



**UNIVERSITÉ
DE GENÈVE**

Archive ouverte UNIGE

<https://archive-ouverte.unige.ch>

Master

2009

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Construction et étalonnage sur population-contrôle de deux versions courtes et parallèles du test de dénomination orale Déno-100

Sintra-Grilo, Daniela

How to cite

SINTRA-GRILO, Daniela. Construction et étalonnage sur population-contrôle de deux versions courtes et parallèles du test de dénomination orale Déno-100. Master, 2009.

This publication URL: <https://archive-ouverte.unige.ch/unige:2501>

© This document is protected by copyright. Please refer to copyright holder(s) for terms of use.



**UNIVERSITÉ
DE GENÈVE**

FACULTÉ DE PSYCHOLOGIE
ET DES SCIENCES DE L'ÉDUCATION

Août 2009

Section de Psychologie

Maîtrise Universitaire en Logopédie

**Construction et étalonnage sur population-
contrôle de deux versions courtes et parallèles
du test de dénomination orale Déno-100**

Mémoire présenté par Daniela Sintra-Grilo

Dirigé par le Professeur Jocelyne Buttet Sovilla

Jury composé de J. Buttet Sovilla, L. Schneider et U. Frauenfelder

Remerciements

Ce travail de recherche n'aurait pu être entrepris sans l'aimable collaboration de diverses personnes, que je tiens ici à remercier.

En premier lieu, merci à Helgard Kremin, qui nous a autorisés à utiliser le *Déno-100*. Mes remerciements vont également aux participants qui ont accepté de donner un peu de leur temps. Enfin, je remercie tout particulièrement ma directrice de mémoire, Madame Buttet Sovilla, pour avoir encadré cette recherche, ainsi que ses collaborateurs du service de neuropsychologie du CHUV pour m'avoir aidée dans le recrutement de participants et dans le traitement statistique des données.

Un grand merci.

Table des matières

LISTE DES ABRÉVIATIONS	3
RÉSUMÉ	4
INTRODUCTION	6
1. ASPECTS THÉORIQUES : LA DÉNOMINATION ORALE EN PSYCHOLINGUISTIQUE ET EN PSYCHOLOGIE COGNITIVE	8
1.1. Trouble de l'évocation lexicale	8
1.2. Approche psycholinguistique et approche cognitive	9
1.2.1. Approche psycholinguistique	9
1.2.2. Approche cognitive ou du traitement de l'information	10
1.3. Troubles spécifiques en dénomination orale d'images	14
1.4. Facteurs influençant les performances en dénomination orale d'images	16
1.4.1. Facteurs individuels.....	16
1.4.2. Facteurs linguistiques et extra-linguistiques.....	17
1.5. Intérêts et limites des épreuves de dénomination orale d'images	20
1.5.1. Intérêts.....	20
1.5.2. Limites.....	21
1.5.3. Comparaison avec les situations de conversation.....	22
1.6. Tests de dénomination orale en français : problèmes actuels	23
1.7. Le Déno-100 : un test de dénomination contrôlé	24
1.7.1. Banque de données pour la dénomination orale d'images et variables cruciales dans l'accès au lexique	24
1.7.2. Caractéristiques du test de dénomination <i>Déno-100</i>	27
2. BUTS DE LA RECHERCHE	29
3. MATÉRIEL ET MÉTHODE	30
3.1. Participants	30
3.2. Matériel	31
3.3. Procédure	32
3.3.1. Conditions de passation.....	32
3.3.2. Mesures et cotation des épreuves	32
3.4. Hypothèses opérationnelles	33

4.	RÉSULTATS	35
4.1.	Résultats des versions à 24 items.....	36
4.2.	Résultats des versions à 20 items.....	46
4.3.	Etalonnage	46
5.	DISCUSSION	48
5.1.	Discussion sur les résultats	48
5.2.	Critiques et limites	51
5.3.	Ouvertures	53
6.	CONCLUSIONS	54
7.	BIBLIOGRAPHIE	55
	ANNEXES	59

Liste des abréviations

AMA :	âge moyen d'acquisition des mots
CAT :	catégorie sémantique des items
CONS :	degré de consensus en dénomination orale
CVIS :	degré de complexité visuelle des dessins
FAM :	degré de familiarité des concepts à dénommer
FREQ :	fréquence d'usage des mots
liv :	« living », caractère animé des concepts représentés
NB SYLL :	nombre de syllabes
obj :	« object », caractère inanimé des concepts représentés
RD :	réponse dominante
RND - :	réponse non dominante
RND + :	réponse non dominante acceptée

Résumé

Le trouble de l'évocation lexicale est l'un des symptômes majeurs rencontrés dans les divers tableaux cliniques de la pathologie aphasique. C'est à l'aide des tests de dénomination orale d'images que l'on peut apprécier la présence ou non de ce trouble. Le présent travail consiste en la construction et l'étalonnage de deux versions courtes et parallèles d'un test de dénomination appelé *Déno-100*. Ce dernier a l'avantage d'être un matériel bien contrôlé, dans lequel les items sont caractérisés par de nombreuses variables normées dans diverses langues, dont le français. Il est toutefois trop long pour des évaluations dans le domaine clinique, d'où l'intérêt de construire des épreuves courtes et parallèles, utilisables dans une perspective de test-retest.

La population ayant participé à cette recherche est composée de 46 sujets-contrôle adultes de langue maternelle française. Trois variables individuelles caractérisant les participants ont été contrôlées en vue de l'étalonnage : l'âge, le sexe et le niveau d'éducation. Les sujets ont eu pour consigne de dénommer les dessins des deux versions courtes et parallèles, constituées de 24 items chacune. La sélection des items a été guidée par deux variables : l'âge d'acquisition des mots et le degré de consensus en dénomination orale. Plusieurs recherches ont mis en évidence que ces deux paramètres auraient une influence sur les performances en dénomination orale d'image, c'est pourquoi ils ont été choisis pour sélectionner les items des versions courtes et parallèles du *Déno-100*.

Les analyses statistiques effectuées ont porté sur l'évaluation des effets des variables individuelles sur les performances aux deux versions courtes. Les résultats obtenus vont dans le sens de ce qui a été rapporté dans la littérature, à savoir qu'il existe un effet d'âge et de niveau d'éducation. Aucun effet de sexe n'a toutefois été constaté. Les hommes et les femmes ont donc des performances similaires en dénomination orale d'images. En ce qui concerne l'âge, le taux de réussite diminue significativement à partir de 80 ans. Le temps de réponse quant à lui, augmente significativement dès 70 ans, puis à nouveau dès 80 ans. Enfin, pour le niveau d'éducation, on constate que les sujets ayant un plus haut niveau d'études ont de

meilleurs taux de réussite. Ils répondent également plus rapidement que leurs pairs ayant eu jusqu'à 9 ans de scolarité.

Ce travail présente de nombreux biais méthodologiques. Il a néanmoins permis de mettre en évidence des effets intéressants en ce qui concerne l'âge et le niveau d'éducation. Cette recherche mériterait ainsi d'être approfondie avec un plus grand nombre de participants et avec des améliorations du point de vue de la sélection des items afin de confirmer et de consolider les résultats obtenus.

Introduction

L'aphasie (du grec *phasis*, parole) est « un trouble du langage ayant pour caractéristiques essentielles d'apparaître suite à une *lésion du système nerveux central* et lorsque *le langage existait* déjà chez l'individu qui a subi l'atteinte cérébrale » (Seron & Feyereisen, 1985, p.132). En d'autres termes, l'aphasie est un trouble acquis qui touche l'hémisphère dominant pour le langage. Il n'est pas aisé de définir ce trouble de façon absolue et de se mettre d'accord sur une définition unanime, car il existe autant d'aphasies qu'il y a de personnes aphasiques. En effet, la perte langagière n'est pas forcément totale et ses manifestations sémiologiques peuvent présenter des variations considérables d'un individu à un autre. Le langage des personnes aphasiques peut être touché à deux niveaux, à savoir dans les capacités d'expression et/ou de compréhension, tant à l'oral qu'à l'écrit (expression orale, lecture et écriture, par exemple). Les étiologies sont également diverses, comprenant les accidents vasculaires cérébraux, les traumatismes crânio-cérébraux, les pathologies dégénératives, les inflammations et les infections au niveau du cerveau ou encore les tumeurs cérébrales. Touchant ainsi la communication, ce trouble langagier a d'importantes conséquences psycho-sociales et familiales qu'il convient de ne pas négliger (www.aphasie.org).

En ce qui concerne plus précisément l'expression orale, les lésions cérébrales peuvent avoir pour conséquence de la perturber à plusieurs niveaux : articulatoire (prononciation du message verbal), lexical (sélection des mots) ou encore syntaxique (agencement des mots au sein d'un énoncé) (Tran, 2007). Cette liste est non exhaustive et dépend des modèles utilisés. Dans ce travail, nous allons principalement nous intéresser aux troubles de l'évocation lexicale chez le sujet adulte en se référant aux approches de la psycholinguistique et de la neuropsychologie cognitive.

La dénomination orale d'images est une épreuve de choix dans le bilan de langage. Elle permet notamment d'évaluer l'évocation lexicale. Souvent remis en question, les tests de dénomination restent largement utilisés, tant par la communauté scientifique que par les cliniciens. Certaines variables caractérisant les items sont considérées comme pouvant avoir une influence significative sur les performances en dénomination orale d'images. C'est le cas notamment de l'âge d'acquisition des mots et du degré de

consensus en dénomination. Il semblerait qu'il soit pertinent de les prendre en compte lors de la sélection des items (Kremin, 2002). C'est d'ailleurs le but du présent travail. Il s'agit de construire deux versions courtes et parallèles du test de dénomination *Déno-100* (Kremin, Perrier, De Wilde & Dordain 2005, version expérimentale non publiée), qui est un test bien contrôlé, mais trop long pour des passations cliniques. Etant donné leur influence sur les performances en dénomination orale d'images, les paramètres d'âge d'acquisition des mots et de degré de consensus en dénomination orale sont deux éléments-clé dans la sélection des items. Dans un deuxième temps, il s'agit d'étalonner ces deux versions courtes et parallèles sur une population adulte normale. Il existe en effet beaucoup de tests de dénomination, mais peu qui comportent un matériel aussi bien contrôlé que le *Déno-100*, qui soient étalonnés en français pour la Suisse romande, et par conséquent représentatifs de cette partie de la francophonie.

1. Aspects théoriques : la dénomination orale en psycholinguistique et en psychologie cognitive

1.1. Trouble de l'évocation lexicale

Le *trouble de l'évocation lexicale* (plus communément appelé « *anomie* » ou « *manque du mot* »¹ par les cliniciens) est l'un des syndromes majeurs rencontrés suite à des lésions cérébrales, qu'elles soient uni- ou multi-focales (Kremin, 2002). Ce trouble correspond à une « difficulté d'évoquer un vocable spécifique adapté à une situation donnée (conversation spontanée, récit, description ou dénomination) » (Tran, 1997, p.17). Cette difficulté est extrêmement courante chez les personnes aphasiques et représente souvent la première plainte évoquée. La fréquence des troubles de l'évocation lexicale s'explique en partie par le fait que l'anomie est le symptôme commun aux différentes formes de la pathologie aphasique. De plus, le manque du mot reste souvent un des derniers éléments de handicap au décours de la récupération (Tran, 1997, 2007). Toutefois, ce symptôme peut également survenir de façon précoce dans le tableau sémiologique de pathologies cérébrales dégénératives ou évolutives. On l'observe par exemple dans la maladie d'Alzheimer, l'aphasie primaire progressive ou encore la démence fronto-temporale (Kremin, 2002 ; Tran, 2007).

Les difficultés de lexicalisation peuvent se manifester de multiples façons : silences, pauses dans le discours, productions erronées ou déformées (paraphasies), ou encore comportements dénominatifs palliatifs tels que les circonlocutions ou les conduites d'approche, par exemple. Les patients auront ainsi recours à diverses stratégies compensatoires pour s'exprimer, selon les modules langagiers préservés (Tran, 1997, 2007). Si le trouble d'évocation lexicale est important et que les stratégies compensatoires sont peu efficaces, les compétences communicationnelles peuvent être affectées de façon significative en raison du manque d'informativité du discours (Tran, 2007 ; Herbert, Hickin, Howard, Osborne & Best, 2008).

¹ L'expression « manque du mot » sera plusieurs fois utilisée dans ce travail en tant que synonyme de « trouble de l'évocation lexicale ». Il convient toutefois de noter que toute personne présentant un manque du mot n'est pas nécessairement aphasique.

L'évaluation de ce déficit inclut habituellement une tâche de dénomination orale d'images, c'est pourquoi on utilise volontiers le terme de *trouble de la dénomination* lorsque l'on parle des difficultés d'évocation lexicale. Ce type d'épreuve est toutefois très souvent critiqué et jugé comme manquant de validité écologique, bien qu'il soit couramment utilisé par les cliniciens et les chercheurs (Herbert et al., 2008). Avant d'aller plus loin dans les spécificités de cette tâche, il convient de resituer le cadre théorique dans lequel on étudie la production lexicale.

1.2. Approche psycholinguistique et approche cognitive

Seules les deux principales approches qui tentent d'expliquer la dénomination orale et ses déficits, et qui ont guidé ce travail, sont présentées ici : l'approche psycholinguistique et l'approche cognitive. Il en existe bien sûr une troisième, l'approche anatomo-clinique, qui s'est consacrée à rechercher les liens entre des lésions cérébrales et des troubles du langage. Etant hors du cadre du présent travail, elle ne sera toutefois pas développée.

1.2.1. Approche psycholinguistique

L'approche psycholinguistique étudie les troubles de l'évocation lexicale en dénomination en se référant à la notion de lexique mental, qui est postulé à titre d'hypothèse. Selon ce courant théorique, le lexique mental est conçu comme « un répertoire mental regroupant, sous la forme d'entrées verbales unitaires et, par référence à une langue donnée, des spécifications de forme (structure phonémique, graphémique), de fonction (appartenance à une ou à plusieurs catégories grammaticales) et de sens (une ou plusieurs significations potentielles) » (Kremin, 1994, p.376). Bien que son objectif principal soit de cerner les principes d'organisation du lexique ainsi que les processus qui contrôlent son fonctionnement, l'approche psycholinguistique tente également de rendre compte des facteurs influençant la production langagière. Sont

ainsi étudiées les caractéristiques propres des sujets (âge, sexe ou encore niveau d'éducation), ainsi que les caractéristiques linguistiques et extra-linguistiques des stimuli présentés. La psycholinguistique a certes permis d'identifier de nombreux facteurs intervenant dans les tâches de dénomination, mais elle ne permet pas de cerner les mécanismes impliqués et perturbant la production langagière (Kremin, 1994, 2002).

1.2.2. Approche cognitive ou du traitement de l'information

Il existe un autre courant théorique qui, à l'aide de modèles, s'intéresse justement à identifier les mécanismes déficitaires dans les troubles de l'évocation lexicale. C'est l'approche de la neuropsychologie cognitive ou du traitement de l'information. Selon les cognitivistes, les processus de la pensée, autrement dit le système cognitif, seraient constitués de différents sous-systèmes de traitement de l'information (ou modules). Ces sous-systèmes fonctionneraient de manière relativement autonome et indépendante pour assurer des fonctions spécifiques. La notion de *modularité* est ainsi centrale dans cette approche. En ce qui concerne le langage, on considère qu'un modèle de traitement de l'information linguistique est conçu comme une architecture fonctionnelle cognitive. Celle-ci serait organisée en un assemblage hiérarchisé de boîtes interconnectées entre elles par des flèches, symbolisant respectivement les sous-systèmes de traitement de l'information (ou modules) et les voies de communications entre ceux-ci (Estienne & Piérart, 2006 ; Kremin, 2002).

Dans le cadre d'un bilan logopédique en neuropsychologie, on procède dans un premier temps de façon descriptive à l'aide de tests standardisés. En effet, ces derniers permettent de mettre en évidence les déficiences langagières ainsi que d'évaluer leur sévérité en se référant à des données normatives selon l'âge, le sexe et le niveau d'éducation, lorsqu'il est établi que ces facteurs ont une influence sur les performances au test utilisé. Dans un deuxième temps, on se base sur une démarche interprétative. Celle-ci permet de confronter entre eux les différents résultats recueillis aux tests de langage, afin de formuler des hypothèses sur le niveau cognitif perturbé (Mazaux, Dehail, Davet, Pradat-Diehl & Brun, 2007). En effet, l'approche cognitive s'intéresse à localiser les lésions dites fonctionnelles au cours du traitement de l'information

(Kremin, 1994). De plus, conformément au postulat de modularité, les différents sous-systèmes de traitement peuvent être sélectivement touchés. Il est ainsi possible d'identifier par déduction le composant déficitaire et responsable de l'altération d'une fonction particulière. Dans le cadre d'un bilan, le clinicien s'appuie donc sur une démarche interprétative et déductive (Estienne & Piérart, 2006 ; Mazaux et al., 2007 ; Tran, 2007) en se basant sur trois types d'observations : l'association de troubles, la double dissociation et le pattern des erreurs (Seron, 2004).

Lors de l'observation d'une *association de troubles ou co-occurrence de troubles*, on postule qu'un même sous-composant de traitement de l'information peut intervenir dans différentes tâches. En effet, la dénomination orale d'images mobilise des composants qui lui sont propres mais qui peuvent également être en partie sollicités dans d'autres épreuves linguistiques. Il s'agit donc d'évaluer les performances d'un patient à l'aide de plusieurs tâches impliquant les mêmes composants que ceux qui seraient en jeu dans la fonction déficitaire. Lorsqu'un patient présente des difficultés d'importance égale dans deux ou plusieurs tâches différentes, il est légitime de penser qu'il existe une altération d'une composante commune de traitement (Estienne & Piérart, 2006 ; Seron, 2004 ; Tran, 2007).

Un deuxième type de preuve fréquemment exploité est l'observation de la *double dissociation*. Lorsque l'on postule l'existence de deux systèmes distincts et indépendants de traitement de l'information (A et B), on peut s'attendre à observer, lors d'une épreuve X, des patients chez qui le système A est lésé tandis que le B est intact, par exemple. Il est toutefois nécessaire de trouver un autre patient qui présente un tableau inverse de performances, à savoir un système A intact tandis que le système B est lésé lors d'une tâche Y également. En effet, si l'on ne peut trouver le tableau inverse, les différences observées pourraient être expliquées par le degré de complexité des épreuves administrées, l'une étant plus difficile que l'autre. Pour être valide, l'utilisation du paradigme de la double dissociation nécessite ainsi toute une série de contrôles méthodologiques (Mazaux et al., 2007 ; Seron, 2004).

Le troisième élément qui peut être apporté dans le travail d'interprétation est l'examen détaillé du *pattern d'erreurs*, et ce dans les diverses tâches soumises au patient. Il s'agit de montrer que les erreurs dans les différentes épreuves présentent les

mêmes caractéristiques et sont sensibles aux mêmes facteurs psycholinguistiques contrôlés dans le matériel présenté (Seron, 2004).

Enfin, outre la mise en évidence des atteintes des modules, les modèles cognitivistes ont l'intérêt d'identifier les composants du système de traitement de l'information qui sont préservés. Cette vision plus « positive » des fonctions langagières est importante, car elle donne des pistes sur les moyens à exploiter pour une éventuelle prise en charge thérapeutique (Tran, 2007).

➤ **Modèle cognitif de la dénomination orale d'images**

Comme il l'a déjà été évoqué, lors d'un bilan neuropsychologique, il est habituel d'évaluer les performances d'un patient à l'aide de la tâche dite de *dénomination orale d'images*. Elle permet en effet d'apprécier la présence ainsi que l'importance d'un éventuel trouble de l'évocation lexicale, ceci à partir d'un matériel imagé (Tran, 1997).

Si l'on se réfère à la neuropsychologie cognitive, la plupart des modèles actuels s'accordent à distinguer trois niveaux centraux dans l'activité de dénomination. Le premier comprend l'analyse perceptive de l'image, qui donne accès au second niveau contenant les informations sémantiques stockées en mémoire. Enfin, le troisième niveau est celui du lexique phonologique, qui conduit à la sélection du nom correspondant à l'image et aux processus périphériques de production verbale. La tâche de dénomination orale d'images impliquerait les mêmes composants dans le traitement de l'information (voir Figure 1) que la parole spontanée, à savoir : *le système sémantique* (B), *le lexique phonologique de sortie* (C) et *la mémoire tampon phonologique* (D). Elle requiert en outre un composant supplémentaire qui est celui de *la reconnaissance visuelle* et donc perceptive des images (A) (Deloche & Hannequin, 1997 ; Estienne & Piérart, 2006 ; Kremin, 2002).

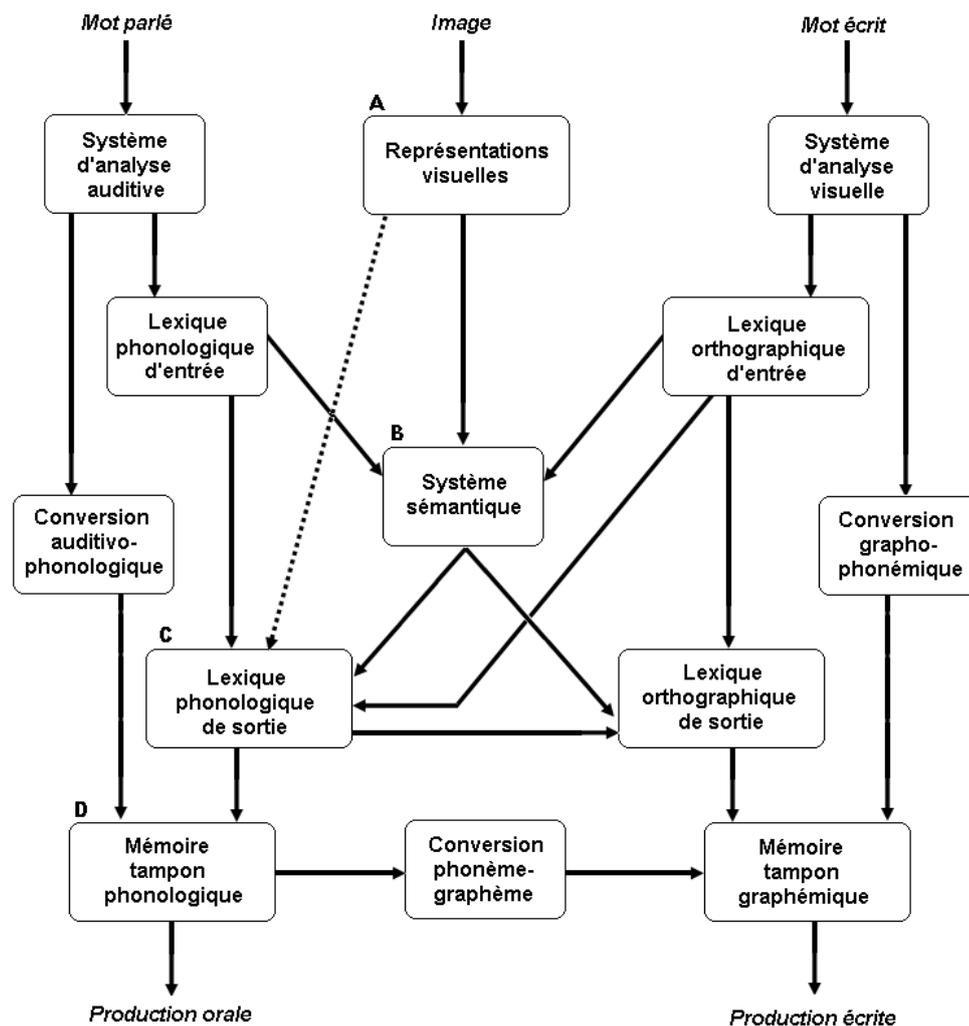


Figure 1 : Modèle cognitif de traitement des mots.

Les composants habituellement considérés pour la dénomination orale d'images sont les boîtes A, B, C et D. En pointillé, représentation de la connexion non sémantique directe entre l'input visuel et l'étiquette verbale, autrement dit la voie visuo-phonologique directe (d'après Kay, Lesser & Coltheart, 1992 ; cités par Estienne & Piérart, 2006).

Il n'existe actuellement aucun accord en ce qui concerne l'activation des processus de traitement de l'information. Certains modèles proposent que les différentes étapes de la dénomination orale d'images soient activées de manière séquentielle (Morton, 1985 ; cité par Kremin, 2002), tandis que d'autres postulent une activation en cascade (Humphreys, Riddoch & Quinlan, 1998 ; cités par Kremin, 2002). Rares sont les auteurs qui ont évoqué une possible connexion non sémantique directe (voir Figure 1, flèche en pointillé) entre l'input visuel et la production orale (Ratcliff & Newcombe, 1982 ; Warren & Morton, 1982 ; Kremin, 1986 ; tous cités par Kremin, 2002).

L'existence de cette voie visuo-phonologique directe (ou non sémantique directe) est en effet compatible avec l'observation de patterns particuliers de performances de certains patients, chez qui la dénomination orale est préservée, malgré la présence d'un trouble majeur de compréhension sémantique (Brennen, David, Fluchaire & Pellat, 1996 ; Kremin, 1986 ; Kremin, Beauchamp & Perrier, 1994 ; Kremin & De Agostini, 1995 ; Shuren, Geldmacher & Heilman, 1993 ; tous cités par Kremin, 2002).

1.3. Troubles spécifiques en dénomination orale d'images

Une fois qu'un trouble de dénomination est mis en évidence, l'intégrité de chacun des composants impliqués dans cette tâche doit être évaluée. Comme mentionné, selon la neuropsychologie cognitive, chaque niveau de traitement peut être sélectivement touché et induire des troubles d'évocation lexicale. Il convient donc d'identifier au mieux le composant responsable du déficit (Estienne & Piérart, 2006).

Les composants évalués en premier lieu sont ceux impliqués dans la *reconnaissance visuelle des images*. Un déficit dans cette étape de traitement de l'information peut conduire à des difficultés de dénomination, les images n'étant d'emblée pas reconnues. Cette évaluation n'est pas directement langagière et s'effectue par exemple à l'aide de la *Birmingham Object Recognition Battery (BORB)*, (Riddoch & Humphreys, 1993 ; cités par Estienne & Piérart, 2006), qui est une batterie de tests évaluant différents composants cognitifs impliqués dans la reconnaissance visuelle d'images et d'objets. En théorie, c'est seulement si cette étape de traitement de l'information est intacte que l'on devrait s'intéresser à l'évaluation des autres systèmes. Toutefois, dans la clinique quotidienne, elle n'est pas systématiquement évaluée. On l'analyse de façon plus approfondie en cas de suspicion de troubles à ce niveau (grand nombre d'erreurs visuelles en dénomination, par exemple). On s'intéresse ensuite à l'évaluation de l'intégrité du système sémantique et du lexique phonologique à l'aide de la dénomination orale d'images plus spécifiquement (voir Annexe V pour une présentation critique de quelques tests de dénomination orale en français). Enfin, il convient également d'examiner la mémoire tampon phonologique et la conversion auditivo-phonologique (Estienne & Piérart, 2006).

Si l'on considère à présent les troubles de l'évocation lexicale à travers le regard de la psycholinguistique, on peut également mettre en évidence des déficits spécifiques. Ces observations s'inscrivent dans l'idée d'une organisation catégorielle du lexique. Cette hypothèse postule que « la pathologie affecterait différemment diverses catégories de stimuli » (Kremin, 2002, p.78). Dans le cadre de la dénomination orale d'images, il existe ainsi des déficits pouvant toucher spécifiquement des catégories sémantiques particulières ou certaines classes grammaticales des mots (Kremin, 1994).

En ce qui concerne l'organisation du lexique en *catégories sémantiques*, on peut observer des dissociations au sein de champs sémantiques précis, tels que les parties du corps, les lettres, les couleurs ou encore les nombres. A titre d'exemple, Dennis (1976 ; cité par Kremin, 2002) a observé un patient qui présentait des difficultés disproportionnées pour la dénomination de parties du corps, tandis que c'était la seule catégorie préservée chez deux autres patients (Denes, Della Barba, Cipolotti & Semenza, 1990 ; Shelton, Fouch & Caramazza, 1998 ; tous cités par Kremin, 2002). Une autre dissociation souvent rapportée concerne les difficultés pour la dénomination d'objets « animés » (plantes ou animaux par exemple), tandis que celle de la catégorie d'objets « inanimés » ou manufacturés par l'homme serait préservée (De Renzi & Lucelli, 1994 ; Hart, Brendt & Caramazza, 1985 ; tous cités par Kremin, 2002). A noter que la dissociation inverse a aussi été observée (Warrington & McCarthy, 1983 ; Warrington & Shallice, 1984 ; tous cités par Kremin, 2002).

Outre l'appartenance sémantique du stimulus, il existe, comme il l'a été évoqué, des dissociations dues à la *catégorie grammaticale* des mots représentés. Les recherches ont principalement porté sur la dénomination de substantifs *versus* la dénomination de verbes (Kremin, 1994). Les dessins représentant des verbes sont souvent moins bien dénommés que ceux représentant des substantifs (Miceli, Silveri, Villa & Caramazza, 1984 ; cités par Kremin, 2002). Toutefois, chez les patients présentant une aphasie primaire progressive fluente, on rencontre la dissociation inverse, à savoir une dénomination de verbes préservée et un manque du mot portant essentiellement sur les substantifs (David, 2007).

Ainsi, l'accès lexical « n'est pas un processus de tout ou rien » (Kremin, 2002, p.87), ni « une simple activité d'étiquetage d'un objet du monde » (Tran, 1997, p.19). Le succès en dénomination orale d'images peut être influencé par divers facteurs

interagissant entre eux, tant extra-linguistiques (représentation imagées ou consigne de l'épreuve, par exemple) que linguistiques (fréquence lexicale ou âge d'acquisition du mot, par exemple). Ces derniers sont développés dans le chapitre suivant. Il est important de noter que le sujet sain peut également avoir des difficultés de lexicalisation, lui donnant l'impression d'avoir le mot « sur le bout de la langue », sans pour autant être aphasique (Kremin, 1994, 2002). En effet, comme également développé dans ce qui suit, on peut ajouter que les facteurs individuels comme l'âge, le sexe ou le niveau d'éducation peuvent aussi avoir une influence sur les performances des sujets, sains comme cérébrolésés (Metz-Lutz et al., 1991 ; Tran, 1997).

1.4. Facteurs influençant les performances en dénomination orale d'images

1.4.1. Facteurs individuels

Comme il vient d'être évoqué, *l'âge, le sexe et le niveau d'éducation* peuvent avoir une influence notable sur les performances des sujets en dénomination orale d'images. Plusieurs auteurs ont examiné les résultats à différents tests en fonction de l'âge et du niveau de scolarité tout particulièrement, le sexe étant moins souvent contrôlé. Pour ne citer qu'un exemple de recherche, Metz-Lutz et al. (1991) se sont intéressés à clarifier le rôle de ces trois paramètres chez des participants adultes normaux ainsi qu'à intégrer l'influence de ces facteurs individuels dans l'élaboration d'une épreuve de dénomination standardisée utilisable dans le cadre de l'examen des patients cérébrolésés. L'analyse de variance effectuée a montré un effet principal des trois facteurs : le pourcentage de réponses dominantes diminue lorsque l'âge augmente, il est également moins élevé pour des sujets avec un niveau de scolarité faible (jusqu'à 9 ans) en comparaison avec ceux ayant eu plus de 9 ans d'études. Enfin les auteurs ont constaté que le score moyen des femmes était plus élevé que celui des hommes.

Toujours dans la même étude, mais d'un point de vue plus qualitatif à présent, les rôles de l'âge et du niveau de scolarité sur la fréquence des erreurs de nature verbale

(« *sandale* » pour « *botte* », par exemple) et perceptive (« *téléphone portable* » pour « *gomme* », par exemple) ont été analysés. Il en découle que les erreurs verbales seraient liées au niveau d'éducation, les sujets ayant été scolarisés moins longtemps en produisant plus. Quant à la fréquence des erreurs perceptives, elle serait liée à l'âge. Ce sont les participants de plus de 60 ans qui en produisent le plus. En conclusion, plusieurs recherches ont confirmé qu'il était important que les effets de ces facteurs individuels soient pris en compte lors d'une évaluation neuropsychologique, pour les tâches de dénomination orale tout particulièrement.

1.4.2. Facteurs linguistiques et extra-linguistiques

Etant donné qu'il n'existe actuellement aucun accord en ce qui concerne les mécanismes déficitaires sous-jacents aux troubles d'évocation lexicale, une des alternatives consiste à étudier les difficultés en dénomination orale à travers les facteurs susceptibles d'influencer la probabilité de réussite à cette tâche (Kremin, 2002).

Une multitude de variables linguistiques et extra-linguistiques, correspondant aux propriétés de l'étiquette verbale, de l'image, ou encore du concept représenté, a été étudiée dans différentes langues par le biais de la dénomination orale (Laganaro, Di Pietro & Schnider, 2006). Les modules du système de traitement de l'information sont par ailleurs diversement sensibles à ces diverses propriétés, que ce soit chez le sujet sain ou cérébrolésé. Par conséquent, l'observation des effets de ces variables peut amener des informations sur la nature du sous-composant altéré (voir Figure 2 pour les descriptions qui suivent). Alario, Ferrand, Laganaro, New, Frauenfelder & Segui (2004) ont ainsi cherché à identifier les effets de neuf différentes variables sur le temps de latence en dénomination orale. Ils conclurent que *degré de complexité visuelle* et le *degré de consensus de l'image* à dénommer sont liés aux propriétés du dessin. Ils ont un effet sur la vitesse de dénomination de l'image au niveau des processus de reconnaissance visuelle. Le *degré de familiarité du concept* à dénommer ainsi que son *imageabilité* sont également connus pour influencer les latences dans l'activation des représentations sémantiques. En ce qui concerne le *degré de consensus sur l'étiquette verbale*, la *fréquence d'usage* ou encore *l'âge d'acquisition des mots-cible*, ces

paramètres auraient un effet sur le temps de latence au niveau des processus d'accès lexical. Enfin, les facteurs phonologiques tels que le *nombre de phonèmes ou de syllabes* pourraient influencer l'encodage phonologique (Alario et al., 2004, Laganaro, et al., 2006).

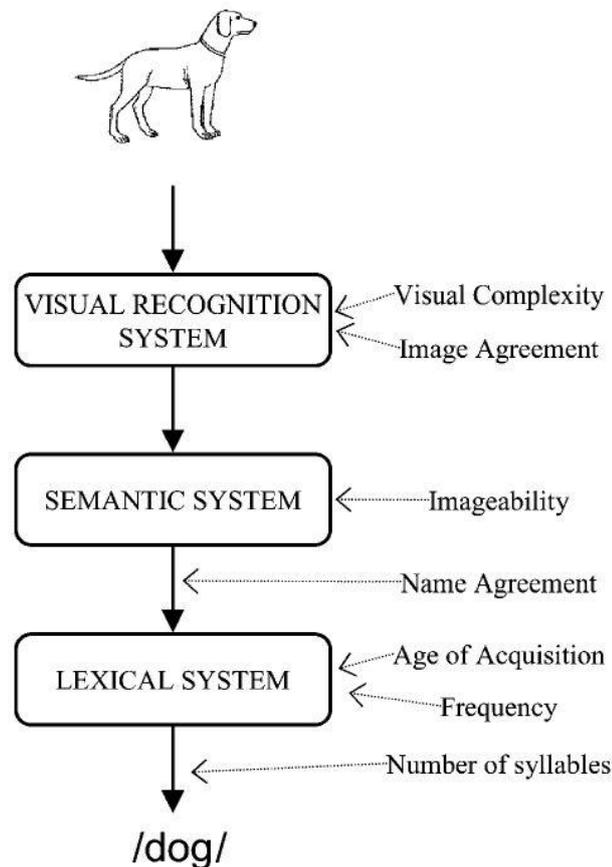


Figure 2 : Modèle de dénomination orale d'images.
Localisation des effets de diverses variables étudiées dans la recherche d'Alario et al. (2004).

Certains de ces paramètres (fréquence d'usage, longueur du mot, etc.) sont ainsi classiquement pris en compte dans les épreuves de bilans de langage (Deloche & Hannequin, 1997 ; Estienne & Piérart, 2006). En effet, au vu de la diversité des variables entrant en jeu dans l'évocation lexicale, les tests de dénomination orale d'images devraient contenir des items de différents types afin de mettre en évidence les possibles dissociations évoquées précédemment (Laiacina, Luzzatti, Zonca, Cuarnaschelli & Capitani, 2001). Toutefois, dans leur étude, Laiacina et al. (2001) sont arrivés à la conclusion que les paramètres visuels n'avaient que peu d'influence sur les performances, alors que les variables d'ordre linguistique (tant lexicales que

sémantiques) influenceraient de façon significative les performances des participants à l'épreuve de dénomination proposée.

Revenons aux facteurs linguistiques pouvant influencer la dénomination orale d'images. Alors que certaines propriétés telles que la fréquence lexicale² étaient considérées comme déterminantes depuis longtemps, d'autres étaient jusque-là plutôt négligées. C'est le cas notamment de l'âge d'acquisition des mots et du degré de consensus en dénomination (Kremin, 2002).

L'âge d'acquisition correspond en moyenne à l'âge auquel un mot a été appris pour la première fois. Les données utilisées par les chercheurs sont des estimations faites par des sujets adultes selon une échelle, ses échelons désignant des tranches d'âge. Elles peuvent sembler peu fiables en apparence, mais ont été validées par un certain nombre de travaux. En effet, il existe une forte corrélation entre l'âge réel d'acquisition des mots et son estimation, par des sujets adultes, à l'aide de cette même échelle (Morrison, Chappell & Ellis, 1997 ; cités par Kremin, 2002). Il semblerait ainsi que les mots appris tôt dans la vie soient non seulement dénommés plus rapidement que les mots appris plus tardivement, mais aussi qu'ils résisteraient mieux aux effets des lésions cérébrales (Morrison, Ellis & Quinlan, 1992).

Le degré de consensus en dénomination orale est une autre variable qui est de plus en plus étudiée et dont le rôle ne devrait pas être négligé. Il correspond à la proportion dans laquelle des sujets témoins s'accordent sur une réponse dominante en dénomination d'images (Kremin, 2002). Dans son étude, Lachman (1973, cité par Kremin, 2002, ainsi que par Morrison et al., 1992) a observé chez le sujet normal une différence dans les temps de réaction entre les mots de faible et de fort degré de consensus. Les images pour lesquelles il n'y a qu'une réponse dominante (fort degré de consensus) sont dénommées plus rapidement que celles ayant des réponses multiples (faible degré de consensus). Cet effet est robuste et indépendant des autres facteurs, n'étant ainsi imputable ni à la fréquence, ni à l'âge d'acquisition des mots.

² Il est à noter que les fréquences lexicales sont souvent estimées à partir du langage écrit et qu'elles sont toujours dépendantes des corpus à partir desquels leurs valeurs ont été calculées (textes littéraires, manuels scolaires, etc.)

L'âge d'acquisition des mots et le degré de consensus en dénomination ont donc fait l'objet d'études plus ou moins récentes et s'avèrent être aujourd'hui des variables pertinentes dans l'étude des troubles de l'évocation lexicale. De ces recherches il ressort également que l'effet de la fréquence des mots sur la dénomination est aujourd'hui remis en cause, alors qu'il était bien établi autrefois. En effet, les facteurs fortement corrélés à la fréquence, comme l'âge d'acquisition, ont été pris en compte dans l'explication des effets observés. Ainsi, ce qui était interprété autrefois comme un effet de fréquence ne serait en réalité qu'un artefact lié à l'influence d'autres facteurs, tout particulièrement l'âge d'acquisition des mots (Kremin, 2002 ; Sirois, Kremin & Cohen, 2006).

1.5. Intérêts et limites des épreuves de dénomination orale d'images

1.5.1. Intérêts

La dénomination orale d'images est une épreuve incontournable lors du bilan de langage dans la clinique aphasiologique. En effet, les tests de dénomination ont plusieurs caractéristiques positives lorsqu'ils sont bien construits. D'un point de vue méthodologique, c'est un type d'épreuve relativement simple à administrer et à coter étant donné que les cibles sont connues, contrairement aux situations de discours spontané. En effet, dans des situations d'évaluation d'un patient avec un langage dont l'expression est jargonnée, son intention communicative peut parfois s'avérer difficile à identifier. La connaissance des mots-cible permet également de contrôler un grand nombre de paramètres susceptibles d'influencer les performances à cette épreuve (Tran, 1997). En ce qui concerne la cotation, une épreuve de dénomination orale fournit un score global permettant d'avoir un indice de sévérité du déficit d'accès lexical. Outre cet aspect quantitatif, elle permet une analyse qualitative des erreurs, que l'on peut mettre en lien avec certaines variables linguistiques connues pour leur influence sur les performances en dénomination (Ellis, Lum & Lambon-Ralph, 1996 ; Nickels & Howard, 1994 ; tous cités par Herbert et al., 2008). Les tests de dénomination ont

également l'avantage d'avoir une bonne fiabilité ainsi que des fortes corrélations test-retest (Howard, Patterson, Franklin, Morton & Orchard-Lisle, 1984 ; Swinburn, Howard & Porter, 2004 ; tous cités par Herbert et al., 2008). Ils représentent ainsi des situations aisément reproductibles et contrôlées.

1.5.2. Limites

Les épreuves de dénomination orale d'images présentent un certain nombre de contraintes méthodologiques quant au choix des items : les référents à dénommer doivent être imageables, pouvoir être facilement représentés, être aisément reconnus par les sujets et ne pas donner lieu à un trop grand nombre de productions différentes (Tran, 1997). On peut encore ajouter qu'il convient qu'elles soient étalonnées dans la langue, voire la région linguistique étudiée, car il existe de nombreux items dépendants de la langue et de la culture (Chainay, Rosenthal, & Goldblum, 1998). Avant le travail de Chainay et al. (1998), les items étaient très souvent choisis dans la batterie de Snodgrass et Vanderwart (1980), étalonnée en anglais, ce qui est aujourd'hui reconnu comme un biais évident quant à l'usage de ces normes pour des recherches ou tests en français. Ce type de tâche a également l'inconvénient d'induire la production de mots isolés et plus particulièrement de substantifs concrets. Il semble d'ailleurs que les personnes aphasiques soient meilleures pour récupérer les mots d'objets dans une situation de production d'énoncés que dans une simple dénomination d'images. La syntaxe aurait ainsi un effet facilitateur lorsqu'il s'agit de produire un nom dans le contexte d'un énoncé (Tran, 1997 ; Zingeser & Berndt, 1988 ; cités par Herbert et al., 2008).

D'un point de vue écologique, les gens ne communiquent pas les uns avec les autres en dénommant des items, ils n'ont pas non plus toujours des images devant eux lorsqu'il s'agit de s'exprimer. En effet, contrairement au discours spontané qui constitue une activité dynamique, la dénomination est quant à elle une activité d'évocation ponctuelle et déterminée par un stimulus non linguistique, à savoir l'image (Tran, 1997). La plus forte critique concernant les tests de dénomination d'images est qu'ils n'offrent pas une évaluation valide du fonctionnement de l'accès lexical dans la

vie quotidienne et manqueraient donc de validité écologique. De plus, tant les sujets que les examinateurs seraient conscients qu'il existe une réponse attendue, contrairement aux situations naturelles de conversation, cela pouvant biaiser la représentation du fonctionnement naturel de l'accès lexical (Helm-Estabrooks, 1997 ; Holland, 1994 ; tous cités par Herbert et al., 2008).

1.5.3. Comparaison avec les situations de conversation

Au vu des critiques sur les épreuves de dénomination, on peut penser que les situations de conversation représentent le moyen idéal pour mesurer les difficultés d'accès lexical dans la vie quotidienne. Toutefois, si on la compare aux tâches de dénomination, la conversation est une activité langagière relativement peu contrôlée (Herbert et al., 2008). Les performances peuvent varier selon le thème de la discussion, ainsi que selon les partenaires impliqués (Perkins, 1995, cité par Herbert et al., 2008). D'un point de vue quantitatif, ceci a pour conséquence une faible fiabilité test-retest. Un second obstacle réside dans le codage des comportements par les examinateurs qui est inévitablement entaché d'une certaine subjectivité et nécessite donc un contrôle rigoureux inter- et intra-examineur. Cet ensemble de difficultés explique sans doute pourquoi peu d'études se sont intéressées à l'aspect quantitatif de l'accès lexical dans des situations plus écologiques, telles que la conversation (Herbert et al., 2008).

Etant donné la remise en cause de la validité des épreuves de dénomination, l'étude d'Herbert et al. (2008) s'est justement intéressée au développement d'une méthode de quantification de l'accès lexical dans des situations de conversation avec des personnes aphasiques. Les performances aux tâches de dénomination ont été comparées à celles d'une situation de conversation, mesurée avec la méthode développée. Les résultats de cette étude indiquent que l'accès lexical est quantitativement similaire dans les deux situations. En effet, les mêmes processus sont en jeu, que ce soit en situation de dénomination ou de conversation. Autrement dit, une personne qui récupère facilement les mots dans une tâche de dénomination en fera de même lors d'une conversation.

En conclusion, cette étude (Herbert et al., 2008) permet de valider l'utilisation des tests de dénomination orale d'images, non seulement en tant que mesure de la sévérité

du trouble, mais aussi comme mesure de la fonctionnalité de l'accès lexical. Bien qu'il existe des différences entre les situations de discours spontané et les épreuves de dénomination, ces dernières peuvent contribuer à mettre en évidence, dans un certain nombre de cas, des troubles de l'évocation lexicale peu perceptibles dans le langage spontané. En effet, les patients peuvent avoir recours à des stratégies de compensation ou à des conduites palliatives qui leur permettent de réduire l'impact de leur manque du mot sur leurs performances communicatives (Tran, 1997). L'usage des épreuves de dénomination dans la clinique reste donc justifié. Ce type de test permet d'identifier la nature et la sévérité du trouble d'accès lexical, ainsi que l'impact potentiel du handicap dans la participation à des conversations dans la vie quotidienne (Herbert et al., 2008, Tran, 1997).

1.6. Tests de dénomination orale en français : problèmes actuels

Il existe plusieurs épreuves en français permettant d'évaluer la dénomination orale d'images³. Certaines sont largement utilisées bien qu'elles ne soient pas toujours satisfaisantes pour une utilisation clinique (pour une description de certaines épreuves couramment utilisées, le lecteur peut consulter l'Annexe V). Le défi actuel réside dans la construction d'épreuves courtes, contrôlées et utilisables dans des situations d'évaluation clinique. En effet, lorsqu'il s'agit de choisir un test de dénomination pour faire un bilan de langage, on se heurte à un certain nombre de contraintes.

Premièrement, il existe une grande diversité de variables psycholinguistiques et individuelles pouvant influencer les performances en dénomination orale, sans que les différents auteurs les ayant étudiées ne parviennent à un consensus. Il s'agit donc de pouvoir sélectionner les variables qui seraient les meilleurs prédicteurs de réussite à cette tâche. Comme il l'a été mentionné plus haut, il a été démontré que les paramètres d'âge d'acquisition et de degré de consensus ont une influence dans le succès en dénomination (Kremin, 2002), or ils ne sont jamais contrôlés conjointement dans les tests existants.

³ Les tests les plus fréquemment utilisés en Suisse romande sont le *Boston Naming Test* (Kaplan, Goodglass & Weintraub, 1983), le *Protocole Montréal-Toulouse* (Nespoulous et al., 1992) ainsi que la version française du *Boston Diagnostic Aphasia Examination* (Mazaux & Orgogozo, 1981).

Deuxièmement, les épreuves (ou sous-épreuves) de dénomination orale ont été étalonnées pour la plupart sur des populations anglophones. Dans le cas des tests en français, nous n'avons actuellement pas d'étalonnage récent pour la Suisse romande. Ils sont pour la plupart normés en France ou en Belgique, ce qui rend parfois la cotation des réponses difficile, certains items n'étant pas dénommés de la même façon dans l'ensemble de la francophonie. En effet, la relation entre les images et leurs noms n'est hélas pas univoque. Il s'agit donc de combler l'absence de normes pour les sujets francophones et, en ce qui concerne le présent travail, pour la Suisse romande tout particulièrement.

Troisièmement, les épreuves bien contrôlées sont souvent trop longues ou construites pour la recherche en psycholinguistique. C'est le cas par exemple de la *Batterie d'examen des troubles en dénomination (ExaDé)*, Bachy-Langedock, 1989) ou des *Normes de dénomination de 315 images en noir et blanc* (Chainay, Rosenthal, & Goldblum, 1998). On trouve encore des bases de données en français, constituées de dessins standardisés pour un grand nombre de variables (degré de consensus en dénomination, âge d'acquisition des mots, complexité visuelle, familiarité avec le concept à dénommer, etc.) telles que les normes d'Alario et Ferrand (1999) ou, plus récemment, de Bonin et al. (2003). A l'autre extrême, la *Batterie Rapide de Dénomination (BARD)*, Croisile, 2005) ne comporte que 10 items. Ce test est supposé pouvoir diagnostiquer une anomie en 15 secondes.

1.7. Le Déno-100 : un test de dénomination contrôlé

1.7.1. Banque de données pour la dénomination orale d'images et variables cruciales dans l'accès au lexique

Comme il l'a déjà été évoqué plus haut, les approches récentes cherchent à aborder le trouble de l'évocation lexicale à travers les facteurs susceptibles d'influencer la réussite aux épreuves de dénomination orale d'images. Or, beaucoup de ces facteurs sont culturellement dépendants et étudiés la plupart du temps chez des sujets

anglophones. Afin de combler l'absence de normes pour les sujets francophones, Kremin et al. (2003) ont participé à la création d'une banque de données dans plusieurs langues pour la dénomination orale d'images : le *Protocole Européen de Dénomination Orale d'Images* (PEDOI).

Cette banque de données est composée de 391 dessins au trait en noir et blanc. Les catégories sémantiques sont variées et la prototypicalité des dessins a été contrôlée. Ces items ont été normalisés auprès de 1040 locuteurs interrogés dans leur langue maternelle (allemand, anglais, espagnol, français, hollandais, italien, suédois et russe). En ce qui concerne le français, les dessins ont été soumis à dénomination auprès de 120 sujets francophones avec un contrôle des variables interindividuelles que sont l'âge (18-39 ans, 40-59 ans, 60-75 ans), le sexe (masculin et féminin) et le niveau d'éducation (plus ou moins de 9 ans d'études). De ces 391 images, seules 269 ont été retenues pour le protocole expérimental du *Déno-100* (Kremin et al., 2005). Cinq paramètres ont été établis pour chaque item :

➤ **Degré de consensus en dénomination orale**

Il correspond au nombre de réponses majoritaires données par les participants dans leur langue maternelle. Ce chiffre est ensuite transformé en pourcentage afin de faciliter les comparaisons entre les langues. En ce qui concerne les sujets francophones, le degré moyen de consensus est de 88% (Kremin et al., 2005).

➤ **Degré de complexité visuelle**

Il se réfère à une estimation de la complexité visuelle graphique à l'aide d'une échelle en sept points (1 = très simple à 7 = très complexe). Un item a été jugé complexe si la moyenne obtenue auprès des personnes interrogées était égale ou supérieure à 3,5 (Kremin et al., 2003 ; Kremin et al., 2005).

➤ **Age d'acquisition des mots.**

Il désigne l'âge moyen auquel un mot a été appris pour la première fois durant l'enfance. Les données utilisées sont des mesures estimées à l'aide d'une échelle en sept points. Il a été demandé aux participants d'estimer l'âge auquel ils avaient appris les mots proposés. Chaque échelon représente une tranche d'âge de deux ans (échelon 1 = de 0 à 2 ans à échelon 7 = plus de 12 ans) et couvre ainsi une période allant de 0 à 13

ans environ (Kremin et al., 2005). Bien que ce soit une mesure estimée, elle s'avère fortement corrélée à l'âge réel d'acquisition. En effet, ces données ont été validées par un certain nombre de chercheurs qui ont notamment comparé ces estimations à des normes tirées de tests de vocabulaire pour enfant (Morrison et al., 1997 ; cités par Kremin, 2002).

➤ **Degré de familiarité avec le concept à dénommer**

Il correspond également à une mesure estimée à l'aide d'une échelle en cinq points (1 = très peu familier à 5 = très familier). Le degré de familiarité fait référence à la fréquence avec laquelle les sujets sont confrontés à l'objet (ou au concept) à dénommer ou amenés à y penser dans la vie quotidienne. La difficulté consiste donc à évaluer le concept lui-même et non pas la représentation imagée proposée, cette dernière étant estimée par le degré de complexité visuelle. Les résultats n'ont montré aucune différence significative en ce qui concerne la langue maternelle des sujets. Ces données ont été établies avec des sujets francophones et germanophones (Kremin et al., 2003 ; Kremin et al., 2005).

➤ **Fréquence d'usage des mots**

Les données sur cette variable proviennent des travaux de Imbs (1971 ; cité par Kremin et al., 2005). En ce qui concerne le *PEDOI*, seule la fréquence relative aux corpus de vocabulaire du XX^{ème} siècle a été considérée (Kremin et al., 2005).

Cette banque de données constitue ainsi une source de matériel standardisé et bien contrôlé du point de vue des paramètres propres aux items pertinents dans l'étude de l'évocation lexicale. C'est en vue de la construction d'un test de dénomination plus économique qu'un sous-ensemble de 100 images issues du *PEDOI* a finalement été sélectionné parmi les 269 images retenues pour le protocole expérimental, pour être appelé par la suite *Déno-100* (Kremin et al., 2003).

1.7.2. Caractéristiques du test de dénomination *Déno-100*

Comme mentionné plus haut, le *Déno-100* (Kremin et al., 2005) se veut être un test de dénomination écologique et contrôlé. Dans cette version à 100 items, les variables considérées ont été les suivantes :

- Caractère animé *versus* inanimé des concepts représentés ;
- Degré de complexité visuelle des dessins ;
- Degré de consensus en dénomination ;
- Age moyen d'acquisition des mots ;
- Fréquence des mots ;
- Longueur des mots (en syllabes et en phonèmes) ;
- Degré de familiarité des concepts représentés.

Tableau I : Caractéristiques du *Déno-100*.

Moyennes ainsi que valeurs maximum et minimum des variables considérées dans la sélection des items (d'après Kremin, 2002).

	Caractère animé		Caractère inanimé	
	moyenne par item	min – max	moyenne par item	min – max
degré de complexité visuelle (échelle de 1 à 7)	4,016	1,9 – 5,5	3,43	1,4 – 5,9
degré de consensus en dénomination (sur 120)	107,52	78 – 120	103,92	75 – 120
âge moyen d'acquisition du mot (échelle de 1 à 7)	2,87	1,37 – 4,81	2,97	1,59 – 4,75
fréquence	1974,7	5 – 28913	2001,06	61 – 15464
log F	127,051	117,853 – 141,893	128,717	106,990 – 144,611
longueur du mot (nombre de syllabes)	1,82	1 – 4	1,92	1 – 3
longueur du mot (nombre de phonèmes)	4,48	2 – 9	4,7	2 – 7
degré de familiarité (échelle de 1 à 5)	2,23	1,04 – 4,29	2,91	1,22 – 4,95

Une analyse de corrélation a été faite afin de mesurer l'interdépendance entre les sept variables considérées dans le test *Déno-100*. Les résultats ont montré que ni le degré de consensus en dénomination, ni la fréquence, ni la longueur des mots n'étaient corrélés à d'autres paramètres. L'âge d'acquisition par contre est fortement corrélé avec la familiarité de l'image. Enfin, tant la familiarité du concept à dénommer que l'âge d'acquisition des mots sont corrélés avec la complexité visuelle. Suite à l'investigation expérimentale sur 16 sujets atteints d'une aphasie progressive de type fluent ou pour lesquels un diagnostic de maladie d'Alzheimer probable a été posé, les analyses de régression multiple montrent également des effets significatifs du degré de consensus et de l'âge d'acquisition. Ces deux paramètres seraient donc les meilleurs prédicteurs du succès en dénomination orale (Kremin, 2002). A noter que ces deux variables jouent également un rôle sur la dénomination du sujet sain (Hodgson & Ellis, 1998 ; Kremin, Hamerel, Dordain, De Wilde, & Perrier, 2000 ; tous cités par Kremin, 2002).

Enfin, il serait utile d'intégrer les différents facteurs linguistiques évoqués dans cette épreuve de dénomination orale d'images, et ce avec l'hypothèse générale que les facteurs individuels (âge, sexe et niveau d'éducation) ont également une influence qui va se répercuter sur l'étalonnage du test.

2. Buts de la recherche

Le but de ce travail a été de construire deux versions courtes et parallèles du test de dénomination *Déno-100* (Kremin et al., 2005), ainsi que de les étalonner en Suisse romande sur une population-contrôle adulte de langue maternelle française.

Le seul test de dénomination actuellement étalonné en Suisse romande est une adaptation française du *Boston Naming Test* (*BNT*, Thuillard Colombo & Assal, 1992, pour la version française). La version originale américaine de cette célèbre épreuve (Kaplan, Goodglass & Weintraub, 1983) date de près de 30 ans et ne tient pas compte des variables précédemment citées et pouvant influencer les performances lors d'une épreuve de dénomination orale. En effet, le *Déno-100* (Kremin et al., 2005) présente l'avantage d'être un matériel bien contrôlé et tenant compte de diverses variables, dont celles qui sont considérées aujourd'hui comme cruciales dans l'étude des troubles de l'évocation lexicale, à savoir l'âge d'acquisition des mots et le degré de consensus en dénomination. Toutefois, bien qu'il soit très utile pour la recherche, le *Déno-100* n'est pas adapté à une utilisation clinique en raison de sa longueur (100 items). C'est pourquoi il a été jugé nécessaire par l'équipe de neuropsychologie du CHUV de se baser sur ce matériel contrôlé afin d'en créer deux versions courtes et parallèles. Cela permettrait non seulement d'évaluer plus rapidement l'évocation lexicale chez les patients cérébrolésés, mais aussi d'en avoir une utilisation de type test-retest pour juger de l'efficacité des thérapies ou d'apprécier l'évolution d'un tableau clinique. Enfin en ce qui concerne l'étalonnage, le but a été d'obtenir des normes en fonction de différentes tranches d'âge, tout en tenant compte des effets possibles du sexe et du niveau d'éducation.

3. Matériel et méthode

3.1. Participants

La population de ce travail est constituée de 46 sujets-contrôle adultes de langue maternelle française, recrutés par l'entremise de tiers personnels. L'échantillon est composé de 19 hommes et 27 femmes répartis en 16 sous-groupes selon l'âge (50-59 ans ; 60-69 ans, 70-79 ans ; 80 ans et plus), le sexe (masculin ; féminin) et le niveau d'éducation (niveau 1 – jusqu'à 9 ans de scolarité ; niveau 2 – 9 ans de scolarité ou plus). Le nombre de sujets par groupe n'est pas équivalent en raison des difficultés de recrutement, notamment en ce qui concerne le niveau 1 d'éducation. Ont été exclues de la recherche les personnes ayant un trouble neuropsychologique ou psychiatrique avéré, ayant subi un accident vasculaire cérébral, vivant dans un établissement médico-social ainsi que n'étant pas de langue maternelle française. Les cas de participants bilingues ont été acceptés pour autant que le français ait été appris dans l'enfance. Le Tableau II présente les détails concernant la répartition des sujets dans les seize différents groupes.

Tableau II : Répartition des 46 sujets-contrôle.
Répartition des sujets en fonction de l'âge, du sexe et du niveau d'éducation. Dans chaque case est représenté le nombre de participants par condition expérimentale

		Hommes	Femmes	Total
50-59 ans (moyenne = 55;6 ans)	Niveau 1	0	0	0
	Niveau 2	5	5	10
60-69 ans (moyenne = 65;0 ans)	Niveau 1	0	3	3
	Niveau 2	5	4	9
70-79 ans (moyenne = 75;1 ans)	Niveau 1	2	4	6
	Niveau 2	5	4	9
80 ans et + (moyenne = 86;3 ans)	Niveau 1	0	6	6
	Niveau 2	2	1	3
Total		19	27	46

3.2. Matériel

Deux sets de 24 dessins au trait en noir et blanc représentant des substantifs animés et inanimés ont été présentés aux participants sous forme de test de dénomination orale (voir Annexe II B pour le cahier de passation). Ils constituent ainsi deux versions courtes et parallèles (version A et version B) du test *Déno-100* (Kremin et al., 2005). Pour rappel, les stimuli du *Déno-100* sont tirés de la base de données du *PEDOI* (Kremin et al., 2003).

Les items ont été sélectionnés en fonction de l'âge moyen d'acquisition des mots (AMA) et du degré de consensus en dénomination orale (CONS). Un premier tri a permis de constituer deux ensembles d'items en fonction de l'âge moyen d'acquisition (AMA précoce *versus* AMA tardif). Les items avec les degrés de consensus et les fréquences les plus faibles, ainsi que ceux représentant des mots-composés ont été éliminés. Enfin, les 24 images ont été choisies de façon à ce que chaque version contienne un nombre égal d'items avec un AMA précoce (12 items) et un AMA tardif (12 items). Les stimuli ont également été pairés en fonction de la fréquence et du caractère animé ou inanimé du concept représenté afin que chaque version contienne des items jugés équivalents d'un point de vue de ces paramètres. Les items de chacune des deux versions ont été ordonnés de façon aléatoire. Les analyses statistiques (voir Annexe I C), avec un seuil de significativité α de 0.05, confirment que les deux versions constituées sont parallèles. Il n'y a pas de différences significatives entre les versions pour les deux principales variables considérées que sont l'âge moyen d'acquisition des mots ($t_{0,975}[46]=2.013, p=.091$) et le degré de consensus en dénomination ($t_{0,975}[46]=2.013, p=.087$), ni pour la fréquence d'usage ($t_{0,975}[46]=2.013, p=.225$).

3.3. Procédure

3.3.1. Conditions de passation

Les deux versions courtes et parallèles du test de dénomination *Déno-100* (Kremin et al., 2005) ont été administrées aux 46 sujets de la recherche. Il a été demandé aux participants de produire le nom des images présentées une à une sur support papier (voir Annexe II A pour la consigne), sans contrainte explicite de temps. Avant de commencer, un item d'exemple leur a été donné. Chaque sujet a passé les deux versions courtes à la suite. Afin d'éviter un éventuel biais lié à l'ordre des épreuves ou à la familiarité de la tâche, la passation des deux versions (dans l'ordre A-B ou B-A) a été attribuée de façon aléatoire aux participants. La répartition finale a été bien équilibrée avec 24 participants ayant passé les épreuves dans l'ordre A-B et 22 dans l'ordre inverse.

3.3.2. Mesures et cotation des épreuves

En ce qui concerne les mesures, chacune des épreuves a été chronométrée et toute réponse ne correspondant pas au mot cible de la liste a été relevé dans le cahier de passation (voir Annexe II B). Un inventaire des réponses fournies (RD, RND+, RND), ainsi que du mode de réponse aboutissant à la cible (par conduite d'approche ou autocorrection, par réponse multiple, par extension de la racine lexicale, etc.) est fourni dans l'Annexe IV C et G. Une échelle en trois points (2 points, 1 point ou 0 point) a été utilisée pour la cotation de chaque item :

➤ **2 points par item**

La note maximum de 2 points a été attribuée pour toutes les réponses correspondant à la cible, appelées *réponses dominantes* (RD). De même, 2 points ont été accordés pour les réponses contenant la racine lexicale de la réponse dominante ou étant une

expansion de celle-ci (par exemple pour « *chaîne* », les réponses « *chaîne de montre* » ou « *chaînette* » ont été acceptées et ont reçu 2 points). Dans le cas de réponses multiples (« *cheval ou poney* » pour « *cheval* » par exemple), les 2 points ont été accordés si la réponse dominante a été fournie. On relève néanmoins les autres productions afin de pouvoir analyser la cause de l'hésitation (erreur visuelle, sémantique, visuo-sémantique, etc.). Un taux de réussite de 100% à chacune des versions du test correspond ainsi à 48 points maximum.

➤ **1 point par item**

Il existe un certain nombre d'items pour lesquels des réponses autres que la cible attendue sont souvent produites. Certaines ont été acceptées, ce sont les *réponses non dominantes acceptées* (RND+), et donnent lieu à 1 point (par exemple « *lettre* » pour « *enveloppe* »). Dans la version initiale à 100 items, l'acceptation des termes génériques comme RND+ n'est pas systématique et n'a pas été maintenue pour les versions courtes et parallèles de ce présent travail. De plus, au vu des réponses recueillies pour cet étalonnage, des réponses considérées acceptables au vu des variations régionales possibles entre pays francophones (« *attaches à souliers* » pour « *lacets* » par exemple) et de la fréquence avec laquelle elles ont été données (« *lettre* » pour « *enveloppe* » par exemple, donnée par 16% des participants environ) ont été ajoutées.

➤ **0 point par item**

Enfin, on ne donne aucun point aux réponses erronées ou aux absences de réponse. Elles sont d'ailleurs considérées comme des *réponses non dominantes* (RND-).

3.4. Hypothèses opérationnelles

Avant d'analyser les effets spécifiques des variables individuelles, il convient tout d'abord de vérifier que les deux versions courtes constituées sont *parallèles*. En effet, les analyses statistiques effectuées au moment de la sélection des items n'ont pas montré de différences significatives entre les deux versions pour les variables considérées. Les versions A et B du test *Déno-100* sont donc *a priori* comparables. Il

s'agit donc de vérifier que cette caractéristique se maintient avec les passations des épreuves courtes.

Le but étant d'obtenir des données normatives pour les deux versions courtes et parallèles du test de dénomination *Déno-100*, il s'agit de déterminer si les variables individuelles ont une influence sur les performances aux deux épreuves. Les principales hypothèses et observations portent ainsi sur les effets de l'âge, du sexe et du niveau d'éducation. Les hypothèses suivantes s'inspirent des résultats présentés précédemment au sujet de la recherche de Metz-Lutz et al. (1991) :

➤ **Hypothèse 1 – effet de l'âge**

Le taux moyen de réussite aux versions courtes A et B du test *Déno-100* devrait décliner avec l'âge.

➤ **Hypothèse 2 – effet du niveau d'éducation**

Le taux moyen de réussite aux versions courtes A et B du test *Déno-100* devrait être meilleur pour les sujets avec un niveau de scolarité élevé (plus de 9 ans de scolarité) que celui des participants avec un plus faible niveau d'études (jusqu'à 9 ans de scolarité).

➤ **Hypothèse 3 – effet de sexe**

Le taux moyen de réussite des femmes aux versions courtes A et B du test *Déno-100* devrait être plus élevé que celui des hommes.

Pour chacune de ces hypothèses, les éventuels effets de ces variables individuelles sur le temps de réponse, qui ne sont habituellement jamais considérés dans les recherches, seront analysés à titre exploratoire. Il s'agit ainsi de déterminer s'il existe un effet de l'âge, du sexe ou du niveau de scolarité sur la rapidité avec laquelle un sujet répond aux items de chacune des deux versions courtes du *Déno-100*.

4. Résultats

Les résultats présentés ci-dessous portent sur deux mesures : le *taux moyen de réussite* (relatif aux scores obtenus aux deux versions courtes et parallèles du test) et le *temps de réponse* à chaque version courte du test (chronométré pour chacune des listes de 24 items). Pour une présentation qualitative des résultats, le lecteur peut se référer à l'Annexe IV.

Une première partie des analyses statistiques porte sur les versions courtes et parallèles telles que constituées pour ce travail, avec 24 items chacune. La seconde partie porte sur des analyses de versions plus courtes, à 20 items, pour lesquelles des images ont été éliminées après les passations. Seul le taux de réussite a pu être considéré pour cette deuxième option de versions courtes et parallèles. Le temps de réponse étant mesuré sur l'ensemble des items de chacune des listes, il est probable que cette mesure ait été influencée par les items éliminés, plus difficiles à dénommer et demandant ainsi plus de temps pour produire le nom de l'image présentée.

En effet, à l'issue des passations, il s'avère que certains items présentent un taux de réponses dominantes nettement en dessous de la moyenne (94,2% de réussite pour la version A et 91,9% pour la version B). Les deux versions constituées (à 24 items) présentent effectivement des taux moyens de réussite relativement élevés, toutefois lorsque l'on observe les résultats de chacune des versions dans les détails, on constate que le taux de réussite à la version A varie entre 76,1% et 100% selon les items (voir Annexe IV A). De même pour la version B, le taux de réussite présente d'importantes variations, entre 50% et 100% (voir Annexe IV E). C'est pourquoi il a été jugé utile d'analyser les résultats obtenus en éliminant les quatre plus mauvais items de chaque version, afin de ne garder que ceux pour lesquels la réponse dominante représentait plus de 90% de réussite. Ce seuil a été retenu afin de pouvoir éliminer un nombre égal d'items avec les plus mauvais résultats dans chacune des versions courtes, tout en garantissant que les moins bonnes performances des sujets-contrôle restent élevées. De plus, ce chiffre est relativement proche du degré moyen de consensus en dénomination orale obtenu par Kremin et al. (2005) pour la version expérimentale du *Déno-100*, qui est de 88%.

4.1. Résultats des versions à 24 items

Les données obtenues aux deux épreuves de dénomination orale d'images des 46 sujets ont été soumises à une analyse de variance à mesures répétées. Cette analyse a montré un effet principal d'âge, mais s'est révélée peu informative pour certaines comparaisons, masquant où se situaient les effets. C'est pourquoi les résultats présentés sont ceux des données traitées avec des tests statistiques bilatéraux de Student à un seuil α de 0.05 (voir Annexe III A).

➤ **Caractère parallèle des deux versions du test**

En ce qui concerne le caractère parallèle des deux versions (A et B), le taux moyen de réussite de la version A est de 94,2% et celui de la version B est de 91,9%. Ces deux résultats semblent relativement proches, pourtant les analyses statistiques montrent une différence significative ($t_{0,975}[45]=2.014$; $p<.05$) entre les taux moyens de réussite obtenus à chacune des deux versions courtes (voir Figure 3). Ces dernières ne sont donc pas parallèles du point de vue des scores obtenus. Ce résultat inattendu est probablement expliqué par le groupe des sujets les plus jeunes (50-59 ans). En effet, si l'on compare les résultats de ces participants aux deux versions du test (âge 1A vs âge 1B, voir Annexe III A pour les analyses statistiques détaillées), les différences sont également statistiquement significatives ($t_{0,975}[9]=2.262$; $p<.05$). Ce groupe de sujets-contrôle présente en effet un taux de réussite plus élevé à la version A qu'à la version B du test *Déno-100*, tandis qu'il n'existe aucune différence statistiquement significative entre les deux versions pour les autres tranches d'âge.

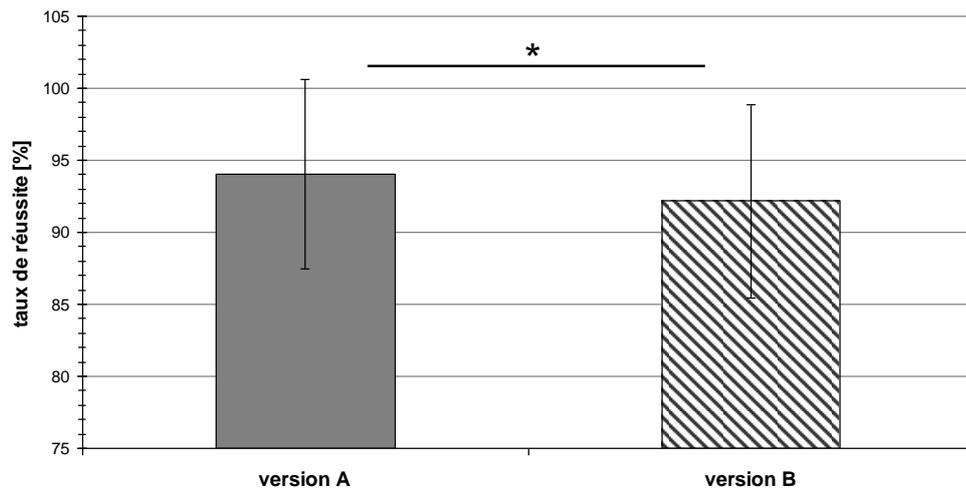


Figure 3 : Moyennes (et écart-types) des taux de réussite [%] pour chacune des deux versions, tous âges confondus.

Si l'on se penche à présent sur les temps moyens de réponse obtenus à chacune des listes de 24 items, les résultats ne présentent pas de différence statistiquement significative ($t_{0.975}[45]=2.014$; $p=.124$). En d'autres termes, les sujets-contrôle mettent en moyenne autant de temps à répondre à la version A du test qu'à la B (voir Figure 4). Les versions sont donc parallèles du point de vue de cette seconde mesure.

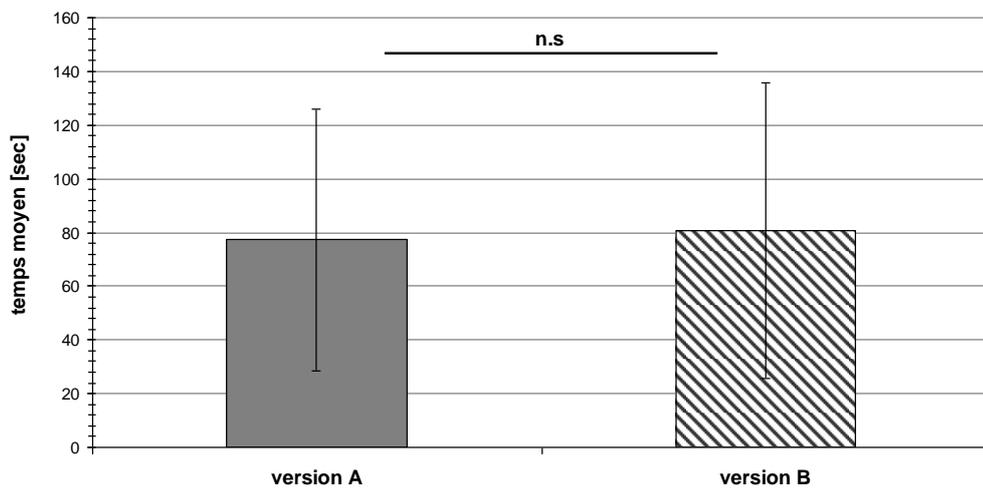


Figure 4 : Moyennes (et écart-types) des temps de réponses [sec] pour chacune des deux versions, tous âges confondus.

Au vu de ces résultats et d'après les analyses statistiques initiales avant les passations, nous pouvons partir du présupposé que les versions constituées sont

parallèles, malgré la différence observée au sein d'une classe d'âge seulement (50-59 ans), et analyser les effets liés aux variables individuelles jugées nécessaires pour un établir les normes d'un test. Ainsi, les principales analyses présentées dans ce qui suit, concernant les effets d'âge, de sexe et de niveau d'éducation, ont été menées sur la base des résultats des deux versions de test confondues, ceci afin d'obtenir les mêmes normes pour les deux versions courtes du *Déno-100*.

➤ **Hypothèse 1 : effet d'âge**

Si l'on considère les scores moyens en fonction de l'âge, on trouve des différences statistiquement significatives, les trois premiers groupes d'âge (50-79 ans) se différenciant du dernier (80 ans et +). En effet, tant les sujets de 50 à 59 ans, ($t_{0,975}[36]=2.028$; $p<.05$) que ceux de 60 à 69 ans ($t_{0,975}[40]=2.021$; $p<.05$) et ceux de 70 à 79 ans ($t_{0,975}[46]=2.013$; $p<.05$) présentent un taux de réussite plus élevé que leur aînés de 80 ans et plus (voir Figure 5). Le score moyen chute donc significativement à partir de 80 ans. En ce qui concerne l'étalonnage des scores, ce résultat nous amène à retenir deux classes d'âge :

- Âge 1-3 regroupés ;
- Âge 4.

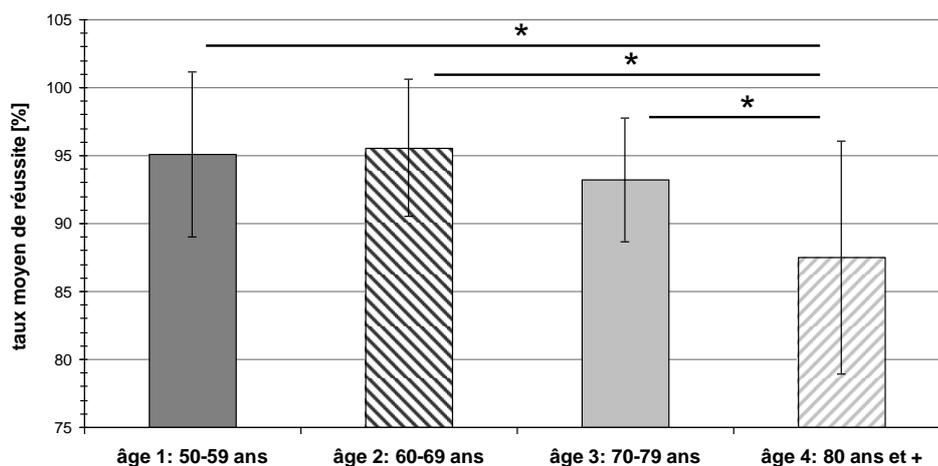


Figure 5 : Moyennes (et écart-types) des taux de réussite [%] en fonction de l'âge.

Du point de vue du temps moyen de réponse, on constate également des différences statistiquement significatives, mais entre d'autres groupes de sujets. De façon générale, le temps moyen de réponse a tendance à augmenter avec l'âge (voir Figure 6). Les participants les plus jeunes (50-59 ans) ont des résultats significativement différents de ceux de 70 à 79 ans ($t_{0,975}[48]=2.011$; $p<.05$) et des plus âgés de la cohorte, de 80 ans et plus ($t_{0,975}[36]=2.028$; $p<.05$). Il en va de même pour les sujets ayant entre 60 et 69 ans : leurs résultats présentent des différences significatives avec leurs aînés de 70 à 79 ans ($t_{0,975}[52]=2.007$; $p<.05$) et de 80 ans et plus ($t_{0,975}[40]=2,021$; $p<.05$). Etant donné qu'il n'y a pas de différence significative entre les résultats des sujets âgés de 50 à 69 ans pour le temps moyen de réponse ($t_{0,975}[42]=2.018$; $p=.658$), on peut regrouper ces deux tranches d'âge en vue de l'établissement des normes pour le temps de réponse. En ce qui concerne les participants de 70 à 79 ans, leurs résultats montrent, comme il l'a été mentionné, un temps de réponse significativement différent des sujets plus jeunes (50-69 ans), mais également des plus âgés, de 80 ans et plus ($t_{0,975}[46]=2.013$; $p<.05$).

Schématiquement et en vue de l'établissement des normes des versions courtes du test *Déno-100* pour le temps de réponse, nous avons trois classes d'âge :

- Âge 1-2 regroupés ;
- Âge 3 ;
- Âge 4.

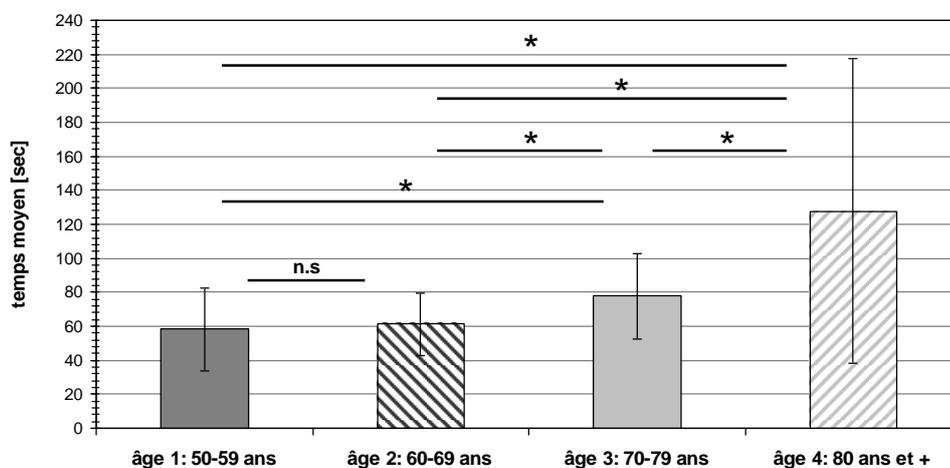


Figure 6 : Moyennes (et écart-types) des temps de réponses [sec] en fonction de l'âge

➤ **Hypothèse 2 : effet du niveau d'éducation**

Les analyses statistiques montrent des différences significatives entre les niveaux d'éducation ($t_{0,975}[90]=1.987$; $p<.05$). En effet, les sujets ayant un niveau de scolarité plus élevé ont un taux de réussite supérieur et répondent plus rapidement que ceux ayant un niveau de scolarité plus faible (voir Figures 7 et 8).

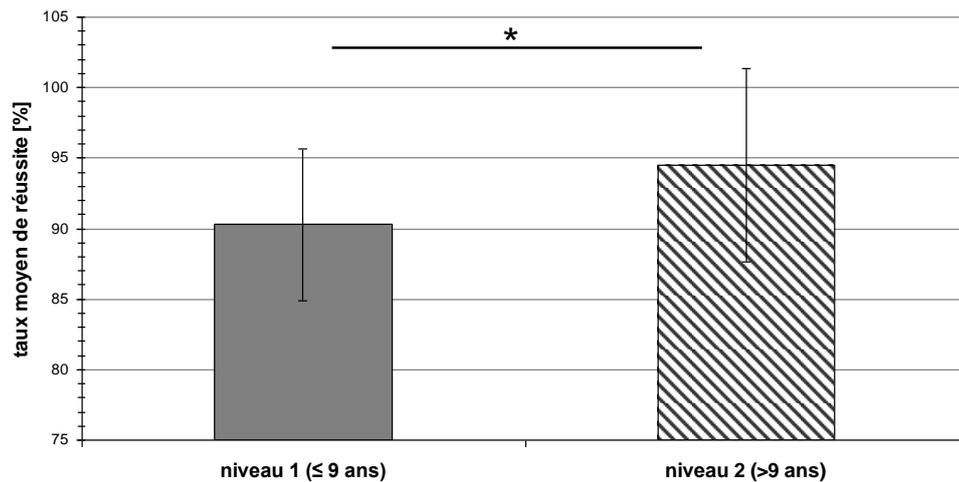


Figure 7 : Moyennes (et écart-types) des taux de réussite [%] pour chacune des deux versions, en fonction du niveau d'éducation, tous âges confondus.

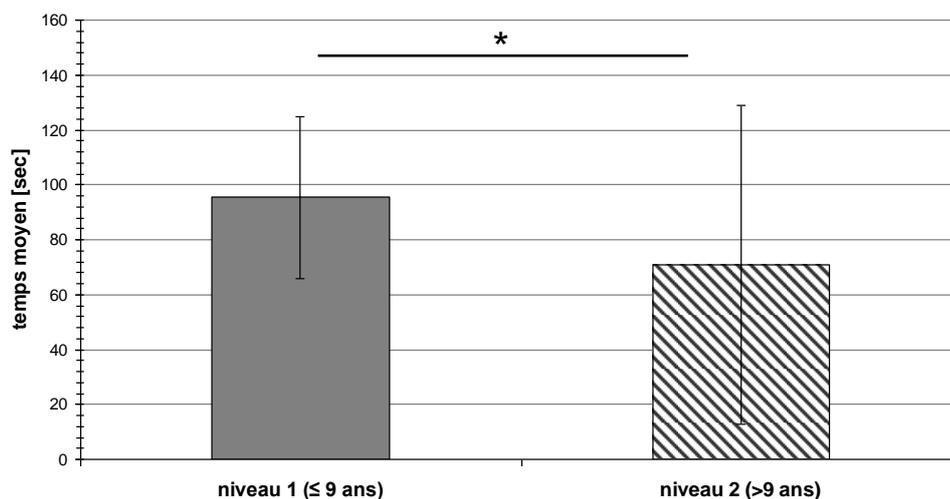


Figure 8 : Moyennes (et écart-types) des temps de réponse [sec] en fonction du niveau d'éducation, tous âges confondus.

Toutefois, si l'on regarde ces résultats à une plus petite échelle, on constate que l'effet du niveau d'éducation n'est pas constant. En effet, une différence significative entre les taux de réussite (voir Figure 9) n'est présente que pour la version B, tous âges confondus ($t_{0,975}[44]=2.015$; $p<.05$). De même, une différence significative pour le temps de réponse (voir Figures 10) n'est présente qu'au sein de la version A, tous âges confondus également ($t_{0,975}[44]=2.015$; $p<.05$).

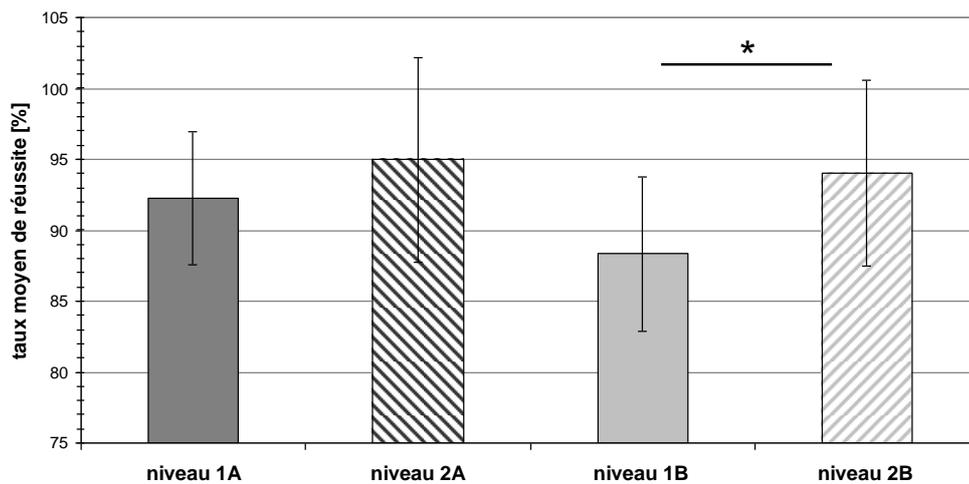


Figure 9 : Moyennes (et écart-types) des taux de réussite [%] en fonction du niveau d'éducation et de la version, tous âges confondus.

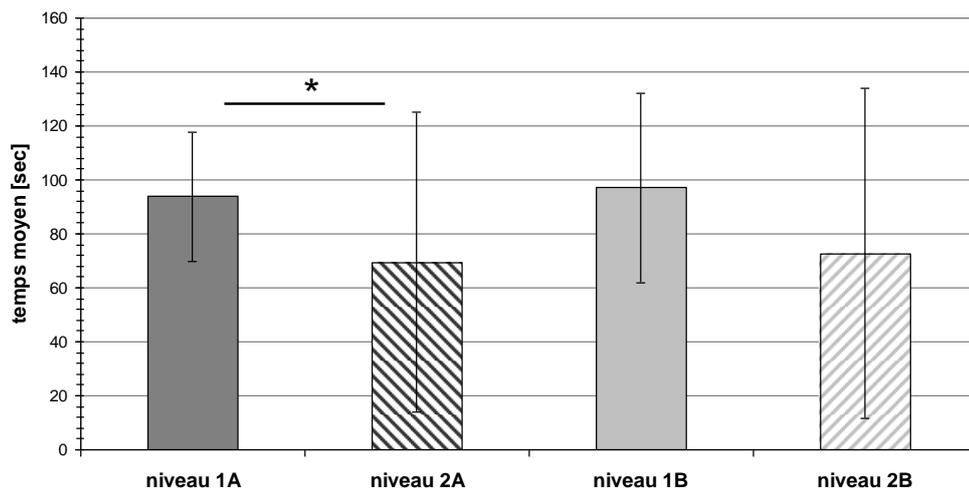


Figure 10 : Moyennes (et écart-types) des temps de réponse [sec] en fonction du niveau d'éducation et de la version, tous âges confondus.

Bien que les résultats généraux aillent dans le sens de ce qui est rapporté dans la littérature, à savoir que le taux de réussite aux tests de dénomination est meilleur pour les personnes ayant un niveau de scolarité élevée, ils sont susceptibles d'entrer en interaction avec les effets d'âge. Il faudrait donc idéalement analyser l'effet croisé de ces deux paramètres, pour chaque tranche d'âge, et exclure toute interaction afin de valider les résultats précédemment exposés. Toutefois, les groupes de sujets n'ayant pu être équilibrés du point de vue du niveau d'éducation, il n'a pas été possible d'analyser systématiquement les effets de cette dernière variable pour chaque tranche d'âge, excepté pour les sujets de 70 à 79 ans (voir Figures 11 et 12). Les analyses statistiques ont d'ailleurs montré des différences significatives dans les résultats de ces participants, tant en ce qui concerne le taux de réussite que les temps de réponse ($t_{0,975}[28]=2.048$; $p<.05$).

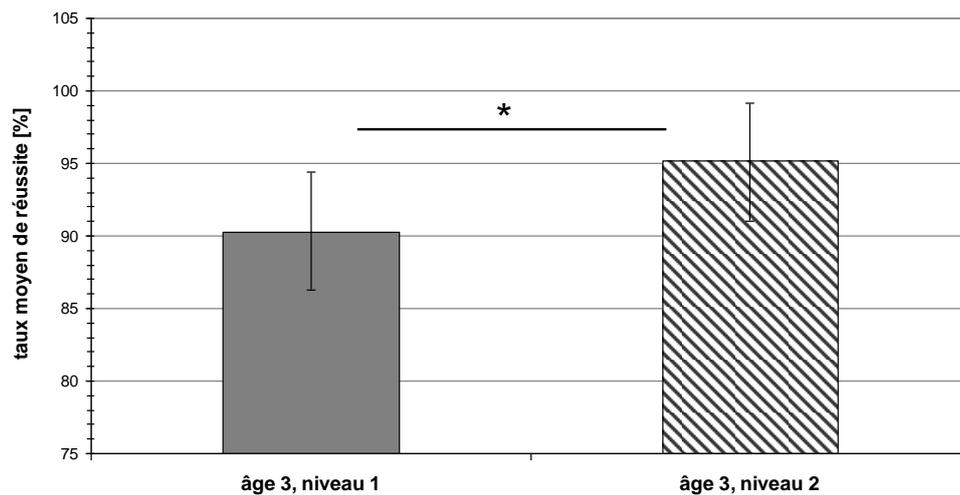


Figure 11 : Moyennes (et écart-types) des taux de réussite [%] des sujets d'âge 3 (70-79 ans) en fonction du niveau d'éducation.

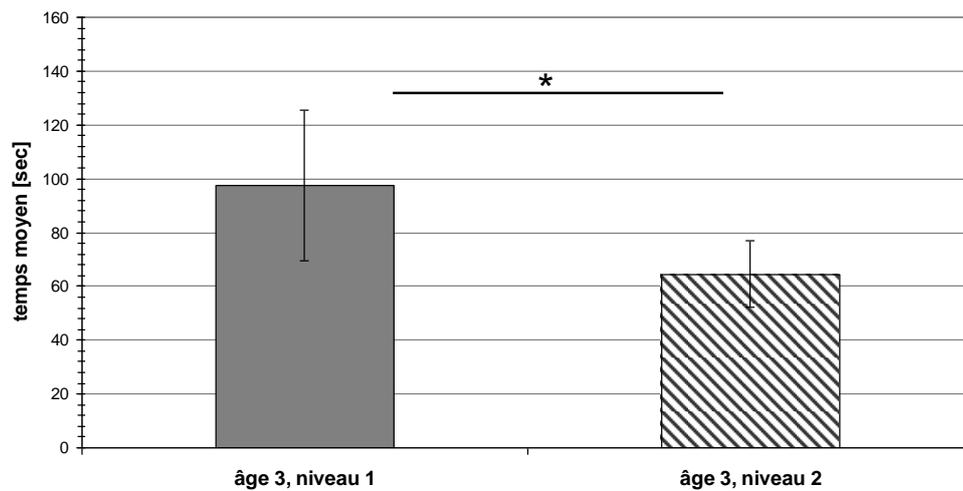


Figure 12 : Moyennes (et écart-types) des temps de réponse [sec] des sujets d'âge 3 (70-79 ans) en fonction du niveau d'éducation.

De plus, les résultats entre les deux listes sont constants au sein de cette tranche d'âge (voir Figures 13 et 14). Les sujets de 70 à 79 ans ayant un niveau de scolarité plus élevé présentent donc un taux de réussite supérieur et répondent plus rapidement que ceux du même âge mais ayant un niveau de scolarité plus faible, que ce soit pour la version A ($t_{0,975}[13]=2.160$; $p<.05$) ou B du test ($t_{0,975}[13]=2.160$; $p<.05$).

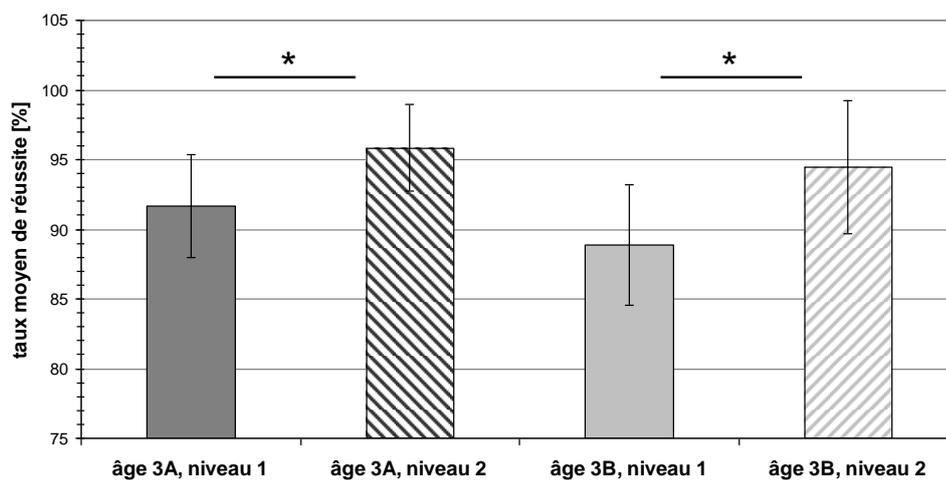


Figure 13 : Moyennes (et écart-types) des taux de réussite [%] des sujets d'âge 3 (70-79 ans) en fonction du niveau d'éducation et de la version.

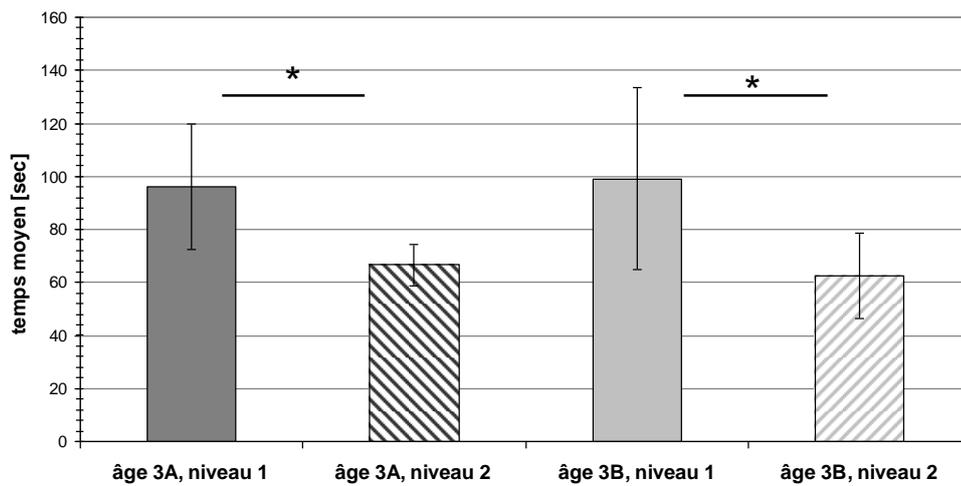


Figure 14 : Moyennes (et écart-types) du temps de réponse [sec] des sujets d'âge 3 (70-79 ans) en fonction du niveau d'éducation et de la version.

➤ Hypothèse 3 : effet du sexe

Aucun effet de sexe n'a été révélé par les analyses statistiques. Toutes les différences sont non significatives (voir Figures 15 et 16), que ce soit pour le taux de réussite ($t_{0.975}[90]=1.987$; $p=.199$) ou pour les temps de réponses ($t_{0.975}[90]=1.987$; $p=.155$).

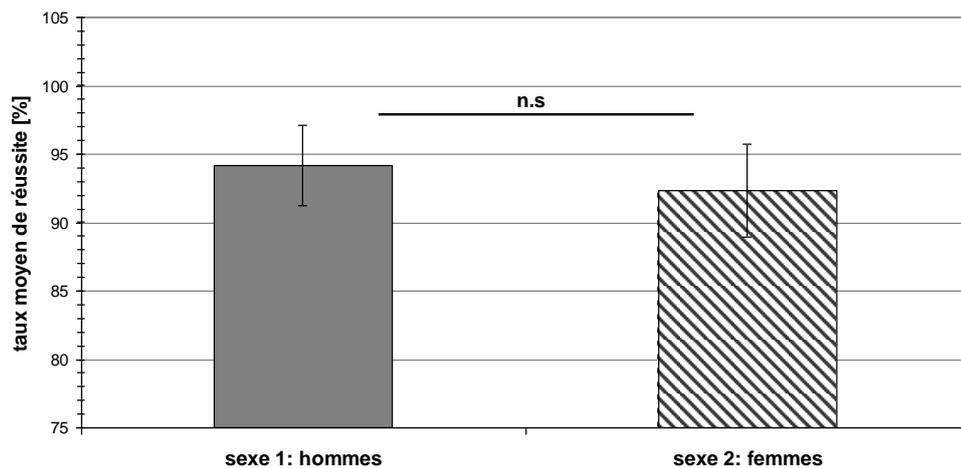


Figure 15 : Moyennes (et écart-types) des taux de réussite [%] en fonction du sexe, tous âges confondus.

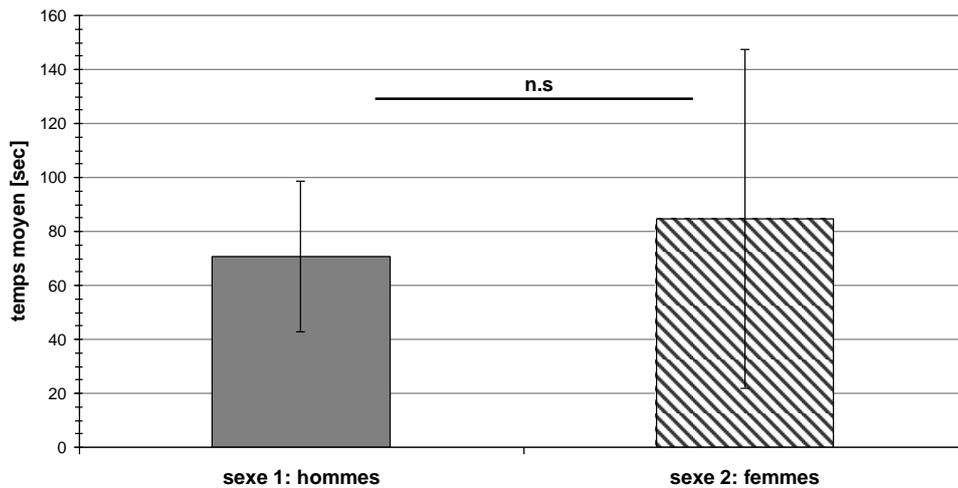


Figure 16 : Moyennes (et écart-types) des temps de réponse [sec] en fonction du sexe, tous âges confondus.

➤ Effet d'ordre

Afin d'exclure de effets parasites liés à l'ordre des versions courtes ou à la familiarité de la tâche, l'ordre de passations (A-B ou B-A) a été attribué de façon aléatoire aux sujets. Les analyses statistiques ne montrent aucun effet d'ordre. En effet, lorsque l'on compare les sujets ayant passé les tests dans l'ordre A-B avec ceux ayant répondu dans l'ordre B-A (voir Figures 17 et 18), les résultats ne présentent pas de différences significatives, ni dans le taux de réussite ($t_{0,975}[90]=1.987$; $p=.345$), ni dans le temps de réponse ($t_{0,975}[90]=1.987$; $p=.141$).

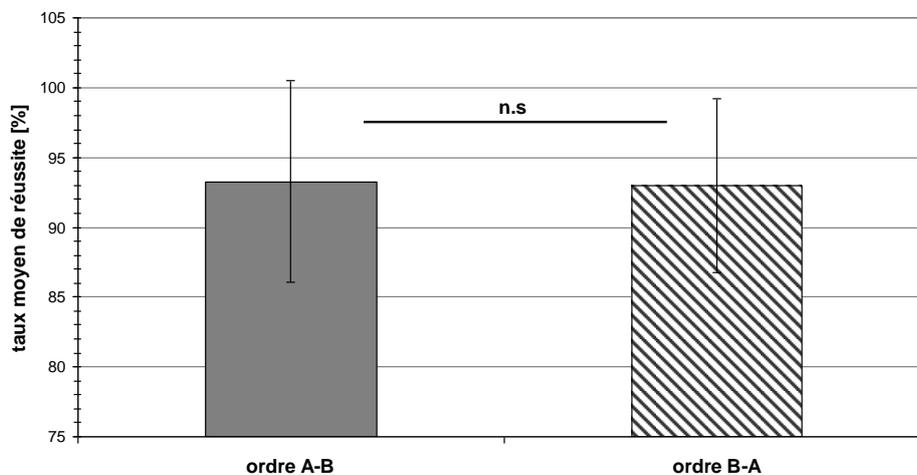


Figure 17 : Moyennes (et écart-types) des taux de réussite [%] en fonction de l'ordre de passation des versions, tous âges confondus.

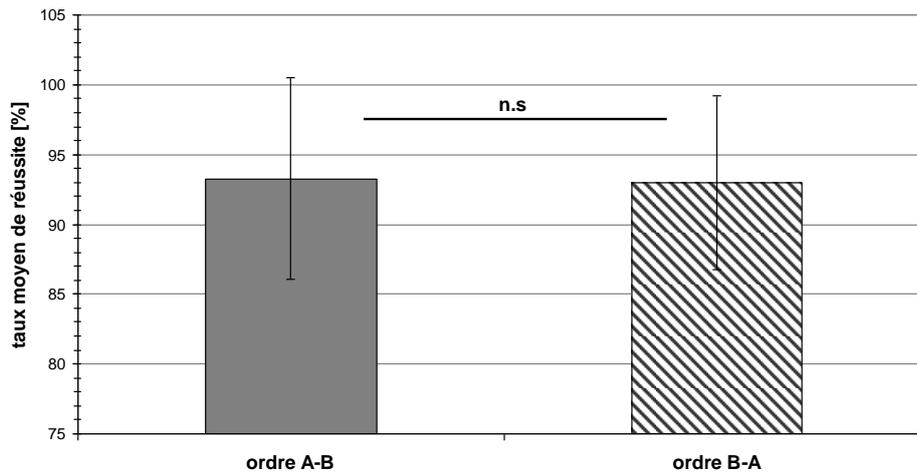


Figure 18 : Moyennes (et écart-types) des temps de réponse [sec] en fonction de l'ordre de passation des versions, tous âges confondus.

4.2. Résultats des versions à 20 items

Etant donné les résultats à certains items (version A : « flûte », « pyjama », « lacets », et « chou-fleur » ; version B : « cigogne », « gomme », « bretelles » et « phare »), avec un taux de réussite bien inférieur à la moyenne, il a été envisagé d'éliminer les plus mauvaises images de chacune des versions et de n'en garder que 20 contre 24 initialement sélectionnées. Toutefois, cette alternative a eu le désavantage de masquer tous les effets observés pour les taux de réussite avec 24 items. Les analyses statistiques ne présentent aucune différence significative. Les deux versions du test sont ainsi trop faciles et perdent tout pouvoir discriminant : les résultats plafonnent.

Pour plus de détails concernant cette partie des analyses, le lecteur peut consulter l'Annexe III B.

4.3. Etalonnage

Pour clore cette présentation de résultats, voici l'étalonnage proposé pour les deux versions courtes et parallèles du test *Déno-100*. Au vu des résultats obtenus, seul l'âge

des sujets a permis d'établir des normes pour la version à 24 items. La version à 20 items masquant tous les résultats permettant de différencier les groupes de sujets n'a donc pas été retenue pour l'étalonnage.

Tableau III : Normes pour le taux de réussite aux épreuves.

	Taux moyen de réussite	Ecart-type	Seuil de normalité
âges 1-3 : 50-79 ans	94 %	5,3	84 % réussite
âge 4 : 80 ans et +	88 %	8,5	70 % réussite

Les scores moyens ont été convertis en taux moyen de réussite. Le seuil de normalité en fonction de l'âge est obtenu en soustrayant 2 écarts-type au taux moyen de réussite des sujets-contrôle. En dessous de ce seuil, les résultats pourraient être considérés comme relevant du pathologique.

Tableau IV : Normes pour le temps de réponse.

	Temps moyen de réponse	Ecart-type	Seuil de normalité
âges 1-2 : 50-69 ans	60 sec	21,4	103 sec
âge 3 : 70-79 ans	78 sec	25,3	128 sec
âges 4 : 80 ans et +	128 sec	89,6	307 sec

Le seuil de normalité pour le temps de réponse en fonction de l'âge est calculé par l'addition de 2 écarts-type au temps moyen de réponse obtenu par les sujets-contrôle. En-dessus de ce seuil, les résultats demanderont une attention particulière et pourraient être considérés comme relevant du pathologique.

5. Discussion

5.1. Discussion sur les résultats

Cette recherche a porté sur l'élaboration de deux versions courtes et parallèles du *Déno-100* (Kremin et al., 2005) ainsi que sur leur étalonnage auprès d'une population-contrôle francophone de Suisse romande, âgée entre 50 et 91 ans. Etant donné le rôle possible des variables individuelles sur les performances en dénomination orale d'images, l'essentiel des résultats s'est consacré à l'analyse des effets d'âge, de sexe et du niveau d'éducation sur les versions à 24 items construites

Comme le montrent les tests statistiques, les résultats obtenus n'ont pas toujours été ceux escomptés pour établir les normes des versions courtes et parallèles du test *Déno-100*. En premier lieu, le problème du supposé caractère parallèle des deux versions courtes du test. Les deux versions constituées ne présentaient au départ pas de différence significative quant aux variables contrôlées dans le choix des items, elles étaient donc *a priori* parallèles. Toutefois, à l'issue des passations, on constate que ce n'est pas le cas lorsque l'on se penche sur les scores moyens obtenus à chacune des versions. Il existe en effet des différences statistiquement significatives. Ces résultats semblent avoir été influencés par un groupe de sujets particulier, à savoir les participants âgés entre 50 et 59 ans, pour qui les résultats présentent une forte variabilité dans le taux de réussite aux deux épreuves. Nous avons tout de même pris le parti de les considérer comme parallèles, étant donné que les résultats semblent être influencés par un groupe de sujets seulement, et d'entreprendre les analyses statistiques prévues pour l'établissement des normes. Il ne semblait en effet pas pertinent d'établir deux étalonnages distincts pour les deux versions courtes étant donné qu'elles ne présentaient pas de différences statistiquement significatives au départ.

En ce qui concerne les trois hypothèses guidant l'étalonnage des épreuves, rappelons qu'elles traitent de l'effet des variables individuelles (âge, sexe et niveau d'éducation) sur le taux moyen de réussite aux deux versions courtes du *Déno-100*

confondues. De plus, et ce à titre exploratoire, les effets des dites variables sur le temps moyen de réponse ont également été analysés.

➤ **Hypothèse 1 : effet de l'âge**

Si l'on compare les quatre tranches d'âge des sujets-contrôle, on constate que le taux moyen de réussite tend à diminuer avec l'âge. Il chute de façon significative à partir de 80 ans (voir Tableau III : 88% de bonnes réponses en moyenne à partir de 80 ans contre 94% entre 50 et 79 ans). On ne note toutefois pas de différence significative entre les groupes de participants plus jeunes, à savoir entre 50 et 79 ans. L'hypothèse selon laquelle les performances des sujets diminueraient avec l'âge se trouve donc confirmée. D'autre part, le temps de réponse a tendance à augmenter de façon significative avec l'âge. Plus les sujets sont âgés et plus ils mettent de temps à répondre. On note la présence de deux seuils significatifs, les différences se marquant dès 70 ans (voir Tableau IV : 78 secondes en moyenne pour 24 items, contre 60 secondes entre 50 et 69 ans) puis dès 80 ans (voir Tableau IV : 128 secondes en moyenne pour 24 items).

Il est intéressant de constater que ces deux versions courtes du test *Déno-100* sont plus sensibles pour discriminer les sujets en fonction de leur âge sur la base de la mesure du temps de réponse plutôt que sur le score obtenu aux épreuves. En effet, les différences entre les âges sont plus marquées pour les temps de réponse que pour les taux de réussite. La mesure du temps de réponse pourrait donc s'avérer être un critère intéressant à prendre en compte lors de l'évaluation d'un patient.

➤ **Hypothèse 2 : niveau d'éducation**

En ce qui concerne la deuxième hypothèse, elle postule que les performances des sujets avec un niveau d'éducation élevé, c'est-à-dire de plus de 9 années d'études, auraient de meilleures performances aux tests de dénomination que leurs pairs de niveau de scolarité plus faible (jusqu'à 9 ans d'études). Cette hypothèse se trouve confirmée. Les analyses statistiques ont en effet montré une différence significative en ce qui concerne le taux de réussite aux deux versions courtes du *Déno-100*, tous âges

confondus. De plus, le niveau d'éducation semble également avoir un effet sur le temps de réponse, plus court pour les participants ayant plus de 9 années d'études.

Il faut rappeler les potentielles interactions avec les autres variables individuelles que sont le sexe et l'âge des participants. On ne constate pas d'effet du niveau d'éducation pour toutes les tranches d'âge. Cependant, ces résultats sont probablement dus à des problèmes d'échantillonnage. En effet, les difficultés de recrutement de participants ayant moins de 9 ans d'études ont eu pour conséquence d'avoir des groupes de sujets mal équilibrés pour pouvoir étudier de telles interactions. Il a néanmoins été possible d'analyser l'effet croisé de l'âge et du niveau d'éducation pour le groupe de sujets-contrôle âgés de 70 à 79 ans, dont la répartition en fonction du sexe et du niveau d'éducation a été la mieux équilibrée. Les analyses statistiques effectuées pour ce groupe de participants ont également confirmé la présente hypothèse. Le niveau de scolarité influence ainsi significativement le taux de réussite ainsi que le temps de réponse aux versions courtes du *Déno-100* chez les participants de 70 à 79 ans. Bien que cet effet ne soit pas présent dans toutes les tranches d'âge, pour les raisons évoquées plus haut, il semble pertinent de maintenir le critère de niveau de scolarité pour étalonner un test de dénomination.

➤ **Hypothèse 3 : effet de sexe**

Enfin, l'hypothèse selon laquelle les femmes auraient un taux de réussite supérieur à celui des hommes, se trouve infirmée dans cette recherche. Ce résultat n'est pas surprenant étant donné que cette variable est en effet peu présente dans les étalonnages de test et la moins souvent contrôlée dans les recherches. De plus, les items sélectionnés ne relèvent pas particulièrement de champs sémantiques à connotation masculine ou féminine pouvant influencer la reconnaissance des objets et leur dénomination.

5.2. Critiques et limites

Les deux versions courtes et parallèles du *Déno-100* constituées de 24 items chacune ont été bien reçues par les participants. Les épreuves sont peu contraignantes, leur passation est agréable et plutôt rapide (1 à 2 minutes en moyenne, selon les âges). Ces aspects-là ne suffisent malheureusement pas à valider ces versions courtes et parallèles. Les résultats obtenus semblent avoir été influencés par trois grands types de facteurs : le recrutement des participants, le choix des items et la méthodologie générale utilisée dans cette recherche.

➤ Recrutement des participants

En premier lieu, le nombre insuffisant de participants, avec comme conséquence des groupes de sujets mal équilibrés si l'on considère les variables individuelles (âge, sexe et niveau d'éducation). Le recrutement a été difficile, tout particulièrement en ce qui concerne le niveau d'éducation dans les tranches d'âge extrêmes (50-59 ans ainsi que 80 ans et plus). En Suisse, il est en effet difficile de trouver des sujets francophones âgés entre 50 et 59 ans avec moins de 9 ans de scolarité, en d'autres termes n'ayant pas fait un apprentissage au moins. A l'opposé, chez les participants âgés de 80 ans et plus, on ne rencontre que peu de femmes ayant fait plus de 9 ans d'études. D'ailleurs les résultats les plus consistants sont ceux pour lesquels les groupes expérimentaux contenaient le plus grand nombre de sujets et étaient le mieux équilibrés (âge 3 : 70 à 79 ans)

➤ Matériel

En second lieu, beaucoup de questions se posent autour des items sélectionnés. En effet, lorsque l'on considère les résultats des versions courtes à 24 items, on constate que certaines images présentent un taux de réussite bien en dessous de la moyenne. Les items concernés (« *bretelles* » ou « *gomme* » par exemple) ont donné lieu à de nombreuses erreurs visuelles et visuo-sémantiques. Les dessins étant d'emblée mal

reconnus, les réponses ont été multiples et variées, faisant chuter le taux de réussite. Toutefois, lorsque l'on élimine les plus mauvais items de chacune des versions courtes, les effets liés aux variables individuelles qui sont utiles pour établir des normes se trouvent masqués et l'on risque d'observer un effet plafond : le test devient alors trop facile et peu discriminant du point de vue de l'âge ou du niveau d'éducation des sujets.

Deux questions se posent alors : doit-on garder les 24 items afin de discriminer les sujets en fonction de leur âge et de leur niveau d'éducation ou doit-on penser ces versions courtes et parallèles comme des tests qui doivent être parfaitement réussis par les sujets normaux et ne garder ainsi que 20 items ? Ces deux options sont développées plus bas, dans les ouvertures proposées pour de prochaines recherches.

➤ **Méthodologie**

Enfin, la méthodologie elle-même et les conditions de passation sont sans doute la cause de la diversité des réponses produites par les sujets-contrôle. On constate que de nombreux participants ont produit des réponses multiples. Le degré de consensus en dénomination orale étant contrôlé, deux hypothèses existent pour expliquer cette constatation. La première remet en cause la consigne, qui n'aurait pas spécifié assez clairement qu'une seule réponse était exigée. La deuxième concerne l'attitude des participants eux-mêmes en situation de test, qui auraient produit plusieurs noms, parfois même des substantifs de niveau subordonné, bref tout cela dans l'optique de « bien faire » et de donner « l'étiquette » la plus précise possible de l'image à dénommer. Rappelons qu'un des inconvénients des tests de dénomination est précisément que tant l'expérimentateur que le sujet sont conscients qu'il y a une réponse attendue et que cela peut biaiser le fonctionnement naturel de l'évocation lexicale.

Cette multitude de productions a eu des conséquences sur la détermination des réponses dominantes et sur la cotation des deux épreuves. C'est pourquoi la décision d'une échelle en trois points a été prise. D'autres réponses non dominantes ont été acceptées (RND+) au vu de la fréquence avec laquelle elles ont été produites. De plus, en cas de réponses multiples, la question de n'accepter que la première production s'est posée, mais ayant l'impression que les sujets, se sachant participer à une recherche,

cherchaient plutôt à « bien paraître », la décision a été prise de tout de même attribuer 2 points dans ces situations, pour autant que la cible ait été produite.

5.3. Ouvertures

Au vu de ces écueils, des améliorations méthodologiques seraient nécessaires. Dans tous les cas, l'étalonnage devrait être repris sur un plus grand nombre de sujets. En ce qui concerne les buts initiaux de la présente recherche, deux grandes options peuvent être envisagées.

➤ **Etude pilote**

La première consiste en des améliorations de la méthodologie pour la sélection des items. Il s'agirait de mener une étude pilote sur un plus grand nombre d'items afin d'éliminer les plus mauvaises images par la suite, à savoir celles qui ont donné lieu à un mauvais taux de réussite au vu de leur bon degré de consensus en dénomination orale. L'étalonnage se ferait ainsi dans un deuxième temps, sur les listes d'items choisies après l'étude pilote.

La présente recherche pourrait justement être considérée comme une étude pilote. L'étalonnage pourrait donc être à nouveau mené mais sur les versions courtes et parallèles à 20 items, afin de déterminer si des effets liés aux variables individuelles choisies existent, notamment en ce qui concerne le temps de réponse qui n'a pu être étudié pour ces versions dans ce travail.

➤ **Nouvelles considérations sur l'étalonnage**

La deuxième option consiste à envisager ce test comme une épreuve du type de la *Batterie Rapide de Dénomination (BARD, Croisile, 2005)*. Il ne s'agirait plus d'avoir des versions courtes et parallèles discriminant les sujets en fonction de leur âge ou de leur niveau d'éducation, mais plutôt qu'elle soient parfaitement réussies par tous les

sujets-contrôles, quels que soient leur âge, leur sexe ou leur niveau d'éducation. D'un point de vue clinique, et selon la *BARD*, la présence d'au moins une erreur pourrait ainsi déjà être caractéristique d'un trouble de l'évocation lexicale. En ce qui concerne le présent travail, les versions à 20 items pourraient être représentatives de cette dernière option.

6. Conclusions

Pour conclure, la conception et la validation d'un test dans le domaine de l'aphasie est une tâche de longue haleine. Selon Mazaux et al. (2007, p.145), cela nécessiterait même entre « 5 à 10 ans en moyenne ». En ce qui concerne le présent travail, on constate que bien que le test soit bien contrôlé au départ, on n'obtient pas aisément les résultats attendus. Les biais sont multiples et le contrôle de tous les paramètres susceptibles d'influencer les performances des participants est difficile. On retiendra que malgré les difficultés rencontrées, il a pu être démontré que l'âge et le niveau d'éducation jouent bel et bien un rôle sur les performances en dénomination orale d'images, tant en ce qui concerne le taux de réussite que le temps de réponse pour des versions à 24 items. A noter toutefois que lorsque l'on élimine certains items après-coup et que l'on ne garde que ceux qui présentent un taux de réussite élevé (plus de 90%), ces effets ne semblent plus exister. Le *Déno-100* représentant un matériel bien contrôlé et ce travail ayant donné des pistes pour l'étalonnage de deux versions courtes parallèles, il serait ainsi pertinent de pouvoir reprendre cette recherche afin d'aboutir à un test de dénomination orale d'images bien contrôlé et utile à la clinique logopédique.

7. Bibliographie

Littérature

Alario, F. X., & Ferrand, L. (1999). A set of 400 pictures standardized for French : Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments & Computers*, *31*, 531-552.

Alario, F. X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments & Computers*, *36* (1), 140-155.

Bachy-Langedock, N. (1989). *Batterie d'examen des troubles en dénomination*. Bruxelles : Editest.

Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies : French norms for agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers*, *35* (1), 158-167

Chainay, H., Rosenthal, V., & Goldblum, M. C. (1998). Normes de dénomination de 315 images en couleur et en noir et blanc. *Revue de Neuropsychologie*, *8* (2), 179-239.

Croisile, B. (2005). Une Batterie Rapide de Dénomination (BARD) en consultation mémoire. *La Revue Francophone de Gériatrie et de Gérontologie*, *12*, 526-530.

David, D. (2007). Les aphasies progressives primaires. In J.-M. Mazaux, M. Allard, P. Pradat-Diehl, & V. Brun (Eds.), *Aphasies et aphasiques* (pp. 97-110). Issy-les-Moulineaux : Masson.

Deloche, G., & Hannequin, D. (1997). *DO 80 : Test de dénomination orale d'images*. Paris : Les Editions du Centre de Psychologie Appliquée.

Estienne, F., & Piérart, B. (2006). *Les bilans de langage et de voix. Fondements théoriques et pratiques*. Paris : Masson.

Herbert, R., Hickin, J., Howard, D., Osborne, F., & Best, W. (2008). Do picture-naming provide a valid assessment of lexical retrieval in conversation in aphasia ? *Aphasiology*, 22 (2), 184-203.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia : Lea & Febiger.

Kremin, H. (1994). Perturbations lexicales : les troubles de la dénomination. In X. Seron, & M. Jeannerod (Eds.), *Neuropsychologie humaine* (pp. 375-389). Liège : Mardaga.

Kremin, H. (2002). L'accès au lexique en dénomination d'images : problèmes actuels. *Psychologie Française*, 47 (2), 77-91.

Kremin, H., Akhutina, T., Basso, A., Davidoff, J., De Wilde, M., Kitzing, P., Lorenz, A., Perrier, D., Van der Sandt-Koenderman, M., Vendrell, J., & Weniger, D. (2003). A crosslinguistic data bank for oral picture naming in Dutch, English, German, French, Italian, Russian, Spanish, and Swedish (PEDOI). *Brain and cognition*, 53, 243-246.

Kremin, H., Perrier, D., De Wilde, M., & Dordain, M. (2005). *Déno-100. Un test de dénomination orale*. Version expérimentale non publiée.

Laiacona, M., Luzzatti, C., Zonca, G., Cuarnaschelli, C., & Capitani, E. (2001). Lexical and semantic factors influencing picture naming in aphasia. *Brain and Cognition*, 46, 184-187.

Laganaro, M., Di Pietro, M., & Schnider, A. (2006). What does recovery from anomia tell us about the underlying impairment : The case study of similar anomic patterns and different recovery. *Neuropsychologia*, 44, 534-545.

Mazaux, J.-M., Dehail, P., Davet, J.-C., Pradat-Diehl, P., Brun, V. (2007). Tests et bilans d'aphasie. In J.-M. Mazaux, P. Pradat-Diehl, & V. Brun (Eds.), *Aphasie et aphasiques* (pp. 144-157). Issy-les-Moulineaux : Masson.

Mazaux, J. M., & Orgogozo, J. M. (1981). *HDAE-F, Boston Diagnostic Aphasia Examination. Echelle française*. Issy-les-Moulineaux : Editions scientifiques et psychologiques.

Metz-Lutz, M. N., Kremin, H., Deloche, G., Hannequin, D., Ferrand, I., Perrier, D., Quint, S., Dordain, M., Bunel, G., Cardebat, D., Larroque, C., Lota, A. M., Pichard, B., & Blavier, A. (1991). Test standardisé de dénomination orale, contrôle des effets de l'âge, du sexe et du niveau de scolarité chez les sujets adultes normaux. *Revue de Neuropsychologie, 1*, 73-95.

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory and Cognition, 20*, 705-714.

Nespoulous, J.-L., Lecours, A. R., Lafond, D., Lemay, A., Puel, M., Joannette, Y., Cott, F., & Rascol, A. (1992). *Protocole Montréal-Toulouse d'examen linguistique de l'aphasie : MT 86 (version M1 bêta)*. L'Ortho Edition : Isbergues.

Seron, X. (2004). *La neuropsychologie cognitive*. Paris : Presses Universitaires de France.

Seron, X., & Feyereisen, P. (1985). Neurolinguistique. In J. A. Rondal & X. Seron (Eds.), *Troubles du langage. Diagnostic et rééducation*. Liège : Mardaga.

Sirois, M., Kremin, H., & Cohen, H. (2006). Picture-naming norms for Canadian French : name agreement, familiarity, visual complexity, and age of acquisition. *Behavior Research Methods, 2*, 300-306.

Snodgrass, J. G., & Vanderwart, M. A. (1980). A standardized set of 260 pictures : norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology : Human Learning and Memory, 6* (2), 174-215.

Thuillard Colombo, F., & Assal, G. (1992). Adaptation française du test de dénomination de Boston. Versions abrégées. *Revue européenne de Psychologie Appliquée*, 42 (1), 67-71.

Tran, T. M. (1997). Intérêts et limites des épreuves de dénomination d'images en pratique clinique aphasiologique. *Glossa*, 59, 16-23.

Tran, T. M. (2007). Rééducation des troubles de la production lexicale. In J.-M. Mazaux, M. Allard, P. Pradat-Diehl, & V. Brun (Eds.), *Aphasies et aphasiques* (pp. 205-215). Issy-les-Moulineaux : Masson.

Internet

www.aphasie.org

Annexes

Table des matières des annexes

ANNEXE I :	Données des versions courtes et parallèles et analyses statistiques préliminaires	61
ANNEXE II :	Consigne et cahier de passation des versions courtes et parallèles du test de dénomination Déno-100.....	63
ANNEXE III :	Résultats des analyses statistiques.....	67
ANNEXE IV :	Analyse qualitative des résultats	71
ANNEXE V :	Présentation critique de quelques tests en français évaluant la dénomination orale	79

**Annexe I : Données des versions courtes et parallèles et analyses
statistiques préliminaires**

A. Version A : caractéristiques des 24 items sélectionnés.

Item	Catégorie	CONS	FREQ	Log F	AMA	CVIS	FAM	NB SYLL	
cuillère	obj	117	857	129,33	1,59	1,4	4,86	2	AMA précoce
chèvre	liv	116	3890	135,9	2,31	4,3	2,15	1	
lion	liv	120	2193	133,4	2,55	4,5	1,24	1	
canard	liv	118	1280	131,07	1,9	5,4	2,99	2	
coq	liv	116	1965	132,9	1,9	3,4	2,51	1	
souris	liv	86	1102	130,42	1,96	4	1,95	2	
rose	liv	120	14375	141,58	2,25	5,5	4,01	1	
brouette	obj	120	454	126,57	2,26	2,8	2,8	2	
lacets	obj	110	600	127,78	2,27	3,7	3,82	2	
dominos	obj	101	268	124,3	2,79	4	1,97	3	
pyjama	obj	111	318	125,02	2,59	4,7	3,78	3	
tabouret	obj	118	714	128,54	2,71	2,3	3,71	3	
artichaut	liv	116	172	122,36	3,12	3,6	2,8	3	AMA tardif
chou fleur	liv	108	66	118,2	3,15	3,4	3,21	2	
enveloppe	obj	94	2552	134,07	3,16	2,1	4,28	3	
piano	obj	117	4581	136,61	3,16	3,8	2,4	2	
perroquet	liv	116	525	127,2	3,4	5,3	1,39	3	
violon	obj	109	1622	132,1	3,43	4,6	1,79	2	
maïs	liv	118	637	128,04	3,62	4,1	2,43	2	
asperge	liv	113	244	123,9	3,71	2,8	2,76	2	
flûte	obj	89	1489	131,73	3,74	2,7	1,82	1	
rhinocéros	liv	110	377	125,76	3,9	4,2	1,04	4	
hache	obj	109	1070	130,3	3,07	2,7	1,97	1	
cactus	liv	99	132	121,21	4,81	3,9	1,82	2	

B. Version B : caractéristiques des 24 items sélectionnés.

Item	Catégorie	CONS	FREQ	Log F	AMA	CVIS	FAM	NB SYLL	
cheval	liv	115	15464	141,9	1,76	4,1	2,64	2	AMA précoce
vache	liv	118	3858	135,86	1,68	4,3	2,95	1	
fourchette	obj	120	602	127,8	1,69	1,4	4,95	2	
âne	liv	114	2453	133,9	2,13	4,1	2	1	
peigne	obj	120	719	128,57	2,06	3,4	4,58	1	
raisin	liv	114	1370	131,37	2,1	2,5	3,5	2	
ours	liv	112	1131	130,53	2,13	3,2	1,41	1	
ciseaux	obj	96	873	129,41	2,25	2,6	4,35	2	
singe	liv	112	1649	132,2	2,5	4,7	1,44	1	
gomme	obj	102	310	124,91	2,55	2,9	3,94	1	
bretelles	obj	95	358	125,54	2,68	2,6	1,8	2	
scie	obj	107	486	126,87	2,88	2,5	2,6	1	
chaîne	obj	111	4020	136,04	2,99	4,3	2,39	1	AMA tardif
flèche	obj	117	1832	132,63	3,04	2,6	1,59	1	
tulipe	liv	118	270	124,3	3,12	3,7	3,26	2	
crocodile	liv	106	326	125,13	3,22	4,5	1,06	3	
zèbre	liv	120	140	121,46	3,26	5,2	1,14	1	
cigogne	liv	94	175	122,43	3,6	4,1	1,35	2	
pingouin	liv	107	687	128,37	3,72	4,3	1,11	2	
téléphone	obj	116	2576	134,11	3,81	3,8	4,91	3	
phare	obj	116	1144	130,58	3,88	3,9	1,85	1	
ananas	liv	117	95	119,78	3,96	4,7	3,14	3	
paon	liv	114	536	127,29	4,03	5,2	1,45	1	
trompette	obj	83	1418	131,5	5,1	1,86	1,97	2	

C. Comparaisons des versions A et B : statistiques préliminaires avec test de Student bilatéraux au seuil α de 0.05. Aucune valeur n'est significative. Les tests sont donc comparables du point de vue des variables ayant guidé la sélection des items.

	CONS	FREQ	AMA	CVIS	FAM
Tests de Student	0,087	0,225	0,091	0,063	0,044
t critiques	2,013	2,013	2,013	2,013	2,013
Degrés de liberté	46	46	46	46	46

Annexe II : Consigne et cahier de passation des versions courtes et parallèles du test de dénomination Déno-100

A. Consigne donnée aux participants

« Vous allez voir des dessins d'objets et d'animaux que vous connaissez probablement. On vous demande de bien vouloir produire le nom de chaque dessin. Ne faites pas de commentaires, donnez seulement le nom correspondant au dessin, précédé par son article – comme par exemple : UNE maison, UN bureau... Alors comment ça s'appelle ? »

L'item donné en exemple (« un crayon ») est le suivant :



B. Cahier de passation**DÉNO-100 : TEST DE DÉNOMINATION ORALE D'IMAGES**

(Kremin et al., version expérimentale de 2005)

Informations sur le participant :

Nom :	Prénom :
Date de naissance :	Âge (mois) :
Sexe :		
Durée de scolarité :	<input type="checkbox"/> ≤ 9 ans	Langue maternelle :
	<input type="checkbox"/> > 9 ans	Autres langues :
		
Date de passation :		

Cotation de l'épreuve :

NOTE TOTALE	Version A :	Réponses correctes : Durée totale :
	Version B :	Réponses correctes : Durée totale :

Version A

ITEMS	REPOSE	RESULTAT
1. Asperge		
2. Dominos		
3. Souris (rat)		
4. Violon		
5. Canard (col vert)		
6. Artichaut		
7. Rhinocéros		
8. Flûte		
9. Maïs		
10. Rose		
11. Lion		
12. Chèvre (bique)		
13. Coq		
14. Piano		
15. Hache		
16. Pyjama		
17. Enveloppe (lettre)		
18. Perroquet		
19. Lacets (attaches à souliers)		
20. Brouette		
21. Chou-fleur		
22. Tabouret		
23. Cuillère		
24. Cactus		
		NOTE TOTALE
		➤ Réponses correctes :
		➤ Durée totale :

Cotation :

- 2pts/item dénommé correctement.
- 1pt/item dénommé avec les propositions entre parenthèses, avec extension de la cible ou contenant la racine lexicale de la cible.

Version B

ITEMS	REPONSE	RESULTAT
1. Scie (égoïne)		
2. Cigogne		
3. Chaîne		
4. Gomme		
5. Flèche		
6. Ours		
7. Ananas		
8. Cheval		
9. Singe (chimpanzé)		
10. Crocodile (alligator)		
11. Vache		
12. Bretelles		
13. Fourchette		
14. Pingouin (manchot)		
15. Ciseaux		
16. Raisin (grappe)		
17. Zèbre		
18. Phare		
19. Trompette		
20. Téléphone		
21. Tulipe		
22. Âne (bourricot)		
23. Peigne		
24. Paon		
		NOTE TOTALE
		➤ Réponses correctes :
		➤ Durée totale :

Cotation :

- 2pts/item dénommé correctement.
- 1pt/item dénommé avec les propositions entre parenthèses, avec extension de la cible ou contenant la racine lexicale de la cible.

Annexe III : Résultats des analyses statistiques

A. Versions à 24 items : résultats des analyses statistiques avec des tests de Student bilatéraux au seuil α de 0.05. Les résultats accompagnés d'un astérisque (*) sont significatifs.

	Tests Student : taux de réussite	Tests de Student : temps de réponse	t critiques	degrés de liberté
<u>Versions parallèles ?</u>				
version A vs B	0,020*	0,124	2,014	45
<u>Effets d'âge ?</u>				
âge 1 vs âge 2	0,786	0,658	2,018	42
âge 1 vs âge 3	0,219	0,011*	2,011	48
âge 1 vs âge 4	0,004*	0,006*	2,028	36
âge 2 vs âge 3	0,082	0,011*	2,007	52
âge 2 vs âge 4	0,002*	0,008*	2,021	40
âge 3 vs âge 4	0,018*	0,037*	2,013	46
âge 1A vs âge 1B	0,011*	1,000	2,262	9
âge 2A vs âge 2B	0,651	0,093	2,201	11
âge 3A vs âge 3B	0,136	0,743	2,145	14
âge 4A vs âge 4B	0,862	0,088	2,306	8
âge 1A vs âge 2A	0,408	0,984	2,086	20
âge 1A vs âge 3A	0,062	0,047*	2,069	23
âge 1A vs âge 4A	0,021*	0,070	2,110	17
âge 2A vs âge 3A	0,317	0,020*	2,060	25
âge 2A vs âge 4A	0,047*	0,068	2,093	19
âge 3A vs âge 4A	0,096	0,188	2,074	22
âge 1B vs âge 2B	0,351	0,546	2,086	20
âge 1B vs âge 3B	0,843	0,119	2,069	23
âge 1B vs âge 4B	0,129	0,056	2,110	17
âge 2B vs âge 3B	0,163	0,200	2,060	25
âge 2B vs âge 4B	0,015*	0,071	2,093	19
âge 3B vs âge 4B	0,079*	0,134	2,074	22

<u>Effets du niveau d'éducation ?</u>				
niveau 1 vs niveau 2	0,004*	0,009*	1,987	90
niveau 1A vs niveau 1B	0,002*	0,412	2,145	14
niveau 2A vs niveau 2B	0,362	0,203	2,042	30
niveau 1A vs niveau 2A	0,190	0,043*	2,015	44
niveau 1B vs niveau 2B	0,006*	0,096	2,015	44
âge 3: niveau 1 vs niveau 2	0,003*	0,001*	2,048	28
âge 3: niveau 1A vs niveau 1B	0,235	0,866	2,571	5
âge 3: niveau 2A vs niveau 2B	0,408	0,136	2,306	8
âge 3: niveau 1A vs 2A	0,035*	0,014*	2,160	13
âge 3: niveau 2A vs 2B	0,039*	0,033*	2,160	13
<u>Effets du sexe ?</u>				
sexe 1 vs sexe 2	0,199	0,155	1,987	90
sexe 1A vs sexe 1B	0,682	0,825	2,110	17
sexe 2A vs sexe 2B	0,102	0,759	2,052	27
sexe 1A vs sexe 2A	0,702	0,330	2,015	44
sexe 1B vs sexe 2B	0,154	0,312	2,015	44
<u>Effet d'ordre ?</u>				
ordre AB vs BA	0,345	0,141	1,987	90
version A : ordre AB vs BA	0,442	0,191	2,015	44
version B : ordre AB vs BA	0,573	0,433	2,015	44
ordre AB vs AB	0,133	0,978	2,069	23
ordre BA vs BA	0,044*	0,540	2,080	21

B. Versions à 20 items : résultats des analyses statistiques pour les versions courtes à 20 items avec des tests de Student bilatéraux au seuil α de 0.05. Les résultats accompagnés d'un astérisque (*) sont significatifs.

	Tests Student : taux de réussite	t critiques	degrés de liberté
<u>Versions parallèles ?</u>			
version A vs B	0,220	2,014	45
<u>Effets d'âge ?</u>			
âge 1 vs âge 2	0,532	2,018	42
âge 1 vs âge 3	0,520	2,011	48
âge 1 vs âge 4	0,384	2,028	36
âge 2 vs âge 3	0,981	2,007	52
âge 2 vs âge 4	0,256	2,021	40
âge 3 vs âge 4	0,256	2,013	46
âge 1A vs âge 1B	0,213	2,262	9
âge 2A vs âge 2B	0,032*	2,201	11
âge 3A vs âge 3B	0,272	2,145	14
âge 4A vs âge 4B	0,482	2,306	8
âge 1A vs âge 2A	0,576	2,086	20
âge 1A vs âge 3A	0,670	2,069	23
âge 1A vs âge 4A	0,171	2,110	17
âge 2A vs âge 3A	0,857	2,060	25
âge 2A vs âge 4A	0,219	2,093	19
âge 3A vs âge 4A	0,219	2,074	22
âge 1B vs âge 2B	0,091	2,086	20
âge 1B vs âge 3B	0,113	2,069	23
âge 1B vs âge 4B	0,904	2,110	17
âge 2B vs âge 3B	0,764	2,060	25
âge 2B vs âge 4B	0,652	2,093	19
âge 3B vs âge 4B	0,706	2,074	22

<u>Effets du niveau d'éducation ?</u>			
niveau 1 vs niveau 2	0,415	1,987	90
niveau 1A vs niveau 1B	0,246	2,145	14
niveau 2A vs niveau 2B	0,389	2,042	30
niveau 1A vs niveau 2A	0,619	2,015	44
niveau 1B vs niveau 2B	0,565	2,015	44
âge 3: niveau 1 vs niveau 2	0,192	2,048	28
âge 3: niveau 1A vs niveau 1B	0,296	2,571	5
âge 3: niveau 2A vs niveau 2B	0,695	2,306	8
âge 3: niveau 1A vs 2A	0,197	2,160	13
âge 3: niveau 2A vs 2B	0,686	2,160	13
<u>Effets du sexe ?</u>			
sexe 1 vs sexe 2	0,5484	1,987	90
sexe 1A vs sexe 1B	0,9117	2,110	17
sexe 2A vs sexe 2B	0,1775	2,052	27
sexe 1A vs sexe 2A	0,6792	2,015	44
sexe 1B vs sexe 2B	0,1978	2,015	44
<u>Effet d'ordre ?</u>			
ordre AB vs BA	0,873	1,987	90
	0,673		
version A : ordre AB vs BA	0,509	2,015	44
version B : ordre AB vs BA	0,290	2,015	44
	0,535		
ordre AB vs AB	0,873	2,069	23
ordre BA vs BA	0,673	2,080	21

Annexe IV : Analyse qualitative des résultats

A. Version A – taux de réussite : inventaire des réponses acceptées (RD et RND+) et pourcentages avec lesquels elles ont été données pour chaque item.

Cible	RD	% RD	RND +	% RND +	% total de réussite
1. asperge	asperge	100,0	-	0	100,0
2. dominos	dominos	95,7	-	0	95,7
3. souris	souris	91,3	rat	4,4	95,7
4. violon	violon	93,5	-	0	93,5
5. canard	canard	97,8	colvert	0	97,8
6. artichaut	artichaut	89,1	-	0	89,1
7. rhinocéros	rhinocéros	89,1	-	0	89,1
8. flûte *	flûte	89,1	-	0	89,1
9. maïs	maïs	95,7	-	0	95,7
10. rose	rose	100,0	-	0	100,0
11. lion	lion	100,0	-	0	100,0
12. chèvre	chèvre	100,0	bique	0	100,0
13. coq	coq	89,1	-	0	89,1
14. piano	piano	100,0	-	0	100,0
15. hache	hache	100,0	-	0	100,0
16. pyjama *	pyjama	76,1	-	0	76,1
17. enveloppe	enveloppe	87,0	lettre	6,5	93,5
18. perroquet	perroquet	97,8	-	0	97,8
19. lacets *	lacets	78,3	attache à souliers	2,2	80,5
20. brouette	brouette	97,8	-	0	97,8
21. chou-fleur *	chou-fleur	89,1	-	0	89,1
22. tabouret	tabouret	95,7	-	0	95,7
23. cuillère	cuillère	100,0	-	0	100,0
24. cactus	cactus	95,7	-	0	95,7
Moyennes		93,7		0,5	94,2

N.B. : les items accompagnés d'un astérisque () sont ceux qui ont été éliminés pour obtenir les versions à 20 items.*

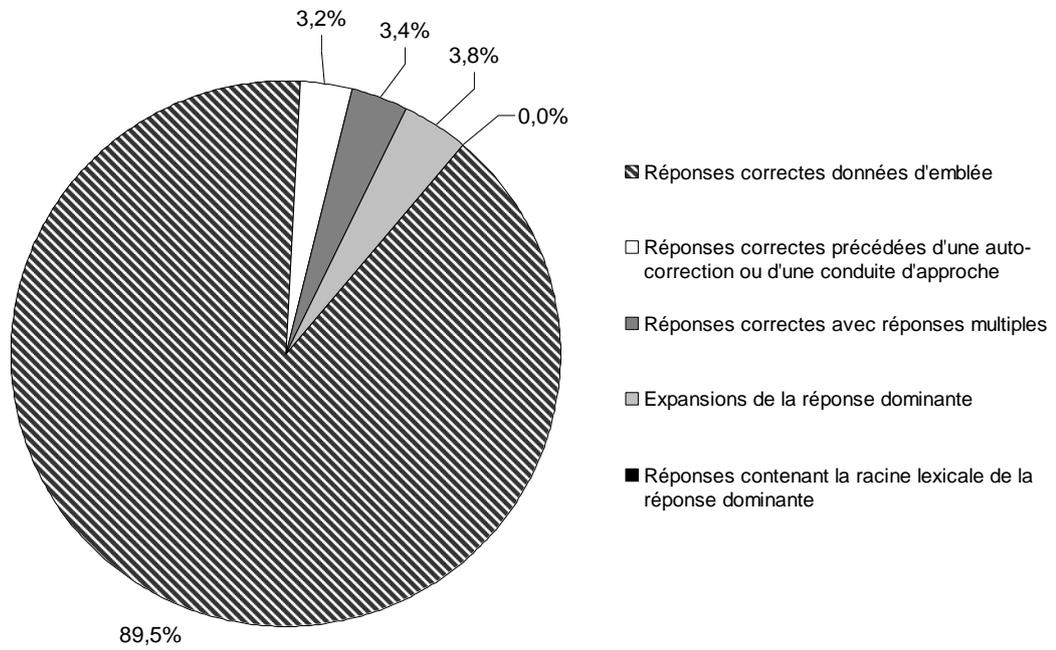
B. Version A – taux d'échec : inventaire des réponses non dominantes (RND- et absences de réponse) et pourcentages avec lesquelles elles ont été données pour chaque item.

Cible	RND -	% RND -	% absences de réponses	% total d'échec
1. asperge	-	0	0	0
2. dominos	jeu, dé	4,3	0	4,3
3. souris	-	4,3	0	4,3
4. violon	mandoline, guitare	6,5	0	6,5
5. canard	pigeon voyageur, cane	2,2	0	2,2
6. artichaut	légume, ananas, rose	6,6	4,3	10,9
7. rhinocéros	hippopotame	6,6	4,3	10,9
8. flûte	hautbois, clarinette, instrument de musique, trompette	10,9	0	10,9
9. maïs	grappe de maïs, pivot de polenta	4,3	0	4,3
10. rose	-	0	0	0
11. lion	-	0	0	0
12. chèvre	-	0	0	0
13. coq	poule	10,9	0	10,9
14. piano	-	0	0	0
15. hache	-	0	0	0
16. pyjama	costume, habit, complet	23,9	0	23,9
17. enveloppe	-	6,5	0	6,5
18. perroquet	faucon, coq	2,2	0	2,2
19. lacets	ficelle, attache, cordon, écheveau	19,5	0	19,5
20. brouette	charrette	2,2	0	2,2
21. chou-fleur	fleur, chou, gâteau	6,6	4,3	10,9
22. tabouret	siège, petit siège	4,3	0	4,3
23. cuillère	-	0	0	0
24. cactus	plante grasse	2,1	2,2	4,3
Moyennes		5,2	0,6	5,8

C. Version A – modes de réponse : inventaire des productions ayant abouti à la cible ou à une RND+.

Cible	% Donnée d'emblée	Avec conduites approche ou auto-corrections	%	Avec réponses multiples	%	Avec expansion de la RD	%	Contenant la racine lexicale de la RD	%
1. asperge	100,0	-	0	-	0	-	0	-	0
2. dominos	88,6	dé	2,3	-	0	jeu de dominos	9,1	-	0
3. souris	77,3	rat	2,3	rat	20,5	-	0	-	0
4. violon	88,4	violoncelle, guitare	4,7	violoncelle, guitare	7,0	-	0	-	0
5. canard	88,9	-	0	colvert, cane	11,1	-	0	-	0
6. artichaut	97,6	« légume...fenouil... »	2,4	-	0	-	0	-	0
7. rhinocéros	95,1	hippopotame	2,4	hippopotame	2,4	-	0	-	0
8. flûte	85,4	clarinette	2,4	clarinette	4,9	flûte à bec	7,3	-	0
9. maïs	63,6	épis de blé	2,3	-	0	épis de maïs	34,1	-	0
10. rose	100,0	-	0	-	0	-	0	-	0
11. lion	97,8	lionne	2,2	-	0	-	0	-	0
12. chèvre	93,5	un bouc	2,2	bouc, bouquetin	4,3	-	0	-	0
13. coq	85,7	poule	4,8	poule	9,5	-	0	-	0
14. piano	91,3	cuisinière	2,2	harmonium	2,2	piano droit	4,3	-	0
15. hache	100,0	-	0	-	0	-	0	-	0
16. pyjama	80,0	complet, costard, costume	14,3	uniforme, costume	5,7	-	0	-	0
17. enveloppe	95,3	-	0	lettre	4,7	-	0	-	0
18. perroquet	95,6	ara, cacatoès	4,4	-	0	-	0	-	0
19. lacets	59,5	fil, ficelle de souliers, attaches à souliers, corde ou écheveau, corde, ficelle	21,6	lasso	2,7	lacets de souliers, lacets de chaussures	16,2	-	0
20. brouette	100,0	-	0	-	0	-	0	-	0
21. chou-fleur	92,7	tourte	2,4	nid, tarte	4,9	-	0	-	0
22. tabouret	81,4	-	0	escabeau	2,3	deux tabourets	16,3	-	0
23. cuillère	95,7	-	0	-	0	cuillère à crème, cuillère en bois	4,3	-	0
24. cactus	95,5	« un pic, j'ai pas le nom, il faut pas s'y frotter... », « c'est piquant... »	4,5	-	0	-	0	-	0
Moyennes	89,5		3,2		3,4		3,8		0

D. Version A – modes de réponse : présentation graphique des processus ayant abouti à la production de la cible ou d'une RND+.



E. Version B – taux de réussite : inventaire des réponses acceptées (RD et RND+) et pourcentages avec lesquels elles ont été données pour chaque item.

Cible	RD	% RD	RND +	% RND +	% total de réussite
1. scie	scie	100	égoïne	0	100
2. cigogne *	cigogne	71,7	-	0	71,7
3. chaîne	chaîne	93,5	-	0	93,5
4. gomme *	gomme	50,0	-	0	50
5. flèche	flèche	97,8	-	0	97,8
6. ours	ours	100	-	0	100
7. ananas	ananas	97,8	-	0	97,8
8. cheval	cheval	100	-	0	100
9. singe	singe	100	chimpanzé	0	100
10. crocodile	crocodile	97,8	alligator	0	97,8
11. vache	vache	97,8	-	0	97,8
12. bretelles *	bretelles	58,7	-	0	58,7
13. fourchette	fourchette	97,8	-	0	97,8
14. pingouin	pingouin	87,0	manchot	8,7	91,3
15. ciseaux	ciseaux	100	-	0	100
16. raisin	raisin	100	grappe	0	100
17. zèbre	zèbre	97,8	-	0	97,8
18. phare *	phare	76,1	-	0	76,1
19. trompette	trompette	87,0	-	0	87
20. téléphone	téléphone	100	-	0	100
21. tulipe	tulipe	95,7	-	0	95,7
22. âne	âne	100	bourricot	0	100
23. peigne	peigne	100	-	0	100
24. paon	paon	95,7	-	0	95,7
Moyennes		91,8		0,4	91,9

N.B. : les items accompagnés d'un astérisque () sont ceux qui ont été éliminés pour obtenir les versions à 20 items.*

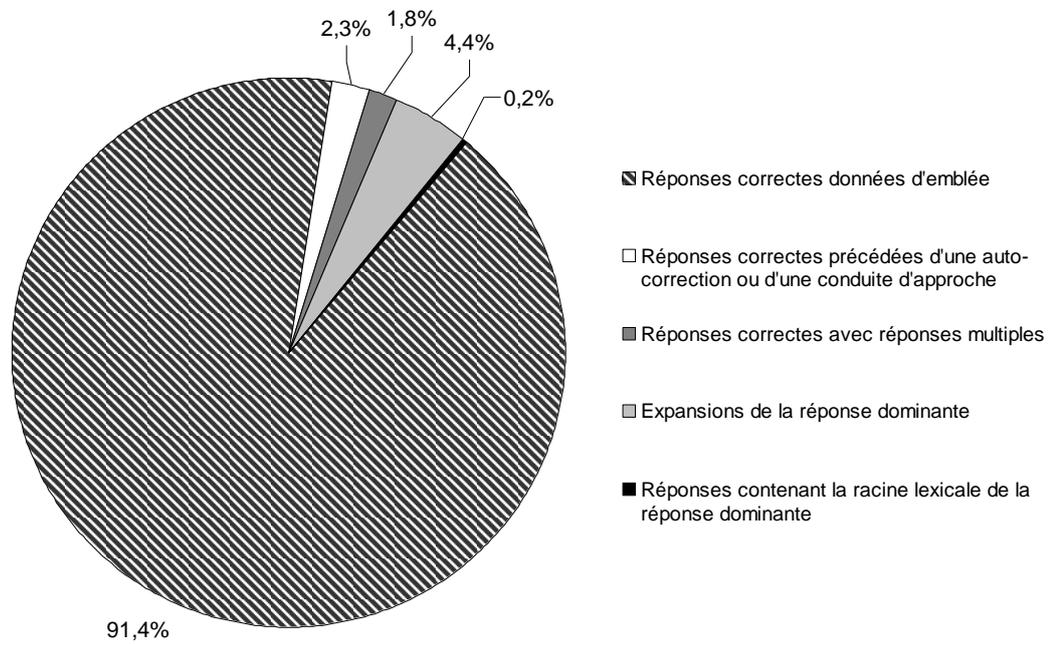
F. Version B – taux d'échec : inventaire des réponses non dominantes (RND- et absences de réponse) et pourcentages avec lesquelles elles ont été données pour chaque item.

Cible	RND -	% RND -	% absences de réponses	% total d'échec
1. scie	-	0	0	0
2. cigogne	ibis, cygne, flamant/flamant rose, pélican, héron, héron, oiseau	26,1	2,2	28,3
3. chaîne	collier	6,5	0	6,5
4. gomme	natel/téléphone/portable, musique à bouche, zappeur/télécommande, radio, beurrier, rasoir	39,1	10,9	50,0
5. flèche	pique	2,2	0	2,2
6. ours	-	0	0	0
7. ananas	pamplemousse	2,2	0	2,2
8. cheval	-	0	0	0
9. singe	-	0	0	0
10. crocodile	-	0	2,2	2,2
11. vache	taureau	2,2	0	2,2
12. bretelles	harnais, siège, fauteuil	26,1	15,2	41,3
13. fourchette	fourche	2,2	0	2,2
14. pingouin	pélican	6,5	2,2	8,7
15. ciseaux	-	0	0	0
16. raisin	-	0	0	0
17. zèbre	-	0	2,2	2,2
18. phare	tour, cheminée, moulin à vent, minaret	23,9	0	23,9
19. trompette	clairon, trombone, clarinette	13,0	0	13,0
20. téléphone	-	0	0	0
21. tulipe	rose, fleur, fleur	4,3	0	4,3
22. âne	-	0	0	0
23. peigne	-	0	0	0
24. paon	-	0	4,3	4,3
Moyennes		6,4	1,6	8,1

G. Version B – modes de réponse : inventaire des productions ayant abouti à la cible ou à une RND+.

Cible	% Donnée d'emblée	Avec conduites approche ou auto-corrections	%	Avec réponses multiples	%	Avec expansion de la RD	%	Contenant la racine lexicale de la RD	%
1. scie	95,7	-	0	-	6,1	scie à bois	4,3	-	0
2. cigogne	84,8	cygne, pélican, "cet oiseau qui..."	9,1	héron	0	-	0	-	0
3. chaîne	79,1	collier pour montre, collier	4,7	-	4,3	chaîne de montre	14,0	chaînette	2,3
4. gomme	91,3	téléphone	4,3	rasoir	2,2	-	0	-	0
5. flèche	97,8	-	0	fléchette	0	-	0	-	0
6. ours	87,0	nounours	4,3	-	0	ours polaire, brun, blanc	6,5	ourson	2,2
7. ananas	100	-	0	-	2,2	-	0	-	0
8. cheval	95,7	un coco	2,2	poney	4,3	-	0	-	0
9. singe	95,7	-	0	gorille, chimpanzé	4,4	-	0	-	0
10. crocodile	95,6	-	0	alligator	2,2	-	0	-	0
11. vache	93,3	taureau, "vache...taureau..."	4,4	taureau	7,4	-	0	-	0
12. bretelles	85,2	"bretelles...harnais..."	3,7	harnais	0	paire de bretelles	3,7	-	0
13. fourchette	97,8	-	0	-	2,4	fourchette à salade	2,2	-	0
14. pingouin	95,2	"il y en a dans les iceberg..."	2,4	manchot	0	-	0	-	0
15. ciseaux	87,0	-	0	-	2,2	paire de ciseaux	13,0	-	0
16. raisin	37,0	-	0	grappe	0	grappe de raisin	60,9	-	0
17. zèbre	100	-	0	-	0	-	0	-	0
18. phare	91,7	tour dans la mer, "tour de contrôle en mer...", "leuchtturm, tour d'appel de bateaux..."	8,3	-	2,4	-	0	-	0
19. trompette	90,2	clarinette, instrument de musique, piston	7,3	saxophone	0	-	0	-	0
20. téléphone	100	-	0	-	0	-	0	-	0
21. tulipe	100	-	0	-	2,2	-	0	-	0
22. âne	97,8	-	0	cheval	0	-	0	-	0
23. peigne	100	-	0	-	0	-	0	-	0
24. paon	95,5	"il ouvre ses ailes...", autruche	4,5	-	0	-	0	-	0
Moyennes	91,4		2,3		1,8		4,4		0,2

H. Version A – modes de réponse : présentation graphique des processus ayant abouti à la production de la cible ou d'une RND+.



Annexe V : Présentation critique de quelques tests en français évaluant la dénomination orale

Afin de mieux comprendre les particularités du test de dénomination *Déno-100*, voici une brève présentation critique de quelques tests fréquemment utilisés dans la francophonie pour évaluer l'aphasie, et plus particulièrement la dénomination orale d'images.

➤ **Adaptation française du test de dénomination de Boston**

Le *Boston Naming Test* (*BNT*, Kaplan, Goodglass & Weintraub, 1983) est une épreuve largement utilisée dans le cadre de l'examen neuropsychologique dans les pays anglo-saxons. Outre son analyse quantitative, il permet une analyse qualitative du type d'erreurs. Certains items sont toutefois fortement marqués culturellement, c'est pourquoi ce test a été étalonné en français (Thuillard Colombo & Assal, 1992). Des versions courtes ont également été développées.

Cette version du *BNT* est le seul test actuellement étalonné en Suisse romande. Il n'est malheureusement pas exempt de critiques. Il date de plus de 30 ans et ne tient pas compte de toutes les variables considérées aujourd'hui comme intéressantes. En effet, seuls le niveau socio-culturel, l'âge et le sexe ont été contrôlés dans la version originale américaine. De plus, dans cette dernière, la passation se fait en temps limité (15 secondes par images) alors que les patients aphasiques ont besoin de plus de temps pour répondre. On risque donc de prendre une réponse qui pourrait être tardive mais correcte comme une absence de réponse.

➤ **Batterie d'examen des troubles en dénomination (ExaDé)**

L'*ExaDé* (Bachy-Langedock, 1989) est, comme son nom l'indique, une batterie pour l'évaluation de la dénomination orale. Elle comporte une épreuve principale et cinq épreuves complémentaires. La première, l'épreuve générale, comporte 90 items sélectionnés selon la fréquence d'usage dans la langue et la longueur des mots. Les cinq autres épreuves évaluent plus spécifiquement le rôle d'autres facteurs pouvant influencer les performances : nature, composition et caractéristiques des items ; temps de réponse et effet des processus de facilitation.

Cette batterie est intéressante et présente de nombreux avantages. L'épreuve générale est destinée à un premier examen approfondi de départ, mais aussi à des contrôles en cours et en fin de programme thérapeutique. L'ensemble des autres épreuves, plus spécifiques, permet une analyse fouillée des troubles, aide à cerner les variables les plus pertinentes pour le sujet examiné et met en relief d'éventuelles dissociations. Enfin les épreuves évaluant l'efficacité de diverses aides donnent des pistes pour les procédures potentiellement utilisables pour la rééducation. Toutefois, *l'ExaDé* présente aussi des inconvénients. La longueur des épreuves est difficile à concilier avec la pratique clinique courante et peut entraver la bonne relation « thérapeute-patient » s'il est utilisé sans ménager celui-ci. En effet, le versant « scientifique » et « humain » du thérapeute peuvent se trouver ici en opposition.

➤ **Protocole Montréal-Toulouse d'examen linguistique de l'aphasie (MT-86)**

Le *MT-86* (Nespoulous et al., 1992) est une batterie qui permet d'évaluer divers aspects du langage. Il comporte d'ailleurs un sous-test pour la dénomination orale. Le souci des auteurs a été de faire un test valide dans toute la francophonie. A l'instar des autres épreuves testant l'évocation lexicale, l'âge, le sexe et le niveau de scolarité ont été considérés *a posteriori* pour l'étalonnage. Ce sous-test explore la dénomination de substantifs et de verbes en contrôlant un grand nombre de paramètres : acceptabilité du mot dans la francophonie, fréquence d'usage, possibilité de représentation iconographique, potentiel associatif dans la langue, etc. Les auteurs du *DO 80* (Deloche & Hannequin, 1997) lui reprochent toutefois un nombre trop restreint d'items.

➤ **DO 80**

Le *DO 80* (Deloche & Hannequin, 1997) est un test qui s'est basé en partie sur les images de la banque de données de Snodgrass et Vanderwart (1980) en tenant compte de la longueur et de la fréquence des mots dans la langue. L'âge, la durée de la scolarité et le sexe des participants ont également été contrôlés. Ce test présente la caractéristique de proposer deux modes de cotation, l'un strict pour la recherche et l'autre plus souple pour la clinique. L'évaluation est quantitative et qualitative. En effet, le *DO 80* propose une grille d'analyse de la nature des erreurs permettant d'appréhender le patient dans son mode de fonctionnement. Enfin, il propose également

une liste des réponses non dominantes constituant un stock lexical dans lequel les personnes aphasiques vont parfois puiser. Contrairement au *BNT* (Kaplan et al., 1983), l'épreuve se déroule sans aucune contrainte de temps. Toutefois, dans la pratique clinique, on peut trouver que ce test reste encore trop long à administrer.

➤ **Normes de dénomination de 315 images en couleur et en noir et blanc**

Les normes utilisées auparavant en psychologie expérimentale étaient celles de la batterie standardisée de 260 dessins au trait en noir et blanc de Snodgrass et Vandewart (1980). Cette batterie a été normée en anglais. Nombreux sont les items dépendants de la culture et de la langue. Son utilisation en français peut donc poser problème, c'est pourquoi il a été jugé nécessaire d'obtenir des normes spécifiques pour la population francophone. De plus, l'utilisation de la batterie de Snodgrass et Vandewart (1980) a été étalonnée sur une population étudiante, ne pouvant ainsi tenir compte des divers niveaux socio-éducatifs ni de l'effet de l'âge sur la dénomination. Enfin, cette batterie est intéressante à des fins de recherche, mais elle est difficilement utilisable en clinique de par sa longueur.

Les normes de dénomination de 315 images en couleur et en noir et blanc (Chainay, Rosenthal, & Goldblum, 1998) veulent ainsi tenir compte des paramètres critiqués dans la batterie de Snodgrass et Vanderwart (1980). On ne peut pas vraiment parler de test de dénomination, à administrer dans sa totalité. Il s'agit plutôt d'une banque de données. Cet ensemble standardisé permet ainsi aux chercheurs francophones de répondre à des exigences méthodologiques et d'avoir un outil destiné à servir de matériel pour des épreuves de dénomination sur une population francophone adulte et âgée. En effet, un grand nombre de variables ont été considérées : diversité catégorielle des objets, nom dominant, familiarité avec l'objet, complexité visuelle de l'image, probabilité d'une réponse acceptable et caractère chromatique de l'image, etc.

➤ **Batterie Rapide de Dénomination (BARD)**

La *BARD* (Croisile, 2005) est un outil qui comporte 10 items, tous issus du *BNT* (Kaplan et al., 1983), constituant ainsi selon les auteurs une version courte de ce dernier, mais étalonnée et validée en français. Facile d'emploi, elle est parfaitement réussie par des sujets normaux (effet plafond constant), quels que soient leur âge, genre

ou niveau sociaux-éducatif. Les auteurs considèrent que la survenue d'au moins une erreur est hautement pathologique et le signe probable de l'existence d'une anomie, dans le cadre d'une aphasie ou d'une démence dégénérative.

L'avantage de la *BARD* est que la passation se fait en très peu de temps (moins de 15 secondes chez les sujets normaux). En effet, son but n'est pas d'évaluer finement la dénomination comme le feraient d'autres tests, tels que *l'ExaDé* (Bachy-Langedock, 1989), le *BNT* (Kaplan et al., 1983) ou le *DO 80* (Deloche & Hannquin, 1997), mais de repérer rapidement l'existence d'une anomie par un médecin dans le cadre d'une consultation mémoire médicale. Ce test cognitif rapide aurait ainsi la particularité de concilier une grande sensibilité et une forte spécificité diagnostique. Cette épreuve est également bien acceptée par les patients. Ils ont le sentiment de bien réussir même lorsqu'ils font deux ou trois erreurs. On peut toutefois se poser la question de la pertinence de l'utilisation de ce test par un logopédiste. A noter aussi que les items sont tirés du *BNT* (Kaplan et al., 1983), mais que les images sont différentes. Les critiques qui sont adressées à ce dernier se retrouvent ainsi également dans le matériel de la *BARD*, à savoir le manque de contrôle de facteurs psycholinguistiques considérés aujourd'hui comme pertinents.

Bibliographie

Bachy-Langedock, N. (1989). *Batterie d'examen des troubles en dénomination*. Bruxelles : Editest.

Chainay, H., Rosenthal, V., & Goldblum, M. C. (1998). Normes de dénomination de 315 images en couleur et en noir et blanc. *Revue de Neuropsychologie*, 8 (2), 179-239.

Croisile, B. (2005). Une Batterie Rapide de Dénomination (BARD) en consultation mémoire. *La Revue Francophone de Gériatrie et de Gérontologie*, 12, 526-530.

Deloche, G., & Hannequin, D. (1997). *DO 80 : Test de dénomination orale d'images*. Paris : Les Editions du Centre de Psychologie Appliquée.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia : Lea & Febiger.

Nespoulous, J.-L., Lecours, A. R., Lafond, D., Lemay, A., Puel, M., Joanette, Y., Cott, F., & Rascol, A. (1992). *Protocole Montréal-Toulouse d'examen linguistique de l'aphasie : MT 86 (version M1 bêta)*. L'Ortho Edition : Isbergues.

Snodgrass, J. G., & Vanderwart, M. A. (1980). A standardized set of 260 pictures : norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6 (2), 174-215.

Thuillard Colombo, F., & Assal, G. (1992). Adaptation française du test de dénomination de Boston. Versions abrégées. *Revue européenne de Psychologie Appliquée*, 42 (1), 67-71.

