ayant été construit à partir de l'épreuve conçue par une équipe basque en langue espagnole, plusieurs erreurs résultent de ce biais culturel. Les items « béret » et « confesser » ont un taux de réussite faible : cela est probablement lié à leur plus faible fréquence en France.

Puisque la normalisation s'est effectuée sur une population originaire de toute la France, certaines réponses non attendues sont liées à des variantes régionales : pour l'item « étendre », des réponses comme « tendre » ou « pendre » sont retrouvées ; pour l'item « se moucher », la réponse « moucher » ; pour l'item « glaçon », la réponse « cube de glace ».

2.1.5 Erreurs liées à la non-intégration de la consigne

Enfin, de multiples erreurs sont dues à la non-intégration de la consigne. En effet, certains sujets ajoutent des compléments ou des qualificatifs aux substantifs et aux verbes : « un bras droit », un « beau cheval », « une cuisine aménagée », « elle range un livre », « il pousse la voiture », « il boit de l'eau », etc. Ces réponses ont été cotées fausses, excepté lorsqu'elles étaient retrouvées très fréquemment (comme détaillé dans la partie Résultats) : elles ont alors fait l'objet d'une décision de groupe.

2.2 Sélection des items composant l'épreuve finale et de leurs acceptions

Plusieurs critères ont guidé nos choix lors de la finalisation de l'épreuve, que nous allons discuter ici : élimination des items dont l'image pouvait prêter à confusion, prise en compte du taux de réussite et de la fréquence des réponses non attendues, prise en compte des termes proches ou synonymes fréquents, prise en compte du temps moyen par item.

2.2.1 Justification des réponses alternatives acceptées

Le détail des réponses alternatives acceptées est disponible dans la partie Résultats.

Pour l'épreuve de dénomination d'objets, les réponses alternatives acceptables par rapport à l'image, présentes chez plus de 8% des sujets, ont été retenues pour faire partie des acceptions : compléments du nom lorsque ceux-ci étaient évidents par rapport à l'image (« bouquet de fleurs », « queue de chat »), synonymes ou termes proches (« bicyclette » pour « vélo », « mug » pour « tasse »), précisions d'un terme (« château fort » pour « château »). Deux réponses alternatives (« timbre-poste » et « trou de serrure ») sont beaucoup moins fréquentes, mais ont été retenues également, car il paraissait difficile de les coter fausses. En effet, « timbre » est une abréviation usuelle du terme « timbre-poste » qui le précise ; « trou de serrure » est également un terme usuel, équivalent à « serrure ».

Pour l'épreuve de dénomination d'actions, ont été acceptées les variantes régionales (étendre / tendre ; se moucher / moucher) ou culturelles (surligner / stabiloter) ainsi que les compléments d'objet lorsque ceux-ci étaient retrouvés fréquemment ou que l'image du verbe appelait naturellement un complément (« regarder la télévision », « tendre le linge », « gagner la course »). Pour l'item « servir », il a été décidé d'accepter le terme proche « verser » retrouvé parmi plus de 13% des réponses ainsi que le complément naturel (« du vin » ou « à boire »). Pour l'item « gagner » a été acceptée la forme du verbe au passé composé (« il a gagné »).

2.2.2 Choix des items de l'épreuve finale

Après l'analyse qualitative détaillée précédemment, les dix items obtenant le plus faible taux de réussite ont été éliminés pour chaque épreuve.

Pour l'épreuve de dénomination d'objets, les taux de réussite du dixième (« dos ») et du onzième item (« manteau ») sont identiques (84,13%). Nous avons décidé d'éliminer l'item « dos » pour plusieurs raisons. D'une part, l'image de l'item « dos » est moins claire et donne

lieu à plus d'erreurs d'identification (« des muscles », « un dos musclé », « un torse », « un buste ») tandis que les erreurs pour l'item « manteau » sont des erreurs plutôt sémantiques (cohyponyme « veste »). Cette difficulté d'analyse de l'image peut être mise en évidence également par le temps de réponse qui est plus long en moyenne de 291 millisecondes pour cet item. D'autre part, l'épreuve de substantifs contient déjà plusieurs items qui concernent les parties du corps humain (« épaule », « bras », « jambe »).

Pour l'épreuve de dénomination d'actions, trois items (« servir », « gagner », « surligner ») n'atteignent pas le seuil de 80% de réussite. Après réflexion, il a été décidé de les conserver tout de même afin d'avoir deux épreuves équivalentes en termes d'items, et avoir un nombre d'items suffisant pour chaque épreuve.

2.3 Analyse de la pertinence et de la sensibilité d'une épreuve évaluant à la fois les noms et les verbes

2.3.1 Analyse de la distribution des données de la population contrôle

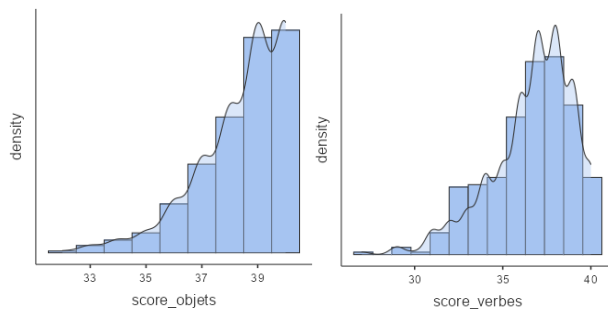


Figure 8 : Distribution des données pour les scores de l'épreuve de dénomination d'objets et de dénomination d'actions

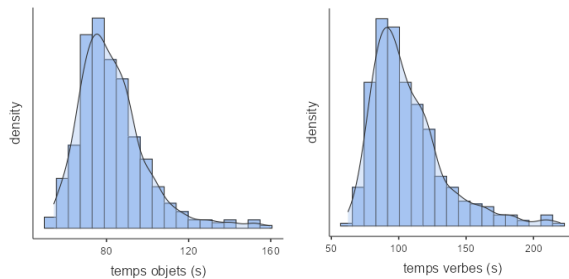


Figure 9 : Distribution des données pour les temps de réponse de l'épreuve de dénomination d'objets et de dénomination d'actions

Comme attendu dans une épreuve de dénomination orale et après sélection des items obtenant un fort taux de consensus, la distribution des scores de la population témoin est marquée par un effet plafond dans chacune des deux épreuves, même s'il est moins marqué pour l'épreuve de dénomination d'actions.

La distribution des temps de réponse n'est pas non plus gaussienne : la courbe est, dans les deux cas, décalée vers la gauche avec davantage de temps de réponse plus longs par rapport à la moyenne que de temps plus courts. Cet effet plancher est retrouvé dans plusieurs autres tests mesurant des temps de réponse et s'explique par le fait qu'il y a un temps minimal pour effectuer une tâche, qui ne peut être réduit, même par les sujets les

plus rapides (Aguert & Capel, 2018). Pour certains tests, cette distribution des temps se normalise si les performances des sujets âgés sont écartées des analyses (moins de valeurs extrêmes).

2.3.2 Corrélation entre les résultats aux épreuves de dénomination d'objets et de dénomination d'actions

Des corrélations significatives modérées ont été observées entre toutes les performances aux épreuves (scores et temps), excepté pour la comparaison entre les temps de réponse à l'épreuve de dénomination d'objets et à celle de dénomination d'actions où une corrélation forte a été observée. Ces corrélations sont positives pour les comparaisons entre les scores, c'est-à-

dire qu'un sujet qui a un score élevé à l'épreuve de dénomination d'objets est susceptible d'avoir un score élevé à l'épreuve de dénomination d'actions. C'est le cas également pour la comparaison entre les temps de réponse aux deux épreuves : un sujet qui dénomme vite les objets est susceptible de dénommer vite les actions. Les corrélations sont en revanche négatives pour les comparaisons entre scores et temps, c'est-à-dire qu'un sujet qui a un score élevé à l'épreuve de substantifs est aussi susceptible de les dénommer plus rapidement.

Ces résultats, qui valident notre hypothèse de départ, sont assez attendus auprès d'une population contrôle puisque les sujets n'ont a priori pas de déficit spécifique d'accès lexical aux noms ou aux verbes. Ils montrent que les deux épreuves testent toutes deux le même type de compétence.

2.3.3 Comparaison des performances des sujets aux deux épreuves

En revanche, nous relevons des différences significatives de performances entre les deux épreuves. En effet, comparée à l'épreuve de dénomination d'actions, la moyenne au score de l'épreuve de dénomination d'objets est supérieure de 2 points et son temps moyen global est moins élevé de 21,65 secondes. La différence entre la production de noms et de verbes a été de nombreuses fois commentée (Bogka et al., 2003; Luzzatti et al., 2002; Rofes, 2015; Rofes & Miceli, 2014). En effet, plusieurs hypothèses sont avancées pour expliquer cette dissociation (Bogka et al., 2003). D'une part, leur organisation lexico-sémantique ne serait pas identique puisque les noms seraient organisés de manière hiérarchique (avec un niveau de subordination entre les catégories : animaux > mammifères > félins > chat) tandis que les verbes auraient plutôt une organisation matricielle (classés en champs sémantiques tels que les verbes de déplacement, de possession, de vision, classés à leur tour sur la base d'autres éléments tels que la direction, la présence ou l'absence d'un instrument, etc.). D'autre part, les verbes sont par nature plus complexes sur le plan grammatical (flexions, transitivité). Enfin, les verbes sont globalement plus abstraits et moins imageables que les noms (Rofes & Miceli, 2014). La production de verbes requerrait ainsi plus de ressources que la production de noms. Il est donc attendu que la dénomination d'actions soit moins réussie et plus lente que la dénomination d'objets concrets. Cette différence est retrouvée dans l'étude de Luzzatti et al. (2002) où les sujets contrôles ont un score à la dénomination de verbes significativement inférieur à la dénomination de noms. En ce qui concerne les temps de réponse, les trois études de Bogka et al. (2003) ont montré qu'il existait également une différence significative entre les noms et les verbes. Cependant, cette différence n'était plus significative dès lors que l'imageabilité et la complexité visuelle des images des verbes étaient contrôlées. Même si ces résultats sont à relativiser du fait de la faible taille des échantillons dans les études précitées, cette hypothèse ne peut pas être écartée. Il serait donc intéressant de contrôler ces variables pour les items de l'épreuve de dénomination d'actions. En outre, comme trois des items n'obtiennent pas 80% de réussite, on ne peut écarter une différence liée à la qualité des images pour cette épreuve. Cependant, cette dissociation confirme l'intérêt d'une épreuve qui évalue à la fois les noms et les verbes, en faveur d'une plus grande sensibilité du test et qui permettra en clinique de mettre en évidence des déficits spécifiques.

3. Analyse des principaux effets des variables démographiques et étalonnage du test

Le deuxième objectif de notre étude était d'analyser les effets des variables démographiques sur les performances des sujets afin d'établir les normes des deux épreuves pour chaque sous-groupe significatif. Nos hypothèses concernant les effets des variables démographiques sont pour la plupart vérifiées.

3.1 Principaux effets des variables démographiques sur les performances des sujets témoins

3.1.1 Aucun effet du sexe

Comme attendu, le sexe n'a aucun effet significatif sur les performances aux deux épreuves, ni pour les scores ni pour les temps de réponse. Ce résultat est concordant avec la littérature : aucun effet du sexe n'est retrouvé sur les performances en dénomination orale du TDQ-30 (Macoir et al., 2021), de la BECLA (Macoir et al., 2016), de la BECS-GRECO (Merck et al., 2011) et seulement un effet mineur dans la BETL, sur les scores mais pas sur les temps de réaction (Tran & Godefroy, 2011), ainsi que dans le Boston Naming Test (Metz-Lutz et al., 1991).

3.1.2 Effet de l'âge

L'âge a un effet significatif sur les performances aux deux épreuves, tant sur les scores que sur les temps de réponse. Les analyses post-hoc ont cependant montré qu'il n'existait aucune différence significative entre les sujets âgés de 18 à 29 ans et les sujets âgés de 30 à 49 ans. Pour l'épreuve de dénomination d'objets, le score des sujets âgés de plus de 70 ans est significativement inférieur au score des sujets jeunes, même si cette différence reste faible au vu de l'effet plafond de l'épreuve (1,2 point en moins par rapport aux sujets de moins de 30 ans). Le temps de réponse est significativement plus long pour la tranche d'âge 50-69 ans (8,28 secondes de plus sur toute l'épreuve par rapport aux sujets de moins de 30 ans) et encore davantage pour la tranche d'âge 70 ans et plus (+17,99 secondes). Pour l'épreuve de dénomination d'actions, l'effet de l'âge est encore plus marqué. Le score des sujets âgés de 50 à 69 ans est significativement inférieur au score des sujets jeunes (-1,13 point), tout comme celui des sujets âgés de plus de 70 ans (-2,47 points). Là encore, la tranche d'âge 50-69 ans a un temps de réponse significativement allongé (+12,34 secondes) et celui des sujets de plus de 70 ans également (+31,67 secondes).

Cet effet de l'âge était attendu et est retrouvé dans toutes les études analysant l'effet des variables démographiques sur les performances en dénomination. En effet, les sujets âgés ont tendance à dénommer correctement moins d'items avec un temps de réponse plus long que les sujets jeunes (Macoir et al., 2016, 2021; Merck et al., 2011; Metz-Lutz et al., 1991; Tran & Godefroy, 2011). L'absence d'effet significatif entre les deux tranches d'âge des sujets de moins de 50 ans est retrouvée également dans la normalisation de la BETL (Tran & Godefroy, 2011).

3.1.3 Effet du niveau d'études

Nous relevons également un effet du niveau d'études sur les performances des sujets, mais uniquement sur l'épreuve de dénomination d'actions. A l'épreuve de dénomination d'actions, les sujets n'ayant pas le Bac obtiennent un score significativement inférieur (-1,20 point) et un temps de réponse plus important (+6 secondes) que les sujets ayant le Bac ou plus.

Aucun effet significatif du niveau d'études n'est retrouvé aux scores ou aux temps de réponses de l'épreuve de dénomination d'objets. Cela peut être mis en relation avec la dissociation retrouvée dans la littérature entre la production de noms et la production de verbes. La production de verbes est en effet plus complexe sur le plan syntaxique et sur le plan de l'abstraction, ce qui pourrait expliquer cet effet du niveau d'études. A contrario, l'épreuve de dénomination d'objets est composée d'items simples et fréquents, ce qui pourrait expliquer l'absence d'effet retrouvé.

Les analyses post-hoc ont montré qu'une différence significative n'était retrouvée que pour le niveau d'études inférieur au Bac : il n'y a pas de différence significative sur les performances en dénomination entre les sujets ayant un niveau compris entre Bac et Bac+3 et les sujets ayant un niveau supérieur à Bac+3. Cela peut s'expliquer par le fait que l'épreuve de dénomination est une épreuve simple qui ne requiert pas de compétences spécifiques et serait donc peu sensible à la durée des études supérieures. Il serait intéressant également d'analyser de manière approfondie les interactions entre les variables pour observer si l'âge n'est pas un facteur de confusion et n'explique pas mieux les différences retrouvées.

3.2 Etalonnage du test

3.2.1 Définition des sous-groupes significatifs

L'analyse des effets a permis de choisir les critères pertinents pour l'étalonnage des deux épreuves. Le sexe n'ayant aucun effet sur les performances, il n'a pas été retenu comme un critère pertinent pour la stratification des normes. Etant donné qu'aucune différence significative n'a été retrouvée entre la tranche d'âge 18-29 ans et la tranche d'âge 30-49 ans, ces deux groupes ont été fusionnés pour ne former qu'un seul groupe : trois tranches d'âge sont donc retenues (18-49 ans ; 50-69 ans ; 70 ans et plus). De la même façon, puisqu'aucune différence significative n'a été retrouvée entre les sujets du niveau d'études Bac-<Bac+3 et les sujets du niveau d'études Bac+3 et plus, les sujets ont été regroupés : deux niveaux d'étude sont donc retenus (< Bac ; > Bac). Enfin, nous avons choisi d'effectuer deux étalonnages différents pour les deux épreuves puisqu'aucun effet du niveau d'études n'a été retrouvé à l'épreuve de dénomination d'objets contrairement à l'épreuve de dénomination d'actions. Pour l'épreuve de dénomination d'objets, les normes ont ainsi été stratifiées en trois groupes avec comme seul critère pertinent l'âge, tandis que pour l'épreuve de dénomination d'actions, les normes ont été stratifiées en six groupes avec comme critères pertinents l'âge et le niveau d'études.

3.2.2 Etalonnages des épreuves

Les données normatives sont présentées en annexe n°9. Au vu de la distribution non gaussienne des données, la répartition des performances en percentiles a été choisie. Les écarts de scores entre les percentiles étant assez réduits, le détail des percentiles 5, 10, 25, 50, 75, 90, 95 plutôt qu'une graduation de 10 en 10 est apparu suffisant. Par ailleurs, la visée à la fois diagnostique et comparative du Multimap justifie l'indication des percentiles plutôt qu'un score seuil. Le seuil pathologique est fixé à P5 (5% des performances les plus faibles) et le seuil de fragilité à P10 (10% des performances les plus faibles). Pour les temps de réponse qui étaient à l'origine inversés (P5 correspondant aux temps de réponse les plus courts et P95 les plus longs),

nous avons fait le choix de transposer les données afin de garantir une meilleure lisibilité et d'éviter ainsi les erreurs d'interprétation en clinique.

4. Limites et directions futures

4.1 Limites et pistes d'amélioration du test

Le test que nous présentons dans sa version finalisée et étalonnée en français offre de nombreux avantages. Il convient néanmoins d'en souligner quelques limites et pistes d'amélioration.

4.1.1 Explicitation de la consigne

L'analyse qualitative des résultats a révélé qu'un certain nombre de sujets produisaient une réponse avec expansion (nom de l'item suivi d'un adjectif ou verbe suivi d'un complément). En effet, certaines expérimentatrices ont donné la consigne aux sujets de ne rien ajouter avant le début de l'épreuve et d'autres l'ont spécifiée seulement lorsque les items d'exemple étaient échoués. Afin d'optimiser la standardisation du test et pour éviter une trop grande variabilité des réponses, peut-être faudrait-il expliciter la consigne écrite précédant le test en spécifiant que pour chacun des items, seuls le nom et le verbe sont attendus.

4.1.2 Contrôle des variables psycholinguistiques

Les auteurs du Multimap original ont contrôlé les variables psycholinguistiques des items (fréquence, nombre de lettres, familiarité, imageabilité, concrétude) en comparant les moyennes de ces variables pour chaque épreuve (substantifs et verbes) et n'ont pas trouvé de différence significative. La longueur et la proximité phonologique des mots ont également été contrôlées pour la liste espagnole, mais pas pour la liste française. Il serait donc intéressant de contrôler au sein de chaque épreuve en français les variables psycholinguistiques comme la longueur ou la fréquence. De plus, l'analyse de la corrélation entre réussite par item et temps moyen de réponse par item a montré qu'en moyenne, plus un item était réussi, plus il était dénommé rapidement. Il serait intéressant d'observer si la réussite et le temps de réponse sont corrélés avec les variables psycholinguistiques classiques (fréquence, concrétude, longueur, imageabilité), dont il a été montré qu'elles avaient un impact significatif sur les performances en dénomination (Tran & Godefroy, 2011).

4.1.3 Effets plafond

La distribution des scores montre d'importants effets plafond, pouvant nuire à la sensibilité du test. En effet, il pourrait ne pas détecter des différences minimes, tout comme d'autres tests retrouvant des effets plafond (Macoir et al., 2016). Cependant, l'étalonnage des temps de réponse permet de contourner cet effet plafond.

4.2 Limites de l'étude : biais méthodologiques

Il est important de noter que certains biais méthodologiques ont pu influencer les résultats de notre étude. Cependant, la taille conséquente de l'échantillon permet de lisser ces irrégularités et les biais présentés ci-dessous ont des répercussions minimales sur les résultats présentés.

4.2.1 Biais liés aux conditions de passation

Malgré la formation des expérimentatrices, il n'est pas exclu qu'il y ait eu des erreurs de cotation lors des passations. En effet, le support digital offre de nombreux avantages, principalement pour l'enregistrement des temps de réponse, mais peut aussi conduire à des erreurs de cotation (sélection de la mauvaise touche pour valider ou invalider la réponse, appuis trop rapides avant que le sujet ait eu le temps de répondre...). En dépit de consignes claires, il est possible également qu'il y ait eu des différences de tolérance dans la cotation des items, notamment pour l'acceptation ou non des réponses avec expansions (qualificatifs, compléments d'objet...). Ces erreurs minimes peuvent avoir de légères répercussions sur les pourcentages de réponses alternatives présentés dans l'analyse qualitative des résultats.

4.2.2 Biais liés au développement du test

Le test étant encore en cours de développement sur la plateforme HappyNeuron, nous avons été confrontés à divers problèmes : l'exportation de données s'est révélée complexe et s'est déroulée en différentes étapes. En effet, plusieurs erreurs s'étaient glissées dans les premiers tableaux de données, ce qui a retardé l'analyse statistique des résultats. De plus, l'enregistrement des réponses alternatives reportées sur le logiciel par les expérimentatrices a été effacé pour quelques sujets. Cela n'a pas eu d'influence sur les scores ni sur les pourcentages de réussite par item puisque le problème n'a pas affecté la cotation correcte ou incorrecte des items, c'est pourquoi nous avons choisi de conserver les sujets concernés dans nos analyses. Cependant, cela a pu une nouvelle fois avoir une influence minime sur les pourcentages d'occurrence des réponses alternatives.

4.3 Directions futures

Cette étude pourrait être améliorée en fonction des limites énoncées précédemment : une analyse approfondie des variables psycholinguistiques des items composant le test serait pertinente et une comparaison de ces variables entre l'épreuve des substantifs et l'épreuve des verbes permettrait de renforcer la robustesse méthodologique du test.

En outre, cette étude doit être poursuivie par une validation auprès de patients et par la vérification des qualités psychométriques indispensables à la construction de tout test. Les résultats des patients au test français du Multimap pourraient être comparés aux résultats des tests classiquement utilisés pour l'évaluation langagière comme la DO 80. En outre, il sera intéressant d'analyser et de comparer les résultats d'éventuelles normalisations du Multimap dans d'autres pays. En effet, le Multimap pourrait devenir un outil international, ce qui permettrait d'analyser les résultats de grandes cohortes de patients entre les centres de chirurgie éveillée pour favoriser les études à grande échelle.

Conclusion

Cette étude a permis de finaliser et d'étalonner un test de dénomination orale comportant une épreuve de dénomination d'objets et une épreuve de dénomination d'actions, le Multimap. Cet outil a été conçu spécifiquement pour être utilisé auprès de patients atteints de gliomes et opérés en chirurgie éveillée, contrairement aux tests de dénomination actuellement utilisés dans

ce cadre. Il offre ainsi une banque de 80 images en couleurs (40 objets et 40 actions), dont le taux d'accord sur la dénomination est élevé (supérieur à 74%). La digitalisation du test permet également de mesurer avec précision les temps de réponse des sujets par item et total.

Les performances de 416 sujets sains, âgés de 18 à 95 ans et répartis de manière équitable selon trois variables démographiques (sexe, âge et niveau d'études) ont fait l'objet d'analyses quantitatives et qualitatives. Celles-ci ont permis de finaliser le test en éliminant les items dont le taux de réussite était insuffisant et en statuant sur les acceptions des différents items. Les effets des variables démographiques âge, sexe et niveau d'études sur les performances des sujets ont ensuite été étudiés afin d'étalonner les normes pour chaque épreuve. De plus, des analyses comparatives et corrélatives des performances des sujets sains aux deux épreuves ont été effectuées.

Nos hypothèses ont été en partie vérifiées : comme retrouvé dans la littérature, le sexe n'a pas d'influence sur les résultats tandis que l'âge a un impact significatif. En outre, comme nous en avions fait l'hypothèse, le niveau d'études influe significativement sur les performances, mais uniquement pour l'épreuve de dénomination d'actions. En effet, la dénomination de verbes, plus abstraits et plus complexes sur le plan syntaxique, requiert probablement davantage de ressources, et met en jeu des réseaux neuroanatomiques au moins partiellement distincts de ceux de la dénomination de noms. Nous retrouvons d'ailleurs une différence significative entre les moyennes des scores et des temps aux deux épreuves, conformément à la littérature, ce qui souligne la pertinence d'un test évaluant à la fois les noms et les verbes. Enfin, comme attendu, nous retrouvons des corrélations entre les performances des sujets sains aux deux épreuves.

Suite aux analyses statistiques, les normes ont ainsi été stratifiées en trois groupes significatifs pour l'épreuve de dénomination d'objets avec comme seul critère pertinent l'âge (18-49 ans ; 50-69 ans ; 70 ans et plus), et en six groupes significatifs pour l'épreuve de dénomination d'actions avec comme critères pertinents l'âge et le niveau d'études (inférieur ou supérieur au baccalauréat). Les percentiles ont été choisis pour rendre compte de la répartition des performances des sujets contrôles en raison de la distribution non normale des données.

Notre étude comporte quelques limitations et différents biais méthodologiques ont pu influencer légèrement sur la fiabilité des mesures. Cependant, la taille conséquente de l'échantillon permet de lisser ces éventuelles irrégularités. Il pourra être intéressant d'explicitier la consigne donnée aux sujets afin d'éviter au maximum les réponses avec expansions. En outre, l'analyse approfondie des variables psycholinguistiques permettrait de renforcer la robustesse méthodologique du test.

Enfin, cette étude doit être poursuivie par la validation du test auprès de patients en chirurgie éveillée et par l'analyse des qualités psychométriques de celui-ci. Les résultats des patients au test devront être comparés aux résultats à d'autres tests de dénomination orale actuellement utilisés, tels que la DO 80. Il pourra également être intéressant de comparer nos résultats aux résultats des normalisations du Multimap dans d'autres pays. A terme, le Multimap pourrait devenir un outil commun aux différents centres de chirurgie éveillée afin d'uniformiser les protocoles. Cela favorisera les analyses de plus grandes cohortes de patients à l'échelle européenne pour consolider les études en chirurgie éveillée et permettre l'avancée des connaissances en neurosciences.

Bibliographie

Aguert, M., & Capel, A. (2018). *Mieux comprendre les scores z pour bien les utiliser. Rééducation Orthophonique, 274*, 61-85.

Bogka, N., Masterson, J., Druks, J., Fragkioudaki, M., Chatziprokopiou, E.-S., & Economou, K. (2003). Object and action picture naming in English and Greek. *European Journal of Cognitive Psychology, 15*(3), 371-403. <https://doi.org/10.1080/09541440303607>

Burke, D. M., & Shafto, M. A. (2004). Aging and Language Production. *Current Directions in Psychological Science, 13*(1), 21-24. <https://doi.org/10.1111/j.0963-7214.2004.01301006.x>

Clément, A., & Perez, A. (2020). *Chirurgie éveillée des gliomes : normalisation d'un bilan de langage oral en périopératoire*. Université Paris-Sorbonne.

Corina, D. P., Gibson, E. K., Martin, R., Poliakov, A., Brinkley, J., & Ojemann, G. A. (2005). Dissociation of action and object naming : Evidence from cortical stimulation mapping. *Human Brain Mapping, 24*(1), 1-10. <https://doi.org/10.1002/hbm.20063>

Corina, D. P., Loudermilk, B. C., Detwiler, L., Martin, R. F., Brinkley, J. F., & Ojemann, G. (2010). Analysis of naming errors during cortical stimulation mapping : Implications for models of language representation. *Brain and Language, 115*(2), 101-112. <https://doi.org/10.1016/j.bandl.2010.04.001>

Deloche, G., & Hannequin, D. (1997). *DO 80 : épreuve de dénomination orale d'images*. ECPA.

De Witt Hamer, P. C., Robles, S. G., Zwinderman, A. H., Duffau, H., & Berger, M. S. (2012). Impact of Intraoperative Stimulation Brain Mapping on Glioma Surgery Outcome : A Meta-Analysis. *Journal of Clinical Oncology, 30*(20), 2559-2565. <https://doi.org/10.1200/JCO.2011.38.4818>

Duffau, H. (2005). Lessons from brain mapping in surgery for low-grade glioma : Insights into associations between tumour and brain plasticity. *The Lancet Neurology, 4*(8), 476-486. [https://doi.org/10.1016/S1474-4422\(05\)70140-X](https://doi.org/10.1016/S1474-4422(05)70140-X)

Duffau, H. (2010). Chirurgie cérébrale en condition éveillée. *La Lettre du neurologue, 14*(2), 43-46.

Duffau, H. (2014). Gliomes diffus de bas grade et neuroplasticité. *Journal de Radiologie Diagnostique et Interventionnelle, 95*(10), 935-945. <https://doi.org/10.1016/j.jradio.2014.07.002>

Duffau, H. (2018). Chirurgie éveillée des gliomes cérébraux. *Interbloc, 37*(3), 156-159. <https://doi.org/10.1016/j.bloc.2018.06.004>

Duffau, H., Moritz-Gasser, S., & Mandonnet, E. (2014). A re-examination of neural basis of language processing : Proposal of a dynamic hodotopical model from data provided by brain stimulation mapping during picture naming. *Brain and Language*, *131*, 1-10. <https://doi.org/10.1016/j.bandl.2013.05.011>

Gisbert-Muñoz, S., Quiñones, I., Amoruso, L., Timofeeva, P., Geng, S., Boudelaa, S., Pomposo, I., Gil-Robles, S., & Carreiras, M. (2021). MULTIMAP : Multilingual picture naming test for mapping eloquent areas during awake surgeries. *Behavior Research Methods*, *53*(2), 918-927. <https://doi.org/10.3758/s13428-020-01467-4>

Gobbo, M., De Pellegrin, S., Bonaudo, C., Semenza, C., Della Puppa, A., & Salillas, E. (2021). Two dissociable semantic mechanisms predict naming errors and their responsive brain sites in awake surgery. DO80 revisited. *Neuropsychologia*, *151*, 107727. <https://doi.org/10.1016/j.neuropsychologia.2020.107727>

Grégoire, J. (2004). *L'examen clinique de l'intelligence de l'adulte*. Mardaga. <https://doi.org/10.3917/mard.grego.2004.01>

Herbet, G., Rigaux-Viodé, O., & Moritz-Gasser, S. (2017). Peri- and intraoperative cognitive and language assessment for surgical resection in brain eloquent structures. *Neurochirurgie*, *63*(3), 135-141. <https://doi.org/10.1016/j.neuchi.2016.10.011>

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>

Jakola, A. S., Myrmet, K. S., Kloster, R., Torp, S. H., Lindal, S., Unsgård, G., & Solheim, O. (2012). Comparison of a Strategy Favoring Early Surgical Resection vs a Strategy Favoring Watchful Waiting in Low-Grade Gliomas. *JAMA*, *308*(18), 1881. <https://doi.org/10.1001/jama.2012.12807>

Lemaitre, A. L., Herbet, G., Ng, S., Moritz-Gasser, S., & Duffau, H. (2022). Cognitive preservation following awake mapping-based neurosurgery for low-grade gliomas: A longitudinal, within-patient design study. *Neuro-oncology*, *24*(5), 781–793.

Lubrano, V., Roux, F.-E., & Démonet, J.-F. (2012). Explorations du langage par stimulations électriques directes peropératoires. *Revue de neuropsychologie*, *4*(2), 97-102.

Lubrano, V., Filleron, T., Démonet, J.-F., & Roux, F.-E. (2014). Anatomical correlates for category-specific naming of objects and actions : A brain stimulation mapping study: Object and Action Naming. *Human Brain Mapping*, *35*(2), 429-443. <https://doi.org/10.1002/hbm.22189>

Luzzatti, C., Raggi, R., Zonca, G., Pistarini, C., Contardi, A., & Pinna, G.-D. (2002). Verb–Noun Double Dissociation in Aphasic Lexical Impairments : The Role of Word Frequency and Imageability. *Brain and Language*, *81*(1-3), 432-444. <https://doi.org/10.1006/brln.2001.2536>

- 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. <https://doi.org/10.1016/j.jns.2016.01.004>
- Macoir, J., Chagnon, A., Hudon, C., Lavoie, M., & Wilson, M. A. (2021). TDQ-30—A New Color Picture-Naming Test for the Diagnostic of Mild Anomia : Validation and Normative Data in Quebec French Adults and Elderly. *Archives of Clinical Neuropsychology*, 36(2), 267-280. <https://doi.org/10.1093/arclin/acz048>
- Mandonnet, E. (2014, 27 mai). *Neurochirurgie guidée par l'image* [conférence]. Informatique et sciences numériques, Amphithéâtre Maurice Halbwachs, Collège de France. <https://www.college-de-france.fr/site/nicholas-ayache/seminar-2014-05-27-17h30.htm>
- Mandonnet, E., Sarubbo, S., & Duffau, H. (2016). Proposal of an optimized strategy for intraoperative testing of speech and language during awake mapping. *Neurosurgical Review*. <https://doi.org/10.1007/s10143-016-0723-x>
- Mandonnet, E., Taillandier, L., & Duffau, H. (2017). Proposition de dépistage des gliomes diffus de bas grade dans la population de 20 à 40 ans. *La Presse Médicale*, 46(10), 911-920. <https://doi.org/10.1016/j.lpm.2017.07.005>
- Mandonnet, E., & Herbet, G. (Eds.). (2021). *Intraoperative Mapping of Cognitive Networks: Which Tasks for Which Locations*. Springer Nature.
- Martín-Monzón, I., Ballagas, Y. R., & Arias-Sánchez, S. (2020). Language mapping : A systematic review of protocols that evaluate linguistic functions in awake surgery. *Applied Neuropsychology: Adult*, 0(0), 1-10. <https://doi.org/10.1080/23279095.2020.1776287>
- Merck, C., Charnallet, A., Auriacombe, S., Belliard, S., Hahn-Barma, V., Kremin, H., Lemesle, B., Mahieux, F., Moreaud, O., Palisson, D. P., Roussel, M., Sellal, F., & Siegwart, H. (2011). La batterie d'évaluation des connaissances sémantiques du GRECO (BECS-GRECO) : Validation et données normatives. *Revue de neuropsychologie*, 3(4), 235. <https://doi.org/10.3917/rne.034.0235>
- Metz-Lutz, M. N., Kremin, H., Deloche, G., Hannequin, D., Ferrand, I., Perrier, D., Quint, S., Dordain, M., Bunel, G., Cardebat, D., & Lota, A. M. (1991). Standardisation d'un test de dénomination orale : Contrôle des effets de l'âge, du sexe et du niveau de scolarité chez les sujets adultes normaux. *Rev Neuropsychol*, 1, 9.
- Mitrushina, M., Boone, K. B., Razani, J. et D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment*. Oxford University Press.
- Moritz-Gasser, S., Herbet, G., Maldonado, I. L., & Duffau, H. (2012). Lexical access speed is significantly correlated with the return to professional activities after awake surgery for low-grade gliomas. *Journal of Neuro-Oncology*, 107(3), 633-641. <https://doi.org/10.1007/s11060-011-0789-9>

Papatzalas, C., Fountas, K., Kapsalaki, E., & Papathanasiou, I. (2021). The Use of Standardized Intraoperative Language Tests in Awake Craniotomies : A Scoping Review. *Neuropsychology Review*, 1-31. <https://doi.org/10.1007/s11065-021-09492-6>

Rofes, A. (2015). *Verbs and nouns in awake neurosurgery: needs and answers* (Doctoral dissertation, University of Trento).

Rofes, A., & Miceli, G. (2014). Language Mapping with Verbs and Sentences in Awake Surgery : A Review. *Neuropsychology Review*, 24(2), 185-199. <https://doi.org/10.1007/s11065-014-9258-5>

Rofes, A., Spena, G., Miozzo, A., Fontanella, M. M., & Miceli, G. (2015). Advantages and disadvantages of intraoperative language tasks in awake surgery : A three-task approach for prefrontal tumors. *Journal of Neurosurgical Sciences*, 59(4), 337-349.

Rofes, A., Mandonnet, E., Godden, J., Baron, M. H., Colle, H., Darlix, A., de Aguiar, V., Duffau, H., Herbet, G., Klein, M., Lubrano, V., Martino, J., Mathew, R., Miceli, G., Moritz-Gasser, S., Pallud, J., Papagno, C., Rech, F., Robert, E., ... Wager, M. (2017). Survey on current cognitive practices within the European Low-Grade Glioma Network : Towards a European assessment protocol. *Acta Neurochirurgica*, 159(7), 1167-1178. <https://doi.org/10.1007/s00701-017-3192-2>

Rofes, A., Mandonnet, E., de Aguiar, V., Rapp, B., Tsapkini, K., & Miceli, G. (2019). Language processing from the perspective of electrical stimulation mapping. *Cognitive Neuropsychology*, 36(3-4), 117-139. <https://doi.org/10.1080/02643294.2018.1485636>

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's Object Pictorial Set : The Role of Surface Detail in Basic-Level Object Recognition. *Perception*, 33(2), 217-236. <https://doi.org/10.1068/p5117>

Satoer, D., De Witte, E., Bastiaanse, R., Vincent, A., Mariën, P. et Visch-Brink, E. (2019). Diagnostic Instrument for Mild Aphasia (DIMA): standardization and clinical application. Manuscript soumis pour publication.

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*, Volume 3(1), 52-69.

Turkstra, L. S., Coelho, C., & Ylvisaker, M. (2005). The Use of Standardized Tests for Individuals with Cognitive-Communication Disorders. *Seminars in Speech and Language*, 26(04), 215-222. <https://doi.org/10.1055/s-2005-922101>

Liste des annexes

Annexe n°1 : Modèle hodotopique dynamique sous-tendant la dénomination orale d'images, proposé par Duffau et al. (2014)

Annexe n°2 : Récépissé de déclaration de conformité à la réglementation des données à caractère personnel

Annexe n°3 : Note d'information et formulaire de consentement

Annexe n°4 : Questionnaire d'inclusion

Annexe n°5 : Pourcentage de réussite par item avant la finalisation de l'épreuve

Annexe n°6 : Mesure des corrélations entre les épreuves (scores et temps)

Annexe n°7 : Analyses descriptives et test de la normalité des données

Annexe n°8 : Effet des variables démographiques sur les performances des sujets témoins

Annexe n°9 : Données normatives

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



 **Université
de Lille**



ANNEXES

DU MEMOIRE

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

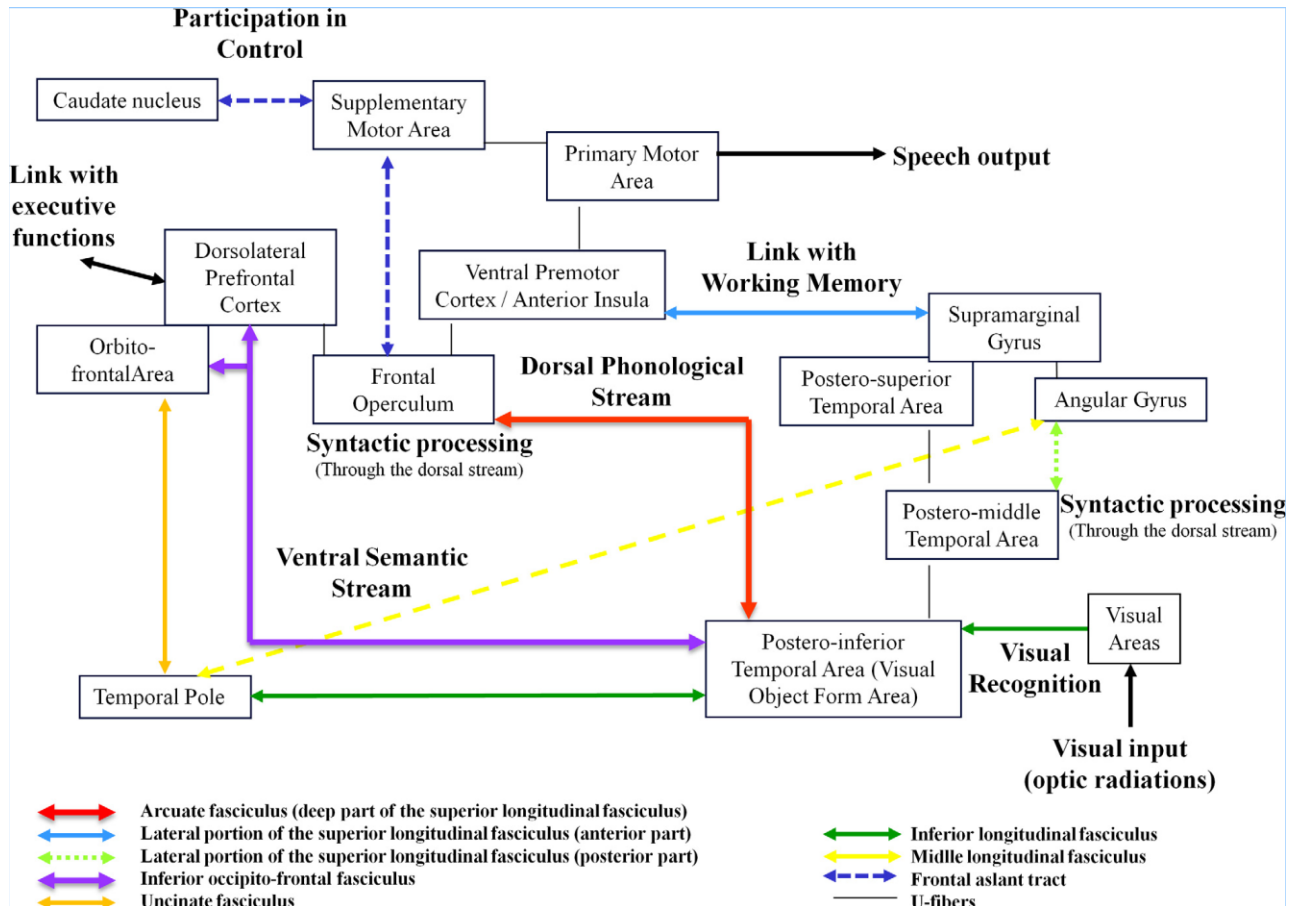
Anouk MALEWSKI

**Evaluation du langage en péri-opératoire de la
chirurgie éveillée des gliomes
Normalisation d'un test de dénomination orale**

MEMOIRE dirigé par

Isabelle POISSON, Orthophoniste, Hôpital Lariboisière, Paris
Corinne HENGUELLE, Orthophoniste, Hôpital Salengro, Lille

Annexe 1 : Modèle hodotopique dynamique sous-tendant la dénomination orale d'images, proposé par Duffau et al. (2014)



Annexe 2 : Récépissé de déclaration de conformité à la réglementation des données à caractère personnel



Direction
Données personnelles
et archives

RÉCÉPISSÉ DÉCLARATION DE CONFORMITÉ

Délégué à la protection des données (DPO) Jean-Luc TESSIER

Responsable administrative Clémentine Dehay

La délivrance de ce récépissé atteste que votre traitement est conforme à la réglementation applicable à la protection des données personnelles. Vous pouvez désormais mettre en œuvre votre traitement dans le strict respect des mesures qui ont été élaborées avec le DPO et qui figurent sur votre déclaration.

Toute modification doit être signalée dans les plus brefs délais : dpo@univ-lille.fr

Responsable du traitement

Nom : Université de Lille	SIREN: 13 00 23583 00011
Adresse : 42 Rue Paul Duez 590000 - LILLE	Code NAF: 8542Z Tél. : +33 (0) 3 62 26 90 00

Traitement déclaré

Intitulé : Normaliser en français, indépendamment de aspects multi linguistiques de l'épreuve, un test de dénomination orale d'images, le MULTIMAP. Établir des normes par rapport test dénomination orale.

Référence Registre DPO : 2021-131

Responsable du traitement / Chargé (e) de la mise en œuvre : M. Dominique LACROIX
Interlocuteur (s) : Mme. Anouk MALEWSKI

Fait à Lille,

Le 17 Juin 2021

Jean-Luc TESSIER

Délégué à la Protection des Données

Annexe 3 : Note d'information et formulaire de consentement



NOTE D'INFORMATION

Projet de recherche **Normalisation d'un test de dénomination orale en chirurgie éveillée**

Vous êtes invité(e) à participer à un projet de recherche dans le cadre du mémoire de fin d'études en orthophonie d'Anouk Malewski, sous la supervision du Professeur Emmanuel Mandonnet (neurochirurgien), de Madame Marion Barberis et de Madame Isabelle Poisson (orthophonistes) exerçant à l'hôpital Lariboisière à Paris, ainsi que de Madame Corinne Henguelle (orthophoniste) exerçant à l'hôpital Salengro à Lille. Ce projet de recherche a pour objectif de constituer une population contrôle afin d'obtenir les normes à un test de dénomination orale d'images destiné à l'évaluation du langage lors des chirurgies éveillées des tumeurs cérébrales.

Cette note d'information vous renseigne sur les modalités de ce projet afin de prendre une décision éclairée. N'hésitez pas à poser toutes les questions nécessaires avant de signer le consentement ci-après.

La chirurgie éveillée ? C'est le traitement de référence de certaines tumeurs cérébrales touchant les zones du langage, de la motricité, des émotions... Durant l'opération, le patient, éveillé, réalise des tâches cognitives ou motrices sous supervision de l'orthophoniste, pendant que le chirurgien pratique une électrostimulation de la zone concernée. Cette technique permet de retirer le plus de tumeur possible tout en limitant les séquelles fonctionnelles afin de garantir une meilleure qualité de vie au patient.

Le test de dénomination orale ? C'est une épreuve qui consiste à donner le mot correspondant à une image présentée (objet ou action), classiquement utilisée lors de l'évaluation du langage. En chirurgie éveillée, elle permet d'identifier d'éventuels troubles chez un patient avant et après l'opération et de repérer les zones essentielles à conserver durant la chirurgie.

Votre participation ? Un entretien d'une trentaine de minutes comportant un court questionnaire et l'épreuve de dénomination.

Les intérêts ? Evaluer plus finement le langage des patients opérés en chirurgie éveillée et participer à la création d'un outil international commun facilitant ainsi la recherche.

Les risques ? Aucun : pas d'utilisation de matériel invasif.

Confidentialité ?

Vos données seront recueillies et codées et ne seront analysées que par nos encadrants et nous-mêmes. Nous nous engageons à ce qu'aucune information permettant de vous identifier ne soit divulguée sous quelque forme que ce soit.



FORMULAIRE DE CONSENTEMENT

Je soussigné(e),

NOM : **Prénom :**

accepte librement de participer à la recherche intitulée « Normalisation d'un test de dénomination orale pour l'évaluation du langage en péri-opératoire de la chirurgie éveillée des gliomes » qui m'est proposée par, et qui est réalisée dans le cadre d'un travail de Master en orthophonie sous la responsabilité de Madame Marion Barberis, Madame Isabelle Poisson et Madame Corinne Henguelle.

- J'ai pris connaissance de la note d'information m'expliquant l'objectif de cette recherche, la façon dont elle sera réalisée et ce que ma participation impliquera.
- J'ai reçu des réponses adaptées à toutes mes questions.
- J'ai disposé d'un temps suffisant pour prendre ma décision.
- J'ai compris que ma participation est libre que je peux y mettre fin à tout moment.
- J'ai noté que je dispose du droit d'accès à mes données prévu par la loi « Informatique et Libertés » (article 39).
- Je conserverai un exemplaire de la note d'information et du consentement.

Signature de la personne participant à la recherche

Signature de l'investigateur (étudiant)

Nom Prénom :

Nom Prénom :

Date :

Signature :

Date :

Signature :

Document réalisé en 2 exemplaires. Un exemplaire doit être conservé par l'investigateur, le deuxième doit être remis à la personne participant à la recherche.

Annexe 4 : Questionnaire d'inclusion



QUESTIONNAIRE D'INCLUSION

SEXE : M F

DATE ET LIEU DE NAISSANCE :

LANGUE MATERNELLE :

SCOLARITE : en France à l'étranger

NIVEAU D'ETUDES : < BAC BAC à BAC+3 > BAC+3

PLUS HAUT DIPLOME OBTENU :

Êtes-vous ou avez-vous déjà été suivi pour une pathologie cardio-vasculaire (trouble du rythme cardiaque, athérosclérose, hypertension artérielle, infarctus du myocarde) ?

Oui Non

Êtes-vous ou avez-vous déjà été suivi pour une pathologie neurologique (tumeur, épilepsie, accident vasculaire cérébral, maladie neurodégénérative) ?

Oui Non

Êtes-vous ou avez-vous déjà été suivi pour une pathologie psychiatrique ?

Oui Non

Êtes-vous ou avez-vous déjà été suivi pour un problème de dépendance à un toxique (alcool, drogues) ?

Oui Non

Êtes-vous ou avez-vous déjà été suivi pour un trouble du langage ?

Oui Non

Je m'engage à l'exactitude des informations communiquées. Les données recueillies seront anonymisées et demeureront strictement confidentielles.

Signature :

Annexe 5 : Pourcentage de réussite par item avant la finalisation de l'épreuve

Substantifs	Taux de réussite (%)	Verbes	Taux de réussite (%)
serviette	55,53	imaginer	1,92
bière	59,86	confesser	22,84
désert	60,34	accélérer	29,09
bouquet	60,58	livrer	42,79
ciseaux	70,67	traverser	44,47
mouche	71,88	aider	49,76
berceau	76,2	transpirer	50,24
cadeau	78,13	fermer	56,25
béret	78,37	attendre	59,13
château	79,33	servir	61,06
tasse	80,77	découper	65,38
plage	81,97	vendre	68,99
lune	82,21	surligner	70,19
dos	84,13	gagner	73,56
manteau	84,13	signer	80,53
vélo	85,34	étendre	80,53
fleur	88,22	construire	82,45
glaçon	88,46	toucher	82,45
oignon	89,66	creuser	83,17
beurre	89,9	saluer	83,17
cuisine	90,87	poster	83,41
jambe	90,87	se moucher	83,65
épaule	91,35	ranger	84,62
nœud	91,59	nettoyer	87,5
bras	93,75	chasser	88,7
queue	94,23	regarder	88,7
serrure	94,23	allumer	89,66
dent	95,43	casser	89,9
cadenas	96,15	brûler	91,83
calendrier	96,15	monter	92,31
cochon	96,39	se marier	93,27
couteau	97,36	écouter	94,95
corde	97,6	peindre	95,19
cheval	98,32	peser	95,91
chemise	98,32	sauter	96,63
journal	98,32	voler	97,12
timbre	98,32	prier	97,36
œuf	98,56	s'asseoir	97,36
fauteuil	98,8	chanter	97,6
poisson	98,8	jouer	97,6
chaise	99,04	fumer	98,32
écureuil	99,04	mesurer	98,56
lapin	99,04	écrire	99,04
os	99,28	pousser	99,04
ceinture	99,52	boire	99,28
parapluie	99,52	conduire	99,52
église	99,76	manger	99,52
feuille	99,76	courir	99,76
papillon	99,76	lire	99,76
chapeau	100	pêcher	99,76

Tableau 6 : Pourcentage de réussite par item avant la finalisation de l'épreuve

Annexe 6 : Mesure des corrélations entre les épreuves

		Score substantifs	Score verbes	Temps substantifs	Temps verbes
Score substantifs	Kendall's Tau B	—			
Score verbes	Kendall's Tau B	0.277***	—		
Temps substantifs	Kendall's Tau B	-0.238***	-0.145***	—	
Temps verbes	Kendall's Tau B	-0.161***	-0.263***	0.586***	—

Note : *** p < 0.001
Corrélation *modérée* (entre -0.1 et -0.5 ou entre 0.1 et 0.5) ; corrélation *forte* (entre -0.5 et -1 ou entre 0.5 et 1)

Tableau 7 : Résultats des corrélations entre les performances aux épreuves (scores et temps)

Annexe 7 : Analyses descriptives et test de la normalité des données

	Score substantifs	Temps global substantifs (s)	Score verbes	Temps global verbes (s)
Moyenne	38.4	82.9	36.7	107
Médiane	39.0	80.2	37.0	99.9
Ecart-type	1.58	16.9	2.33	27.1
Minimum	32	54.3	27	62.3
Maximum	40	159	40	220
Shapiro-Wilk W	0.854	0.910	0.925	0.890
Shapiro-Wilk p	< .001	< .001	< .001	< .001

Tableau 8 : Analyses descriptives et test de la normalité des données

Annexe 8 : Effet des variables démographiques sur les performances des sujets témoins

			Score substantifs	Temps substantifs	Score verbes	Temps verbes
Sexe	Global		0.845	0.298	0.545	0.367
NSC	Global		0.742	0.130	< .001*	0.035*
	< Bac	> Bac	0.995	0.127	< .001*	0.024*
	< Bac	> Bac+3	0.774	0.367	< .001*	0.268
	> Bac	> Bac+3	0.802	0.705	0.957	0.567
Âge	Global		< .001*	< .001*	< .001*	< .001*
	18-29	30-49	0.966	0.596	0.960	0.979
	18-29	50-69	0.206	< .001*	0.002*	< .001*
	18-29	70+	< .001	< .001*	< .001*	< .001*
	30-49	50-69	0.478	0.011*	< .001*	< .001*
	30-49	70+	< .001*	< .001*	< .001*	< .001*
	50-69	70+	0.017*	< .001*	< .001*	< .001*

Tableau 9 : Effet des variables démographiques sur les performances des sujets (*p<0.05)

Annexe 9 : Données normatives

MULTIMAP DENOMINATION ORALE D'OBJETS (SUBSTANTIFS)						
Total N=416	SCORES (/40)			TEMPS (s)		
	18-49 ans N=207	50-69 ans N=109	70 ans et plus N=100	18-49 ans N=207	50-69 ans N=109	70 ans et plus N=100
Moy.	38.8	38.4	37.7	76.9	84.2	93.9
ET	1.27	1.63	1.85	13.3	16.0	18.6
Min.	33	32	33	54.3	55.0	65.7
Max.	40	40	40	125	159	152
P5	36.0	36.0	34.0	102	105	133
P10	37.0	36.0	35.0	94.7	99.0	119
P25	38.0	37.0	37.0	84.0	89.3	102
P50	39.0	39.0	38.0	74.1	82.0	88.9
P75	40.0	40.0	39.0	67.6	75.4	81.8
P90	40.0	40.0	40.0	61.8	69.1	76.3
P95	40.0	40.0	40.0	60.1	67.3	72.9

Note : N = effectif ; Moy. = moyenne ; ET = écart-type ; Min. = minimum ; Max. = maximum ; P = percentile.

Tableau 10 : Données normatives pour l'épreuve de substantifs

MULTIMAP – DENOMINATION ORALE D'ACTIONS (VERBES)												
Total N=416	SCORES (/40)						TEMPS (s)					
	18-49 ans		50-69 ans		70 ans et plus		18-49 ans		50-69 ans		70 ans et plus	
	< BAC N=62	> BAC N=145	< BAC N=40	> BAC N=69	< BAC N=32	> BAC N=68	< BAC N=62	> BAC N=145	< BAC N=40	> BAC N=69	< BAC N=32	> BAC N=68
Moy.	37.0	37.8	35.8	36.9	33.8	35.8	98.6	94.1	112	106	132	125
ET	2.05	1.63	2.71	2.00	2.49	2.16	17.7	20.1	27.6	21.8	30.5	30.7
Min.	31	32	27	29	29	31	69.1	62.3	72.2	74.5	79.7	85.7
Max.	40	40	40	40	38	40	152	179	210	220	185	214
P5	34.0	35.0	30.9	33.4	30.5	32.0	131	129	156	141	184	192
P10	34.1	36.0	33.0	34.0	31.0	33.0	119	120	136	127	168	167
P25	36.0	37.0	34.8	36.0	32.0	34.0	110	100	123	117	157	138
P50	37.0	38.0	36.0	37.0	34.0	36.0	94.9	89.3	109	103	129	121
P75	38.0	39.0	37.3	38.0	36.0	37.0	87.1	81.6	94.3	91.5	108	105
P90	39.0	40.0	38.1	39.0	37.0	38.3	79.5	76.8	82.2	85.7	94.6	95.4
P95	40.0	40.0	39.0	39.0	37.5	39.0	75.4	70.4	78.4	79.7	92.1	89.8

Note : N = effectif ; Moy. = moyenne ; ET = écart-type ; Min. = minimum ; Max. = maximum ; P = percentile.

Tableau 11 : Données normatives pour l'épreuve de verbes