

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

CONNAISSANCES LEXICALES ET EN MORPHOLOGIE DÉRIVATIONNELLE
CHEZ DES APPRENANTS ADULTES AVANCÉS EN FRANÇAIS LANGUE
SECONDE

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAITRISE EN LINGUISTIQUE

PAR

DORIS ALLARD

JUILLET 2008

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

*Les mots ne tiennent pas sur cette mélodie,
Je les mets sur des notes le matin, mais la nuit
Ils tombent et je découvre, épars sur le tapis,
De pauvres lettres mortes et mes idées enfuies.*

Georges Chelon, Les Mots

REMERCIEMENTS

La réalisation d'un mémoire de maîtrise demande des efforts et du temps. C'est un parcours parsemé de moments de doutes et d'errements. C'est pourquoi nous avons besoin d'être guidés. J'ai eu la chance que Mme Lori Morris, professeure au département de linguistique de l'UQAM, accepte de diriger mon travail et je dois la remercier. Son enthousiasme pour la recherche ainsi que sa confiance ont été des sources précieuses d'encouragement.

Je dois aussi saluer l'ouverture dont ont fait preuve Mme Anne-Marie Benoit, coordonnatrice responsable des cours d'appoint en grammaire à l'UQAM, ainsi que toutes les chargées de cours et professeures du cours *Connaissances de base en français écrit* à la session d'automne 2005. Leur collaboration m'a permis de réaliser une recherche intéressante auprès d'un grand nombre de sujets.

Merci enfin à mon ami Jose pour sa patience et sa compréhension tout au long de cette entreprise particulière.

TABLE DES MATIÈRES

LISTE DES TABLEAUX.....	vii
LISTE DES FIGURES.....	ix
LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES.....	xi
RÉSUMÉ.....	xii
INTRODUCTION.....	1
CHAPITRE I PROBLÉMATIQUE.....	6
CHAPITRE II CADRE THÉORIQUE.....	11
2.1 Le lexique mental.....	12
2.1.1 La morphologie et le lexique mental.....	17
2.1.2 Le lexique mental du bilingue.....	24
2.1.3 L'hypothèse des deux lexiques.....	26
2.2 La connaissance lexicale.....	30
2.2.1 La connaissance lexicale en L2.....	34
2.3 L'acquisition du vocabulaire en L2.....	39
2.3.1 Les conditions d'acquisition.....	42
2.4 La morphologie dérivationnelle : description.....	46
2.4.1 Morphologie dérivationnelle et L2.....	53

2.5	Résumé.....	57
CHAPITRE III		
	MÉTHODOLOGIE.....	60
3.1	Introduction.....	60
3.2	Les participants.....	60
3.3	Le choix et l'élaboration des instruments de mesure.....	63
3.4	L'administration des épreuves.....	71
3.5	Codage des données.....	73
	3.5.1 La tâche 1 : connaissances lexicales réceptives.....	73
	3.5.2 La tâche 2 : connaissances lexicales productives.....	74
	3.5.3 La tâche 3 : connaissances morphologiques productives.....	74
	3.5.4. La tâche 4 : mots apparentés.....	75
	3.5.5 La tâche 5 : production morphologique.....	75
3.6	Outils d'analyse.....	75
CHAPITRE IV		
	PRÉSENTATION DES RÉSULTATS.....	77
4.1	Résultats globaux.....	78
4.2	Analyses détaillées.....	84
	4.2.1 La tâche 1 : connaissances lexicales réceptives.....	84
	4.2.2 La tâche 2 : connaissances lexicales productives.....	85
	4.2.3 La Tâche 3 : connaissances morphologiques productives.....	94
	4.2.4 La tâche 4 : mots apparentés.....	101
	4.2.5 La tâche 5 : production morphologique.....	108
	4.2.6 Résultats selon le type de tâche.....	121
4.3	Les résultats au cours d'appoint en grammaire.....	124
CHAPITRE V		
	DISCUSSION.....	129
	CONCLUSION.....	138

APPENDICES.....	143
RÉFÉRENCES.....	157

LISTE DES TABLEAUX

Tableau		Page
3.1	Portrait de l'ensemble des sujets selon le sexe et le groupe linguistique	62
3.2	Langue maternelle déclarée des sujets LNN.....	63
3.3	Sommaire des tâches de l'épreuve.....	71
4.1	Comparaison des résultats, temps 1 et temps 2.....	79
4.2	Analyse de variance des facteurs groupe et version	81
4.3	Comparaison des résultats selon le groupe et selon la version du test	82
4.4	Comparaison des meilleurs résultats pour toutes les tâches.....	83
4.5	Tâche 1 : comparaison du nombre moyen d'erreurs selon le type	85
4.6	Tâche 2 : comparaison des résultats selon le type de réponse.....	87
4.7	Tâche 2: fréquence des items lexicaux.....	93
4.8	Tâche 3, version 1: comparaison des résultats selon le type de réponse.....	95
4.9	Tâche 3: fréquence des items lexicaux.....	101
4.10	Tâche 4: comparaison des erreurs selon le type.....	102
4.11	Tâche 4: fréquence des items lexicaux des vraies paires de mots....	105
4.12	Tâche 4: fréquence des items lexicaux des pseudo-paires de mots.....	108
4.13	Tâche 5: comparaison du taux de réussite et d'absence de réponse	109
4.14	Tâche 5: comparaison du taux d'erreur.....	118
4.15	Tâche 5: fréquence des items lexicaux	120

4.16	Comparaison des résultats selon le type de tâche.....	122
4.17	Comparaison des résultats selon les connaissances ciblées.....	123
4.18	Corrélations entre les tâches, LN et LNN.....	124
4.19	Barème de notation du cours de grammaire.....	125
4.20	Comparaison de la moyenne des résultats au cours de grammaire.....	127

LISTE DES FIGURES

Figure		Page
2.1	Le modèle de Kroll et Stewart.....	27
2.2	Le modèle de Paradis (2004)	28
4.1	Tâche 2, version 1: comparaison du taux de réussite par item.....	88
4.2	Tâche 2, version 1: comparaison du taux d'absence de réponse par item.....	89
4.3	Tâche 2, version 2: comparaison du taux de réussite par item.....	90
4.4	Tâche 2, version 2: comparaison du taux d'absence de réponse par item.....	91
4.5	Tâche 3, version 1: comparaison du taux de réussite par item.....	96
4.6	Tâche 3, version 1: comparaison du taux d'absence de réponse par item.....	97
4.7	Tâche 3, version 1: comparaison du taux d'erreurs item.....	98
4.8	Tâche 3: taux d'erreur et d'absence de réponse pour chaque item, LNN.....	100
4.9	Tâche 4, version 1: comparaison du taux de réussite pour les vraies paires de mots.....	103
4.10	Tâche 4, version 2: comparaison du taux de réussite pour les vraies paires de mots.....	104
4.11	Tâche 4, version 1: comparaison du taux de réussite pour les pseudo-paires de mots.....	106
4.12	Tâche 4, version 2: comparaison du taux de réussite pour les pseudo-paires de mots.....	107

4.13	Tâche 5, version 1: comparaison du taux réussite par item.....	110
4.14	Tâche 5, version 1: comparaison du taux d'absence de réponse par item.....	112
4.15	Tâche 5, version 2: comparaison du taux de réussite par item....	114
4.16	Tâche 5, version 2: comparaison du taux d'absence de réponse par item.....	115
4.17	Tâche 5: comparaison du taux de réussite par catégorie grammaticale	116
4.18	Tâche 5: comparaison du taux d'absence de réponse par catégorie grammaticale	117
4.19	Tâche 5 : comparaison du taux d'erreur selon la catégorie grammaticale.....	119
4.20	Comparaison des performances selon le type de tâche.....	123
4.21	Comparaison des résultats au cours de grammaire.....	126

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

ANOVA : abréviation anglaise de *Analysis of Variance*

GREG : Groupe de Réflexion sur l'Enseignement de la Grammaire

L1 : langue première

L2 : langue seconde

LIN-1002 : sigle du cours *Connaissances de base en grammaire du français écrit*

LN : locuteur(s) natif(s)

LNN : locuteur(s) non natif(s)

UQAM : Université du Québec à Montréal

RÉSUMÉ

Au Québec, toutes les universités francophones exigent une attestation des connaissances linguistiques aux étudiants qui désirent être admis à un programme d'études de premier cycle. Depuis dix ans, de nombreuses institutions d'enseignement supérieur ont mis sur pieds des cours d'appoint en grammaire pour les étudiants qui n'arrivent pas à satisfaire aux exigences en matière de français écrit. De plus en plus d'étudiants dont la langue maternelle n'est pas le français doivent suivre un de ces cours et plusieurs d'entre eux ne le réussissent pas.

Dans cette recherche, nous avons vérifié quelles étaient les connaissances lexicales et en morphologie dérivationnelle chez des étudiants locuteurs non natifs du français ayant échoué au test de français écrit de l'université, en posant comme hypothèse de départ que leurs connaissances seraient plutôt faibles. Pour vérifier cette hypothèse, nous avons soumis à deux reprises des sujets locuteurs natifs et locuteurs non natifs, inscrits au même cours d'appoint en grammaire, à une épreuve de cinq tâches mesurant leurs connaissances, tant en compréhension qu'en production, d'items lexicaux fréquents.

Dans l'ensemble de notre épreuve, les sujets locuteurs non natifs ont obtenu des résultats inférieurs à ceux des locuteurs natifs et les différences de performance ont été significatives dans quatre des cinq tâches. Particulièrement, les sujets locuteurs non natifs ont démontré que leurs connaissances lexicales productives étaient très faibles, ce qui soulève l'hypothèse que leurs connaissances lexicales limitées seraient à la base de leurs difficultés en morphologie dérivationnelle. Les résultats obtenus par les sujets locuteurs non natifs laissent à penser que certains d'entre eux ne possèdent peut-être pas une base lexicale suffisamment étendue pour poursuivre avec succès des études supérieures dans une université francophone.

langue seconde lexicale morphologie dérivationnelle apprenants adultes

INTRODUCTION

Dans le système d'éducation au Québec, on estime que la maîtrise de la langue d'enseignement est une habileté indispensable à la réussite des études, quel qu'en soit le niveau. L'admission à tous les programmes universitaires de premier cycle est conditionnelle à une attestation des connaissances linguistiques. Depuis 1998, les établissements d'enseignement supérieur exigent la réussite de l'épreuve uniforme de français, langue d'enseignement et littérature, menant à l'obtention du diplôme d'études collégiales. Pour les étudiants ayant terminé leurs études collégiales avant 1998, on exige d'avoir réussi le test de français écrit du ministère de l'Éducation du Québec. Pour les candidats ayant déjà fréquenté une université, la détention d'un grade universitaire d'une institution francophone ou la réussite d'un test de maîtrise du français d'un établissement universitaire québécois sont acceptés. À l'Université du Québec à Montréal, le test de français est basé sur les connaissances en français écrit que devrait posséder un étudiant de cinquième secondaire. De nombreux étudiants n'obtiennent pas la note de passage au test de français et l'UQAM, à l'instar de beaucoup d'autres universités, a mis sur pied des cours de rattrapage ou de mise à niveau, comme par exemple celui intitulé *Connaissances de base en grammaire du français écrit* (LIN-1002), pour les candidats qui ne satisfont pas aux exigences de maîtrise de la langue. Depuis 1998, la réussite du cours LIN-1002 confirme aux étudiants ayant auparavant échoué au test de français écrit qu'ils respectent désormais les exigences de la politique linguistique.

De plus en plus d'étudiants dont le français n'est pas la langue maternelle fréquentent des universités francophones au Québec. Un bon nombre de ces étudiants n'arrivent pas à satisfaire aux exigences de connaissance en français écrit

au moment de leur demande d'admission, soit parce qu'ils n'ont pas été scolarisés en français, soit parce que leur scolarisation en français n'a été que de courte durée ou insuffisante pour la maîtrise de la langue cible; ils doivent donc suivre des cours de rattrapage. À titre d'exemple, à l'UQAM, pour le seul cours LIN-1002, la proportion d'étudiants non francophones est passée de 33% en 1993 à 45% en 2003 selon le rapport du *Groupe de réflexion sur l'enseignement de la grammaire*, produit en 2005.

Nous possédons peu d'informations sur les connaissances en vocabulaire des étudiants allophones qui fréquentent les universités francophones au Québec. Par exemple, dans le test mesurant les connaissances en français écrit de l'UQAM, sur les onze sections, une seule, intitulée « choix des mots », touche le vocabulaire. Les autres sections mesurent les connaissances orthographiques et grammaticales ainsi que la ponctuation. Quant aux cours d'appoint en grammaire, ils visent surtout la maîtrise de l'orthographe grammaticale. Ainsi, les connaissances lexicales des étudiants allophones ne font l'objet d'une évaluation spécifique ni dans le test de français écrit, ni dans le cadre du cours LIN-1002, dont le contenu est axé sur des exercices d'application de règles de grammaire et qui a été conçu, à l'origine pour des francophones faibles en français écrit.

La connaissance lexicale n'est qu'une des composantes du savoir linguistique, elle est néanmoins indispensable à toute performance dans la langue cible (Cobb, 2000). La connaissance du lexique aide à l'usage de la langue cible, ce qui permet, en retour, la croissance du vocabulaire (Waring, 1998). Le vocabulaire est une composante essentielle dans l'acquisition d'une langue seconde, car les mots sont la plus importante source de sens (Laufer et Goldstein, 2004). L'acquisition du vocabulaire serait au moins aussi importante que l'acquisition de la grammaire (Jiang, 2002). Plus important encore, Schmitt et Meara (1997) ont établi qu'il y a une

corrélation entre la taille du vocabulaire L2 et la compétence globale dans la langue cible.

Nous ne savons pas dans quel contexte les étudiants allophones ont appris le français: en milieu scolaire ou naturel, dans un contexte de langue seconde ou de langue étrangère, de façon implicite ou explicite. On sait cependant que bien souvent, en classe de langue seconde, l'enseignement du lexique tient bien peu de place dans les curriculums (Tschirner, 2004): on insiste sur l'enseignement de la grammaire, mais pour ce qui est du vocabulaire, les apprenants sont laissés à eux-mêmes (Waring, 1998) et, généralement, ils investissent peu dans l'apprentissage du vocabulaire (Tschirner, 2004). Cobb et Horst (2000) ont soumis à des tests de connaissances lexicales des étudiants de niveau intermédiaire et avancé, de première et deuxième année, inscrits à un programme universitaire d'apprentissage de l'anglais L2 pour fins de communication professionnelles. Leur recherche visait à mesurer les connaissances des sujets selon les niveaux de fréquence des mots de vocabulaire. Les résultats obtenus au pré-test et au post-test ont mis en évidence le fait que les apprenants avaient fait très peu d'acquisition de mots moins fréquents entre la première et la deuxième année d'études.

Comparativement à la grammaire, système fermé, le vocabulaire d'une langue est un système ouvert dont l'apprentissage est quasi infini. Selon Waring et Nation (2004), même après avoir étudié la L2 durant plusieurs années, bien peu d'apprenants réussissent à acquérir un vocabulaire aussi étendu que les locuteurs natifs. S'il y a des adultes apprenants de L2 qui réussissent à avoir un vocabulaire aussi étendu que des locuteurs natifs scolarisés, il s'agit là de l'exception et non de la règle (Waring, 1998).

La poursuite d'études supérieures signifie nécessairement avoir à produire, ainsi qu'à lire et à comprendre, de nombreux textes. En L1, la taille du vocabulaire est un reflet du niveau de scolarité du locuteur (Waring et Nation, 2004) et la taille du vocabulaire en L2 est un facteur important dans la réussite académique générale (Saville-Troike, 1984; Laufer, 1992; Morris et Cobb, 2004; Laufer et Goldstein, 2004). Tschirner (2004) a mené une étude sur les connaissances lexicales en L2 auprès d'étudiants universitaires inscrits au programme de langue et littérature anglaises. Les résultats obtenus font dire à l'auteur que, même si ces étudiants avaient étudié la L2 en moyenne durant huit ans et qu'ils avaient réussi le test d'admission au programme, ils n'étaient pas suffisamment préparés pour entreprendre des études universitaires dans le domaine choisi à cause de leurs faiblesses en vocabulaire.

En français, la grande majorité des mots sont des mots construits par dérivation (Rey-Debove, 1984). La dérivation est un processus qui consiste à former des mots par l'addition d'un ou de plusieurs affixes à une base, selon des règles bien précises. Dans les mots dérivés, la base ainsi que les affixes sont tous porteurs de sens. Comparativement aux mots simples, non construits, les mots dérivés peuvent être analysés par la décomposition de leurs parties. La connaissance des principes de la morphologie dérivationnelle et la capacité d'analyse des mots construits confère donc un avantage au locuteur face à un mot qui ne lui est pas familier. En effet, la morphologie dérivationnelle, si elle est complexe, contribue aux compétences lexicales parce que ses règles impliquent de la redondance (Frost et Grainger, 2000). La langue française est riche sur le plan de la morphologie et une bonne connaissance de cette langue suppose donc la maîtrise des principes de base en morphologie dérivationnelle.

À notre connaissance, peu d'études ont porté sur les connaissances lexicales en français langue seconde chez des étudiants universitaires. La connaissance lexicale

est complexe et il ne serait pas possible d'en mesurer tous les aspects dans une seule et même étude. Dans cette recherche, nous avons voulu savoir quelles étaient les connaissances lexicales d'étudiants non francophones faibles en français écrit. Plus précisément, nous souhaitons mesurer certains aspects de la connaissance lexicale et de la morphologie dérivationnelle, sur les plans réceptif et productif, chez des étudiants dont le français n'est pas la langue maternelle, inscrits à un cours de rattrapage en grammaire du français dans une université francophone. Nous avons voulu savoir s'il y avait une corrélation entre leurs résultats aux tests lexicaux et la note obtenue au cours de rattrapage. Enfin, nous avons comparé leurs performances avec celles des étudiants francophones faibles en français écrits qui suivaient ce même cours.

Dans le prochain chapitre, nous décrivons le problème spécifique auquel nous nous sommes intéressée et nous établissons les questions principales à la base de notre travail. Le chapitre trois sera consacré aux concepts et aux aspects théoriques qui sous-tendent cette recherche. Au chapitre quatre, nous présenterons la démarche méthodologique adoptée alors que les résultats détaillés feront l'objet du chapitre cinq. Enfin, la discussion des résultats obtenus sera présentée au sixième chapitre.

CHAPITRE I

PROBLÉMATIQUE

Au cours des dix dernières années, le nombre d'étudiants ayant échoué au test de français requis pour l'admission définitive à un programme d'études à l'UQAM est demeuré important. Le taux d'échec au test d'admission en français a justifié la mise sur pied de plusieurs cours correctifs en français écrit afin d'aider les étudiants à remédier à leurs lacunes dans la maîtrise de la grammaire de base. Ces cours de rattrapage et de mise à niveau sont centraux dans la Politique de la langue française de l'UQAM. Comme le rapportait le *Groupe de réflexion sur l'enseignement de la grammaire* en 2005, l'achalandage à ces cours représente une part importante des crédits d'enseignements dispensés chaque année par le Département de linguistique et de didactique des langues de l'UQAM. Par exemple, il y a en moyenne, quelque 700 inscriptions par année académique pour le seul cours *Connaissances de base en grammaire du français écrit* (LIN-1002).

Le test de français écrit de l'Université évalue la connaissance des règles de grammaire de base de l'écrit. Le cours LIN-1002 est lui aussi axé sur la maîtrise de la grammaire de base. D'année en année, le taux de réussite à ce cours se situe à près de 60%. Les taux d'échec au test de français de l'Université ainsi qu'au cours LIN-1002 démontrent effectivement que beaucoup d'étudiants sont faibles en

grammaire écrite. Parmi ces étudiants, il se trouve des francophones et des non-francophones et nous ne savons pas si les causes d'échec sont les mêmes pour les deux groupes linguistiques. Si les étudiants non francophones sont capables de s'exprimer oralement en français, cela ne garantit pas que leurs compétences sont solides (Koda 1999). Et qu'en est-il de leurs connaissances lexicales? Se pourrait-il que ces étudiants aient aussi des faiblesses sur ce plan? Est-ce que ces faiblesses lexicales pourraient avoir un impact sur les performances grammaticales des étudiants?

La proportion d'étudiants non francophones dans les cours de rattrapage en français est importante à l'UQAM. Leur présence dans des cours conçus à l'origine pour des locuteurs natifs soulève quelques questions. Selon les auteurs du *Rapport du groupe de réflexion sur l'enseignement de la grammaire* du Département de linguistique et de didactique des langues de l'Université du Québec à Montréal (2005), les difficultés en français écrit des étudiants non francophones aux cours de rattrapage en grammaire sont, généralement, différentes de celles des francophones; le taux d'abandon serait par ailleurs plus important chez les étudiants allophones. Comme le soulignent les membres du Groupe de réflexion sur l'enseignement de la grammaire (GREG) à propos du cours LIN-1002, la méthodologie présuppose que les étudiants allophones possèdent la même intuition linguistique que les locuteurs natifs du français. Si l'on suppose que des faiblesses sur le plan lexical pourraient avoir un lien avec les difficultés en grammaire chez les étudiants, comment se situent les étudiants allophones à cet égard et comment se comparent-ils aux étudiants francophones ?

On décrit couramment les connaissances lexicales selon deux grandes dimensions : la taille ou l'étendue, c'est-à-dire, le nombre d'items connus; la profondeur, soit la connaissance des diverses propriétés d'un item lexical, par exemple, sur le plan syntaxique, sémantique, morphologique, connotatif. Beaucoup d'études sur les

connaissances lexicales ont porté sur la taille du vocabulaire (Laufer, 1992, Nation et Newton, 1997; Liu et Nation, 1985; Hu et Nation, 2000; Nation, 2001). En langue seconde, des chercheurs se sont demandé, notamment, de combien de mots doit disposer un locuteur pour être en mesure de réaliser avec succès une tâche (Laufer et Yano, 2001; Waring et Nation, 2004; Huckin et Coady, 1999). On sait par ailleurs qu'il faut avoir atteint un niveau-seuil en termes d'étendue de vocabulaire pour que les habiletés acquises en L1 puissent être transférées (Laufer, 1997; Cobb, 1999; Tschirner, 2004).

On distingue deux grands domaines dans la morphologie en français: la morphologie flexionnelle et la morphologie dérivationnelle ou lexicale. Alors que la flexion ajoute des traits grammaticaux à une racine sans créer un mot nouveau, le procédé de dérivation consiste à créer de nouveaux mots par l'addition d'affixes à une racine ou base. La dérivation est un procédé très productif en français et le lexique compte une proportion importante de mots dérivés, d'où l'importance d'en connaître les principes. Par exemple, le suffixe *-ion*, qui sert à former des noms, se retrouve dans quelques 2,400 mots dérivés (Huot, 2001). Les connaissances en morphologie dérivationnelle contribuent à la fois à l'étendue et à la profondeur du savoir lexical et en langue seconde, on sait maintenant que la connaissance de la morphologie joue un rôle dans la reconnaissance lexicale (Frost et Grainger, 2000).

Nous avons vu que les connaissances lexicales et morphologiques ne font l'objet d'aucune évaluation à l'examen d'admission à l'Université et que les cours de rattrapage pour les étudiants faibles en français écrit sont axés sur la maîtrise de la grammaire de base. La connaissance lexicale est une composante importante des compétences en L2 et il nous a semblé important de connaître quelle était cette connaissance chez des étudiants allophones, à l'aise avec la L2, mais faibles en français écrit.

Questions de recherche

- ◆ Quelles sont les connaissances lexicales et en morphologie dérivationnelle chez des étudiants dont le français n'est pas la langue maternelle désirant entreprendre des études supérieures dans la L2?
- ◆ Comment se comparent des étudiants non francophones et francophones, faibles en français écrit, au chapitre des connaissances lexicales et morphologie dérivationnelle?

Objectifs spécifiques

Il serait très ambitieux et peu réaliste de vouloir évaluer toutes les composantes de la compétence lexicale dans une seule et même étude. Nous avons ciblé certains aspects de la taille et de la profondeur des connaissances en vocabulaire, sur les plans productif et réceptif, dans le but de contribuer à une meilleure compréhension de cette question.

Certains mots sont plus fréquents que d'autres dans la langue parlée et dans la langue écrite. Des chercheurs se sont intéressés non seulement à l'étendue du vocabulaire, mais aussi au rang de fréquence des items connus, soit simplement pour mesurer le degré de sophistication du vocabulaire des locuteurs, soit pour expliquer le niveau de performance dans une tâche donnée (Cobb, 1999; Nation, 1993). Dans notre recherche, nous avons balisé l'évaluation de l'étendue du vocabulaire par la fréquence des items lexicaux ciblés.

En ce qui a trait à la morphologie dérivationnelle, nous avons limité notre recherche aux suffixes parce que les préfixes et les suffixes ne sont pas semblables sur le plan formel (Hay, 2001) et qu'ils font appel à des procédés de niveaux différents (Slobin,

1973). Nous avons voulu vérifier la connaissance des principes suivants: 1) plusieurs mots peuvent partager une même base et avoir une signification en commun : ils font alors partie de la même famille de mots; 2) les suffixes déterminent la catégorie grammaticale du mot dérivé; 3) certaines structures comme *déroule+ure* sont impossibles en français, alors que d'autres comme *double+ure* sont tout à fait légales. Et puis, toutes les formes dérivées n'ont pas le même degré de complexité (Pinker, 1994; Caramazza, 1997) : le processus de dérivation d'un mot peut entraîner des changements formels. La transparence dans la dérivation joue-t-elle un rôle dans la reconnaissance des mots chez les apprenants adultes ?

Et puis, nous souhaitons savoir s'il existait une corrélation entre les connaissances lexicales et en morphologie dérivationnelle et les compétences grammaticales chez nos sujets. Enfin, nous voulions vérifier s'il y avait des différences dans les connaissances lexicales et en morphologie dérivationnelle entre des apprenants de L2 relativement avancés et des locuteurs natifs faibles en français écrit et identifier quelles étaient ces différences.

Dans le chapitre qui suit, nous présenterons les différents aspects qui ont constitué notre cadre théorique. Nous verrons tout d'abord le concept de lexique mental dont nous décrirons le contenu et le fonctionnement. Suivra une section faisant état des différentes propositions théoriques sur la place qu'occupe la morphologie dans le lexique mental. Dans la section consacrée au lexique mental du bilingue, il sera question des différentes discussions autour de son organisation et nous présenterons le modèle théorique que nous avons retenu. Par la suite, nous traiterons de la connaissance lexicale en L1 puis en L2, de l'acquisition du vocabulaire d'une langue seconde et, finalement, nous décrirons dans ses grandes lignes la morphologie dérivationnelle du français et sa contribution dans la compréhension et l'acquisition du vocabulaire chez les apprenants du français L2.

CHAPITRE II

CADRE THÉORIQUE

Il sera question dans ce chapitre des éléments théoriques qui nous apparaissent nécessaires à la définition de l'objet de notre étude, ainsi que des résultats de recherches qui ont orienté notre travail.

Nous tenterons tout d'abord de comprendre comment le lexique est emmagasiné en mémoire, ce que contient le lexique mental et comment sont gérées ces informations. Ensuite, nous verrons quel est le rôle de la morphologie dans le traitement linguistique et nous passerons en revue les principales théories sur l'intégration de la morphologie dans le lexique mental. Puisque notre recherche s'intéresse aux locuteurs non natifs, la section suivante abordera des questions qui se posent sur le fonctionnement du lexique mental chez un bilingue. Puis, nous verrons que la connaissance lexicale, c'est beaucoup plus que de savoir un mot et nous en détaillerons les composantes. Ensuite, il sera question de ce qui caractérise le savoir lexical du bilingue et de quelques recherches ayant établi le niveau de connaissance requis chez un adulte apprenant de L2. Nous toucherons par la suite aux processus d'acquisition du vocabulaire de L2 chez l'adulte ainsi qu'aux conditions d'apprentissage et nous verrons, notamment, comment elles peuvent influencer l'acquisition. Enfin, l'avant-dernière section sera consacrée à la description des grands principes de la morphologie dérivationnelle en français et de

l'atout que constitue leur connaissance dans la compréhension des mots complexes. Nous terminerons par la présentation de recherches mettant en parallèle l'application de l'analyse morphologique et la compétence en L2.

2.1 Le lexique mental

Le langage, au même titre que les autres facultés cognitives, fait appel à la mémoire : la mémoire enregistre des représentations linguistiques qui sont le résultat de traitements cognitifs. Toutes nos connaissances à propos des mots ainsi que les concepts que représentent ces mots sont stockés en mémoire dans ce qui est appelé le *lexique mental* (Schreuder et Flores d'Arcais, 1989). On conçoit globalement la structure du lexique mental comme étant constituée de deux grands systèmes : un système linguistique (la forme, le sens) et un système non linguistique (les concepts) en constante relation (Tréville, 2000). Tous les modèles théoriques sur le lexique mental ont pour base les relations des unités lexicales entre elles, fondées sur leurs traits caractéristiques (Tréville, 2000), et les relations entre les unités lexicales et les concepts (Schreuder et Flores D'Arcais 1989).

« The general idea underlying the various theories and approaches is that the various lexical units, and/or the underlying concepts, are represented mentally as a set of elements or entries bearing some relations to one another. All theories make, implicitly or explicitly, some assumptions about the organization of the units.» (Schreuder et Flores D'Arcais, 1989, p.411).

L'étude du lexique mental permet de cerner la nature des connaissances enregistrées, de vérifier des hypothèses sur l'acquisition du langage, son mode d'organisation et son usage (Libben et Jarema, 2002). Le lexique mental est vu comme étant l'ossature même de l'habileté langagière (Libben et Jarena, 2002), car il possède la capacité d'intégrer des informations et de les classer, de les retrouver et de les conserver en mémoire à travers le temps, cela sans qu'il n'y ait nécessairement d'intervention consciente (Libben, Jarena, 2002).

On utilise les termes « unités » ou « entrées » lexicales plutôt que le terme habituel « mot » pour nommer les représentations stockées dans le lexique mental. L'unité lexicale peut être simple ou complexe, constituée d'une ou de plusieurs parties : mots composés, locutions, expressions, marqueurs, phrases syntagmatiques. (Tréville, 2000). L'entrée lexicale correspond, en quelque sorte, à l'adresse précise d'un mot dans la mémoire (Pinker, 1999) et c'est l'encodage des informations qui lui attribuerait son identité. Chaque entrée contient au moins quatre composantes informationnelles: le sens, la catégorie grammaticale, la structure morphologique et la forme phonologique et graphique (Clark, 1993). Évidemment, il arrive que cette information soit incomplète, dans le cas, par exemple, d'un mot nouveau. Certains auteurs définissent le contenu de l'entrée lexicale comme étant le lemme (*lemma*) (aspects sémantique et syntaxique) et le lexème (composantes formelles) (Clark, 1993; Caramazza, 1997; Jiang, 2000; Hillis, 2000). L'aspect sémantique joue un rôle central dans l'établissement d'une unité lexicale (Tréville, 2000) et dans son intégration au lexique mental : un mot dont le sens n'est pas connu n'est pas tout à fait un mot (Jiang, 2000).

La nature ainsi que la forme de la représentation lexicale seraient déterminantes pour le processus d'accès au lexique mental (Henderson, 1989). Il semble que l'efficacité du rappel ou de la remémoration dépende de la forme spécifique de l'encodage (Thomson et Tulvig, 1970) et on suppose que retrouver un item en mémoire ce n'est, ni plus ni moins, qu'une récapitulation du processus d'encodage (Mulligan et Lorito, 2006). L'encodage peut être imparfait ou incomplet et cela aura un impact sur l'accès et le rappel. L'accès suppose aussi une certaine organisation des composantes de cette vaste base de données. On reconnaît généralement que les informations sémantiques, syntaxiques et lexicales sont des ensembles indépendants, de niveaux différents (Hillis, 2000). Dans le modèle de réseaux indépendants (*independent network model*) de Caramazza (1997) par exemple, les connaissances lexicales sont organisées en une série de réseaux distincts (représentations orthographiques, phonologiques, syntaxiques, sémantiques) qui

sont connectés entre eux par un nœud (*node*) spécifique à la modalité (orale ou écrite). Le réseau syntaxique comprend des sous-réseaux avec des nœuds pour chaque aspect: genre, catégorie grammaticale et ainsi de suite. Les réseaux orthographique et phonologique contiennent des représentations des items lexicaux (les racines) spécifiques à la modalité. À l'intérieur du réseau, l'activation d'un nœud empêche l'activation des autres candidats. Les nœuds de ces réseaux ont aussi pour effet d'inhiber les items en compétition. Encore ici, nous voyons l'importance de la qualité de l'encodage : ce sont les composantes de chaque item qui feront en sorte qu'il sera activé ou pas.

L'accès au lexique mental serait aussi conditionné par la fréquence des mots : les mots les plus fréquents auraient une plus grande force latente d'activation. Lorsqu'en production une représentation sémantique active plusieurs candidats, le mot le plus fréquent aurait plus de possibilités d'être sélectionné (Hillis, 2000). Dans le cas d'un apprenant d'une L2, on peut penser que la fréquence d'un item lexical donné, conséquente à son expérience linguistique, n'est pas nécessairement la même que la fréquence de cet item dans le lexique de la langue cible.

L'existence de liens entre les différents systèmes de même que la spécificité des réseaux dans le lexique mental peuvent expliquer les performances et les erreurs dans des tâches linguistiques données. Ainsi, la distinction minimale entre deux pôles d'information – lemme et lexème - pour toute entrée lexicale fournirait une explication à notre capacité à différencier deux mots dont la forme est identique, mais le sens différent (les homophones), de même qu'elle éclairerait le phénomène du « mot sur le bout de la langue », alors que le lemme est disponible, mais que le lexème fait défaut (Caramazza, 1997).

L'organisation du lexique mental est centrale pour le fonctionnement du langage et selon Meara (1996b), elle est tout aussi importante que les items eux-mêmes : c'est ce qui permet au locuteur, à partir de qui est entreposé en mémoire, d'accomplir une tâche donnée : par exemple, identifier des mots en situation de compréhension, ou encore, choisir les items pertinents pour le sens du message qu'il veut transmettre (Clark, 1993). Chaque mot peut avoir des liens avec plusieurs réseaux et il y a dans le lexique mental plusieurs types de classification (Tréville, 2000), lesquelles permettraient de réaliser différentes tâches. Il semble qu'en production, par exemple, c'est une organisation sémantique qui serait sollicitée alors qu'en compréhension, ce serait la forme, notamment les premières lettres du mot, qui servirait de repère (Clark, 1993; Henderson, 1989). En écriture, différents systèmes seraient sollicités selon qu'il s'agit d'une production libre ou d'une dictée (Hillis, 2000).

Les processus mentaux nécessaires à la réalisation d'une tâche vont varier selon qu'il s'agit de compréhension ou de production. Ainsi, la reconnaissance d'un mot demande d'associer les traits formels d'une représentation enregistrée en mémoire à l'information générée par le segment graphique ou sonore (Grainger, 1993). En production, le choix d'un item lexical nécessite la sélection d'un certain nombre de caractéristiques ou traits pour activer une unité lexicale pleinement fonctionnelle (Tréville, 2000). Comme il y aurait indépendance fonctionnelle entre les représentations sémantiques et formelles, l'accès aux lexiques phonologiques et orthographiques se ferait par des procédures distinctes pour la production et la compréhension (Hillis, 2000).

Dans l'acquisition, le processus de lexicalisation, c'est-à-dire la consolidation des traits qui ont été encodés pour constituer une unité lexicale, va permettre l'intégration des nouvelles entrées dans le lexique mental (Dumay et al., 2004). Ces données seront classées et mises en relation avec les autres, deux opérations fondamentales.

L'acquisition de nouveaux éléments lexicaux et leur survie dans la mémoire consciente seraient conditionnées par la création, entre eux et avec les réseaux préexistants, de liens puissants (sémantiques, morphologiques, psychologiques et autres). (Tréville, 2000, p. 44).

L'acquisition de nouveaux éléments a donc des effets sur l'organisation du lexique mental : les liens entre les représentations peuvent se modifier en intensité avec l'intégration de nouveaux mots ou de nouveaux concepts (Tréville, 2000). Le locuteur qui développe son vocabulaire dans sa langue maternelle acquiert continuellement des connaissances sur les mots qu'il possède déjà: c'est ainsi que les représentations changent elles aussi et la construction continue du savoir lexical se fait donc en intégrant des informations nouvelles à ce qui est déjà connu, enregistré et classifié (Clark, 1993). On peut penser qu'il en va de même pour le locuteur adulte apprenant d'une langue seconde au fur et à mesure de son apprentissage et de son exposition à la L2.

Des auteurs font l'hypothèse d'un autre système impliqué dans l'acquisition des items lexicaux, d'un autre processus qui pourrait être antérieur à l'enregistrement en mémoire long terme. Jiang (2000) et Jiang et Forster (2001) proposent qu'il existe une mémoire épisodique, non lexicale, qui a pour fonction d'enregistrer les événements, en plus d'une mémoire sémantique, lexicale. Selon cette conception, l'information lexicale peut être représentée dans les deux systèmes. Les items acquis de façon récente sont représentés en mémoire épisodique, mais pas encore intégrés au lexique mental. Lorsqu'un mot est appris, il est enregistré dans le lexique mental, mais il resterait des traces dans la mémoire épisodique. Il semble que la lexicalisation ne soit pas instantanée puisqu'il faudrait que soient consolidées les données épisodiques (Dumay et al., 2004). Enfin, le rappel en mémoire épisodique (aussi désignée comme étant la mémoire de l'action) demanderait un effort conscient, ce qui n'est pas le cas pour la mémoire sémantique (ou mémoire verbale) (Kormi-Nouri et al., 2003).

Nous avons vu que la qualité de l'encodage, l'intégration dans un réseau et la fréquence des unités lexicales conditionnent l'accès au lexique mental et que différentes composantes seront sollicitées selon la tâche à accomplir. Le lexique mental doit intégrer et gérer une grande quantité d'informations, et ce, pour chaque entrée lexicale. Les mots morphologiquement complexes constituent une part importante du lexique en français. Il sera question dans la section suivante de la façon dont la morphologie des unités lexicales est enregistrée en mémoire, du rôle de la morphologie dans le traitement langagier et des principales théories à cet égard.

2.1.1 La morphologie et le lexique mental

La morphologie fait partie des informations comprises dans l'entrée lexicale stockée en mémoire et plusieurs recherches ont établi qu'elle joue un rôle dans le traitement langagier. Des expériences utilisant la technique d'amorçage ont démontré que la reconnaissance du mot ciblé était facilitée par la présentation d'un mot amorce dérivé. Longtin et Meunier (2005), dans une étude en français utilisant en amorçage des dérivés et des pseudos dérivés, ont démontré que la morphologie joue un rôle dans la reconnaissance lexicale. Dans une étude comparant les effets de deux types d'amorçage, on a trouvé que le temps de réponse était plus rapide avec une amorce morphologique qu'avec une amorce sémantique (McQueen et Cutler, 1998). Beauvillain (1996), dans une recherche réalisée auprès d'adultes francophones a mesuré le temps de fixation visuelle des mots et a conclu que la morphologie est utilisée pour l'intégration de l'information dans une tâche de lecture. Selon ce qu'ont trouvé Rastle et al., (2000), la reconnaissance des mots passe par un niveau de représentation où les morphèmes sont traités différemment des mots entiers.

S'il est certain que la morphologie a un effet sur le traitement langagier (Frost et al., 2005), quel est son rôle et comment est-elle emmagasinée dans le lexique mental ?

Selon différentes études menées en langue maternelle, le lexique mental possède une organisation morphologique (Lecoq et al., 1996; McQueen et Cutler, 1998). D'autres auteurs sont d'avis que la morphologie ne jouerait qu'un rôle indirect dans l'organisation du lexique mental.

« It remains to be seen, however, whether morphology needs to be explicitly represented in the system or whether morphological effects emerge from interactions within and between units representing form and meaning. » (Frost et al., 2005).

Ce questionnement nous amène à présenter les deux visions principales sur le traitement et l'organisation des mots morphologiquement complexes, selon que c'est par le mot entier que se fait l'accès à ses propriétés lexicales et syntaxiques, ou selon que l'on considère que le mot complexe doit être réassemblé, ce qui suppose l'accès à ses composantes. La première position (Butterworth, 1983; Bybee, 1988) dite non compositionnelle (ou *full listing*), considère que ce sont les mots entiers qui sont enregistrés dans le lexique mental et que chaque mot morphologiquement complexe est inscrit indépendamment des autres dérivés de la même famille, sans encodage de la structure morphologique. De ce fait, les formes simples et complexes sont traitées de la même façon et la reconnaissance se produit quand il y a correspondance exacte entre l'item de l'input et l'item du lexique mental. Selon la conception de Bybee dans son *Lexical network model* (1988), chaque mot est une association de propriétés sémantiques et phonologiques et la force des relations entre les mots du lexique mental va dépendre de la force des liens entre des caractéristiques communes. Selon Bybee (1988) la rapidité dans la reconnaissance d'un mot est liée à sa fréquence et la force lexicale (*lexical strength*), attribuable à la fréquence, expliquerait la persistance en mémoire de formes complexes irrégulières, mais couramment utilisées. Face à des mots complexes non familiers, il y aurait cependant une forme d'analyse primaire qui servirait de procédé de vérification (Hankamer, 1989), une identification de ses composantes.

Une autre vision non compositionnelle se retrouve chez les tenants de la théorie connexionniste de l'organisation du lexique mental (*parallel distributed processing*). Dans cette conception, on considère que la morphologie n'est pas explicitement représentée (Raveh, 2002). Les mots peuvent être reconnus sans avoir recours à la morphologie, dans la mesure où l'orthographe et la phonologie nous fournissent suffisamment d'information (Seidenberg et Gonnerman, 2000; Raveh, 2002). Le « traitement » morphologique traduit une sensibilité aux relations systématiques entre les structures de surface et leur sens (Plaut et Gonnerman, 2000) et les notions de morphologie émergent par l'apprentissage des régularités que celle-ci produit dans le lexique (Frost et al., 2000). Certains, comme Plaut et Gonnerman (2000), reconnaissent un certain traitement morphologique qu'ils désignent comme un niveau intermédiaire ou de médiation situé entre les représentations orthographiques et sémantiques.

Selon le point de vue compositionnel (ou *root-listing*) cette fois, les mots sont décomposés en unités morphémiques (racines ou bases et affixes). Dans le modèle de Taft et Forster (1975) et de Taft (1991), les unités morphémiques ont des représentations indépendantes, mais sont interreliées et tous les dérivés d'une même base partagent une même entrée dans le lexique mental. La structure morphologique est perçue dès la reconnaissance d'un mot et elle est utilisée dans l'accès au vocabulaire stocké en mémoire. L'identification d'un mot se fait par la première syllabe de la racine, la racine étant isolée par un procédé d'analyse de gauche à droite et, après vérification de la racine dans le lexique mental, s'opère un processus de vérification des affixes. Le nœud morphémique de la racine est la plus importante source d'activation, car c'est ce qui va fournir les indices les plus pertinents pour ensuite restreindre la sélection du nœud conceptuel approprié.

« For example, the HAND morpheme node will be activated whether the presented word is HAND, HANDS, HANDED, HANDY, HANDBOOK, or UNDERHAND, and therefore, one would need to determine which concept node was more active in order to determine which of these words was the one presented. The appropriate concept node will become more active than any others on the basis of activation passing up from the nodes representing the rest of the word (e.g., the morpheme node for ED, BOOK, UNDER, etc.). (Taft, 1991, p. 115).

Les auteurs considèrent que la décomposition est le seul chemin d'accès possible pour les mots complexes: c'est la théorie dite de *single route*. Pour Taft (1991), il ne fait pas de doute que les représentations ont des liens entre elles dans le lexique mental; cependant, la nature de ces liens est différente selon qu'il s'agit d'items d'un même paradigme morphologique ou d'une famille sémantique sans relations morphologiques.

Un des arguments plaçant en faveur de l'hypothèse de la compositionnalité vient de résultats d'expériences avec des pseudo-mots. Longtin et al. (2003) ont comparé l'effet d'amorçage sur le temps de réponse des participants confrontés à des paires de mots en français avec ou sans lien morphologique. Leurs résultats appuient l'idée qu'un processus de décomposition joue dans la reconnaissance des items lexicaux complexes et que ce n'est pas seulement un phénomène de similarité orthographique.

« Our results suggest that in masked conditions a decomposition process takes place, inducing activation of the corresponding morphemic and pseudo-morphemic constituents. If it is actually the case, then the observed priming effect is a consequence of the decomposition process and not the reflection of a relationship between the prime and the target in the mental lexicon. » (Longtin et al., 2003, p.328).

Des chercheurs ont fait valoir que toutes les formes construites ne sont pas égales en termes de complexité, ce qui impliquerait que tous les mots morphologiquement complexes ne feraient pas appel aux mêmes traitements : cette vision hybride se

retrouve chez Pinker (1994) qui propose une théorie dite de « règle et mémoire associative » dans laquelle les formes régulières seraient générées, par défaut, par un système qui applique la règle aux racines qui ne sont pas irrégulières; les formes irrégulières doivent être mémorisées. Les formes irrégulières seraient affectées par les propriétés de la mémoire associative, telles que la fréquence et la similarité, cette dernière caractéristique permettant une forme de généralisation par analogie. Dans la conception d'Ullman (2001), qui met de l'avant un modèle déclaratif/procédural, la mémorisation et l'usage des mots simples (donc, non construits ou composés) dépendraient d'une mémoire associative des représentations, alors que la mémoire procédurale aurait comme spécificité de gérer les règles de grammaire dans la construction séquentielle et hiérarchique des structures de formes complexes. Les transformations morphologiques pourraient être comprises et générées par l'un ou l'autre système selon la nature des changements et selon qu'il s'agisse d'une règle de construction plus ou moins productive.

Parmi les modèles qui proposent que l'accès aux mots morphologiquement complexes pourrait se faire de deux façons (*dual route*) selon leur nature, le plus complet nous semble être l'*Augmented Addressed Morphology* (AAM) de Caramazza et al., (1988) et Chialant et Caramazza (1995). Selon les auteurs, l'hypothèse non compositionnelle n'explique ni la compréhension et la reconnaissance des mots inconnus, ni le jugement sur les pseudo-mots. Au départ, ils postulent que les mots d'une même famille partagent une représentation lexicale commune. Ils appuient cette affirmation sur des recherches ayant démontré que dans des tâches de décision lexicale, le temps de réponse était significativement plus court lorsqu'il s'agissait de paires de mots morphologiquement apparentés. Ils conçoivent aussi des entrées morphémiques distinctes pour les racines et les affixes. Cependant, les auteurs sont d'avis que les connaissances morphologiques ne sont pas activées à chaque rencontre d'un mot complexe, que ce soit à l'oral ou à l'écrit. Contrairement au modèle de Taft et Forster (1975), la décomposition morphologique n'est pas le chemin d'accès obligé.

Chialant et Caramazza (1995) font observer que certaines formes répondent à des règles très peu fréquentes qui ne s'appliquent qu'à un nombre restreint de mots. Le processus systématique de décomposition serait alors peu efficace. Le principe de productivité des affixes aurait donc un rôle à jouer dans le traitement lexical. Les auteurs émettent l'hypothèse que les formes irrégulières et régulières font appel à des traitements différents et proposent que les divers types de formes morphologiques sont représentés et traités de façon différente, car tous les niveaux d'information morphologique ne sont pas d'égale pertinence pour l'encodage et l'analyse lexicale. Le modèle AAM reconnaît deux modes de traitement des unités lexicales : global et décomposé. Ainsi, dans les premières phases du traitement d'un mot, le locuteur s'appuie sur ce qui est explicite dans la structure de surface, c'est-à-dire, l'information orthographique de l'input, cependant, en l'absence d'indices sur la racine, il ne peut y avoir de décomposition lexicale.

Le modèle AAM emprunte aux conceptions non compositionnelles la notion de similarité : l'activation d'une unité lexicale sera proportionnelle à la similarité entre l'input et la représentation en mémoire et s'il y a transparence orthographique, cela permet au principe de similarité de s'appliquer. Pour les formes transparentes, les deux accès - global et décomposé - seront activés et le degré d'activation sera conséquent à la fréquence de l'item. L'accès aux mots morphologiquement complexes se ferait d'une façon globale pour les mots connus et fréquents et de façon décomposée pour les mots non familiers mais réguliers; seuls les affixes et les racines très productifs seraient représentés sous une forme décomposée. Ainsi la fréquence et la transparence seraient des facteurs déterminants pour le traitement. Enfin, selon cette conception, l'activation de la représentation du mot entier serait plus rapide que l'activation de ses composantes morphémiques.

« [...] for all orthographically transparent forms both whole-word and morpheme-sized access units will be activated, to an extent which is directly proportional to the frequency of the access unit. » (Chialant et Caramazza, 1995, p 63).

Longtin et al. (2003) sont aussi d'avis que la transparence est un facteur influençant le traitement des mots morphologiquement complexes : la décomposition ne serait utile que pour les dérivés transparents. D'autre part, la transparence sémantique aurait un rôle à jouer dans l'organisation du lexique mental et les mots morphologiquement opaques ne seraient pas nécessairement reliés à la famille morphologique dans le lexique mental d'un locuteur, alors que les dérivés transparents le seraient. Il en découle qu'il pourrait y avoir représentation non décomposée dans le cas des mots opaques.

Des chercheurs ont tenté d'établir des principes généraux pour les représentations et pour le traitement lexical qui pourraient s'appliquer à toutes les langues humaines, même si l'on sait que chaque langue a ses particularités morphologiques, syntaxiques et phonologiques et que l'on suppose que cela implique un traitement cognitif particulier (Libben et Jarena, 2002). Clark (1993) avance l'idée que les deux modes de traitement de la morphologie pourraient être valables et efficaces selon la typologie de la langue. Le processus non décompositionnel ne serait pas adéquat dans le cas de langues morphologiquement complexes comme le turc, étant donné l'énorme quantité d'items lexicaux différents qu'il faudrait enregistrer en mémoire (Clark, 1993; Hankamer, 1989). Marslen-Wilson (2001) a trouvé que dans les langues sémitiques, la représentation morphologique jouerait un rôle structurel fondamental, alors que ce ne serait pas le cas pour le mandarin qui ne possède pas de principes combinatoires pour construire des mots.

Il se pourrait que pour des langues morphologiquement complexes, l'analyse morphologique soit, par défaut et de façon automatique, appliquée par le locuteur (Longtin et al., 2003). Longtin et Meunier (2005) et Longtin et al. (2003) sont d'avis que dans une langue comme le français, la décomposition survient très tôt dans la reconnaissance lexicale et que c'est un processus qui se produit automatiquement face à tout mot qui semble morphologiquement construit. Cela démontre que le

cerveau humain possède les facultés nécessaires au traitement de mots dont la construction est complexe.

« The existence of languages with agglutinative morphology indicates that the human mind has, or is capable of acquiring, a parsing mechanism that allows the recognition and understanding of words of impressive complexity. » (Hankamer, 1989, p. 405).

Les deux sections précédentes ont permis de décrire le contenu du lexique mental et à expliquer son fonctionnement chez un locuteur natif. Nous savons que la plupart des êtres humains ont la capacité d'apprendre deux, voire plusieurs langues. Cela implique que le cerveau peut gérer un lexique mental autre que celui de la langue maternelle. Nous verrons dans la section suivante comment fonctionne le lexique mental chez la personne bilingue et en quoi il est différent.

2.1.2 Le lexique mental du bilingue

Selon Paradis (2004), le cerveau humain possède tout ce qu'il faut pour comprendre et produire des énoncés dans une L2. Des auteurs estiment qu'il n'y a pas de différences qualitatives entre le cerveau d'un monolingue et celui d'un bilingue et dans les deux situations, les activités cérébrales doivent devenir le plus automatiques possibles (Franceschini et al., 2003). Nous avons vu dans les sections précédentes que le lexique mental est une structure centrale de la langue, mais une structure complexe qui gère plusieurs réseaux. Que se passe-t-il chez le locuteur qui acquiert ou a acquis une deuxième langue, qui doit utiliser un deuxième lexique? Comment gère-t-il les lexiques des deux langues et quelles sont les difficultés que cela peut poser? Grosjean (1998) fait remarquer qu'un bilingue, ce n'est pas deux monolingues dans une personne. Les bilingues sont différents des monolingues par le fait qu'il y a contact et interactions entre les deux langues. Comment est organisé le lexique mental du bilingue, est-il différent de celui du monolingue?

Des différentes discussions portant sur l'organisation du lexique mental du bilingue, il semble qu'une constatation ait fait l'accord parmi les chercheurs: tout comme nous l'avons vu pour le lexique mental en L1, il faut distinguer les composantes conceptuelles des composantes lexicales. Le système conceptuel est commun aux deux langues, alors que les représentations lexicales sont spécifiques à l'une et l'autre langue (French et Jacquet, 2004).

Sur le plan développemental, il semble qu'au fur et à mesure de l'apprentissage, le lexique mental de la L2 se transforme (O'Gorman, 1996; Meara, 1999; Singleton, 1997). Si chez le jeune enfant les associations de mots dans la L1 ont surtout une base formelle, évoluant graduellement vers des associations sémantiques, il semble que ce changement se passe aussi au cours de l'apprentissage de la L2 (Singleton, 1997). Selon Kroll et Stewart (1994), du moins pour les locuteurs ayant appris la L2 après la puberté, il y aurait une composante développementale dans le mode d'accès au lexique. Les résultats de tests en neurologie ont démontré que plus une L2 est bien maîtrisée, plus les activations cérébrales dues à un stimulus en L1 ou en L2 sont similaires, que ce soit en compréhension ou en production (Pallier et Argenti, 2003). Chez les bilingues qui ne le sont pas parfaitement, les activités cérébrales en L2 seraient plus variables, une plus grande activation traduirait le fait que de travailler dans une langue qui n'est pas bien maîtrisée demanderait plus d'efforts mentaux (Pallier et Argenti, 2003).

Tous les chercheurs ne conçoivent pas de la même façon la constitution et le fonctionnement de ce lexique mental L2 et le débat n'est pas encore terminé, semble-t-il, entre deux hypothèses principales. La première, qui porte différentes appellations (entrepasage partagé, interdépendant ou en commun,) affirme qu'il n'y a qu'un seul et même système lexical étendu (*extended*) qui comprend les éléments des deux langues. Il n'y a qu'un seul système sémantique, mais chaque mot est étiqueté selon la langue à laquelle il appartient (van Heuven, et al., 1998). L'autre

position, dite de *dual system*, propose que le bilingue possède des représentations séparées pour les mots et leurs traductions au niveau lexical, mais que ces mots partagent les mêmes représentations au niveau sémantique (Smith, 1991). À notre avis, plusieurs recherches ont apporté des arguments solides en faveur de l'hypothèse des deux lexiques et des tenants de cette position ont proposé des modèles de l'organisation du lexique mental du bilingue capables d'expliquer les erreurs et les hésitations du locuteur dans une tâche en L2.

2.1.3 L'hypothèse des deux lexiques

Sur le plan neurologique, l'hypothèse des deux lexiques est appuyée par l'observation chez des patients bilingues atteints d'aphasie, de cas de récupération sélective, successive ou antagoniste de l'une et l'autre langue (Grosjean, 1982). Chez certains sujets, il y avait perte d'une langue, mais l'autre demeurait relativement intacte. Il peut y avoir différents types d'aphasie dans l'une et l'autre langue chez un même patient et le pattern de récupération peut varier pour la L1 et la L2. Ces constatations portent donc à croire que les deux langues peuvent fonctionner indépendamment et que cette séparation se situerait au niveau lexical (Chapnik-Smith, 1997).

S'il y a deux lexiques séparés, comment se font les liens entre les deux et avec les concepts sous-jacents? Différentes hypothèses ont été avancées au fil des ans dont des modèles dits de médiation conceptuelle proposant que les connexions entre les items lexicaux de L1 et L2 se font au niveau des représentations conceptuelles (Potter et al., 1984; Kroll et Curley, 1988; Kroll et Stewart, 1994; De Groot, 1993; Kroll et Tokowicz, 2001). Parmi ces modèles, certains diffèrent des autres dans leur définition des liens qui existent entre les deux lexiques et entre le lexique de L2 et le système conceptuel. Kroll et Stewart (1991) et Kroll et Tokowicz (2001), proposent un principe d'asymétrie dans les liens entre les items lexicaux de L1 et L2 et avec le

système conceptuel. Dans leur modèle hiérarchique, tel qu'illustré ci-dessous, les mots de la L1 et de la L2 sont reliés par des liens lexicaux et conceptuels. Cependant, les liens lexicaux de la L2 vers la L1 sont plus forts que ceux de la L1 vers la L2. Les liens conceptuels avec la L1 sont aussi plus solides.

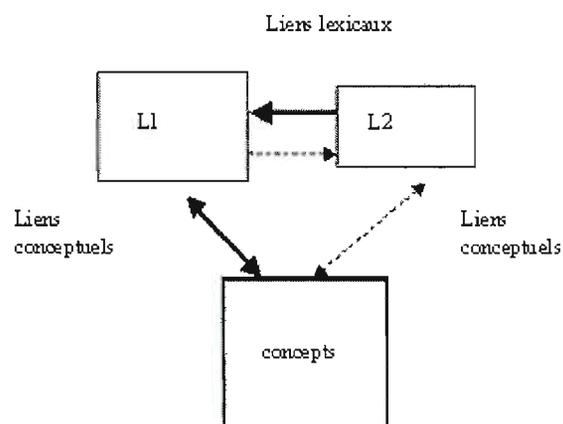


Figure 2.1 Le modèle de Kroll et Stewart. (Schéma adapté du modèle de Kroll et Stewart (1994) et Kroll et Tokowicz (2001)).

Leur modèle, qui s'applique davantage aux apprenants novices, explique qu'au début ces derniers n'arriveraient pas à traduire des concepts directement dans des mots de la L2: ils devraient procéder par médiation de la L1. Avec le temps, va se développer la capacité d'une médiation conceptuelle avec la L2 et l'information conceptuelle disponible pour la L1 deviendra la même pour la L2.

Les hypothèses de Paradis (2004) nous fournissent davantage d'explications sur les erreurs commises par les locuteurs bilingues avancés et sur les difficultés qu'ils rencontrent. Tout d'abord, l'auteur émet l'hypothèse des sous-systèmes linguistiques

chez le bilingue et propose une conception du lexique mental à trois unités de stockage: les concepts; les représentations sémantiques de L1, les représentations sémantiques de L2, lesquelles sont connectées au module conceptuel, tel qu'illustré ci dessous.

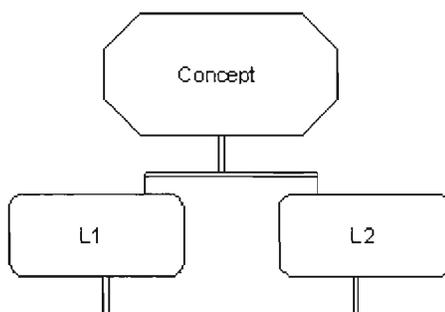


Figure 2.2 Le modèle de Paradis. (Schéma adapté du modèle de Paradis (2004)).

Sa conception se base sur la proposition d'une mémoire sémantique spécifique à la langue et distincte de la mémoire conceptuelle qui elle, est indépendante de la langue ou des langues. Pour Paradis, si le contenu du système conceptuel n'est pas linguistique, cependant l'organisation des unités verbalisables est spécifique à la langue, car il n'y pas toujours de correspondance exacte entre un mot de la L1 et sa traduction en L2. Les concepts sous-jacents à un mot et à sa traduction dans la L2 vont différer dans la mesure où l'organisation lexicale sémantique n'est pas la même dans la L1 et dans la L2. Les représentations conceptuelles qui correspondent aux représentations sémantiques d'un item lexical forment un sous-ensemble dans le système conceptuel. Les éléments compris dans ce sous-ensemble et leurs traductions dans la L2 ne se superposent que partiellement et le degré de superposition dépend de la proximité des deux langues tout comme il peut varier d'un mot à l'autre. Il se peut que certains concepts n'aient de réalisation linguistique

que dans une langue et si une langue ne possède pas de mot pour un concept, il sera difficile de le verbaliser. Dans ce cas, le locuteur bilingue aura tendance à utiliser le mot dans la langue qui possède cette correspondance avec un concept : c'est le phénomène de l'emprunt.

Selon l'hypothèse du niveau seuil de Paradis (2004), un item du lexique mental est activé lorsqu'il y a suffisamment d'impulsion neurale. Cette quantité d'impulsion nécessaire représente le niveau seuil d'activation. À chaque fois qu'un item lexical est activé, son niveau seuil est abaissé et moins d'impulsion est nécessaire pour sa réactivation. En contrepartie, un item qui n'est pas souvent activé sera plus difficile à stimuler. La sélection d'un item lexical particulier demande que son activation dépasse celle des autres candidats potentiels et que ces compétiteurs soient inhibés, c'est-à-dire, que leur niveau seuil d'activation soit haussé. Enfin, la disponibilité d'un item est fonction de la fréquence et de la récence de son activation, ce qui implique que des items peuvent être davantage disponibles que d'autres.

Par rapport à d'autres conceptions du lexique mental, celle de Paradis nous apparaît plus complète et mieux en mesure de décrire ce qui se passe chez des apprenants avancés de L2. Elle a l'avantage d'intégrer des facteurs importants de l'input (fréquence, récence) et de prendre en compte la distance L1-L2, autant d'éléments qui peuvent expliquer les difficultés et les erreurs de l'apprenant d'une L2.

Après avoir vu les opérations mentales impliquées dans le traitement des items lexicaux, nous nous intéresserons maintenant à ce que contient le lexique : les mots. Plus exactement, nous tenterons de définir ce que signifie connaître un mot, nous considérerons le savoir lexical selon ses deux dimensions, soit la profondeur et l'étendue et nous verrons quelle est leur importance.

2.2 La connaissance lexicale

Connaître un mot, *grosso modo*, c'est posséder l'information sur les aspects de la forme et du sens dont nous avons besoin pour l'identification et l'utilisation correcte de ce mot (Chialant et Caramazza, 1995). Bien que le contenu sémantique soit central dans le lexique (Laufer et Goldstein, 2004), connaître un mot c'est beaucoup plus que savoir sa définition (Beck et McKeown, 1991; Nagy et Scott, 2004) ou en connaître le sens (Anderson et Nagy, 1991). Nation (1990) identifie huit composantes dans la connaissance d'un mot : la forme orale, la forme écrite, le comportement grammatical, la collocation, la fréquence, le registre, le sens conceptuel, son association avec d'autres mots. Laufer (1990) regroupe les composantes de la connaissance lexicale comme suit : la forme (phonologique, graphique, morphologique); le comportement syntaxique; le sens (référentiel, associatif, pragmatique); les relations avec les autres mots (syntagmatique et paradigmatic). Qian (1999) parle plutôt de la connaissance des propriétés d'un mot : ses propriétés formelles, sémantiques et syntaxiques. Toutes ces descriptions font référence à l'une des notions centrale du savoir lexical, c'est-à-dire, la profondeur (*depth*). Connaître un mot en profondeur voudrait donc dire posséder toutes les informations sur ce mot, ce n'est pas toujours le cas chez les locuteurs. Un locuteur peut ne connaître que certains aspects d'un mot et cela va se traduire dans sa performance relative dans une tâche donnée. Différentes composantes de la connaissance lexicale sont impliquées dans différentes tâches et des connaissances partielles peuvent mener tout de même à une bonne performance (Qian, 2002). Dans une tâche simple de reconnaissance lexicale par exemple, le sujet doit être capable d'identifier la forme du mot, il n'a pas besoin d'en connaître le sens (Henriksen, 1999).

Comme nous l'avons vu dans la section consacrée au lexique mental, on ne peut pas concevoir les mots d'une langue comme des items isolés, mais plutôt comme des éléments faisant partie d'un système, d'une organisation. La connaissance

approfondie d'un mot suppose aussi la capacité du locuteur à le relier à d'autres mots ou à l'en dissocier, ce qui signifie connaître le réseau de ce mot (Read, 2004). Selon Read (2004), la profondeur des connaissances lexicales peut aussi être définie en termes de qualité des divers liens construits entre les mots dans le lexique mental. Savoir des mots et connaître le système lexical sont deux aspects différents (Nation, 2001).

« Words are not isolated units of language, but fit into many interlocking systems and levels. Because of this, there are many things to know about any particular word and there are many degree of knowing. » (Nation, 2001, p. 23).

La connaissance lexicale est complexe (Nagy et Scott, 2004) et on dit qu'elle est de nature progressive (*incremental*) (Schmitt 1998; Cobb, 2000; Nation, 2001; Qian 2002 ; Laufer et Goldstein, 2004), c'est-à-dire, que c'est un savoir qui se développe en plusieurs étapes, allant de l'absence totale de connaissance du mot jusqu'à la capacité à l'utiliser de façon productive. En effet, il faut distinguer entre reconnaître un mot et le comprendre (Chall, 1987), et comprendre un mot ne signifie pas nécessairement qu'on saura l'utiliser (Laufer et Goldstein 2004).

Il est très courant de distinguer vocabulaire actif ou productif et vocabulaire passif ou réceptif. Le vocabulaire réceptif d'un individu, c'est sa capacité à percevoir la forme d'un mot et à retrouver sa signification; le vocabulaire productif lui, implique de retrouver en mémoire puis de produire la forme appropriée pour exprimer une signification (Nation, 2001). Cependant, cette distinction classique entre vocabulaire actif/productif et vocabulaire passif/réceptif n'est pas purement dichotomique (Laufer et Goldstein, 2004; Read, 2004; Henriksen, 1999). En fait, Il faut voir le savoir lexical sur un continuum, ayant plusieurs niveaux et plusieurs dimensions (Henriksen, 1999, Beck et McKeown, 1991). La connaissance passive/réceptive précède généralement la connaissance active/productive et sur le continuum, des zones se superposent et des mots qui sont acquis sur le plan réceptif deviennent peu à peu disponibles pour

la production (Read, 2004). Il se peut aussi que certains mots connus sur le plan réceptif ne deviennent jamais du vocabulaire actif (Laufer et Goldstein, 2004) et certains aspects d'un mot peuvent faire partie du vocabulaire productif alors que d'autres demeureront au stade passif (Henriksen, 1999).

Laufer et Goldstein (2004) conçoivent que savoir le sens d'un mot doit être représenté sous une forme hiérarchique allant du plus actif au plus passif et soulignent la différence entre pouvoir reconnaître un mot parmi une liste (plus passif) et pouvoir s'en rappeler la forme et le sens (plus actif).

« We hypothesize that knowledge of meaning constitutes a hierarchy in which some degrees of knowledge are more advanced than others and presuppose the less advanced degrees of knowledge. » (Laufer et Goldstein, 2004 p. 407).

Cette hiérarchie dans la connaissance lexicale se traduira évidemment dans les comportements linguistiques : un locuteur peut être performant en lecture mais piètre en production. Alors qu'en lecture une compréhension approximative du sens d'un mot peut suffire, en situation de production, le locuteur devra choisir le terme exact s'il veut que son message soit bien compris. De plus, il est possible que le locuteur ne connaisse qu'un des aspects sémantiques d'un mot, qu'un usage dans un contexte précis et pas les autres.

Laufer et Paribakht (1998) différencient deux types de vocabulaire actif : le vocabulaire actif « contrôlé » (*controlled*) et le vocabulaire actif « libre » (*free*). Le vocabulaire « actif libre », c'est ce que le locuteur choisit lui-même de produire ou d'utiliser. En expression libre, le locuteur peut éviter d'utiliser des mots qu'il juge complexes ou qu'il ne maîtrise pas bien. Le vocabulaire contrôlé, c'est ce que le locuteur peut utiliser si on lui demande de le faire dans une tâche. Ainsi, une tâche

contrôlée dans un test, par exemple, peut traduire plus fidèlement les connaissances d'un sujet par rapport à des items lexicaux donnés que la production d'un texte.

Un autre aspect central lorsqu'il est question de connaissances lexicales, c'est la notion d'étendue (*breadth*) que Read (2004) définit comme étant le nombre de mots pour lesquels on connaît au moins certains traits sémantiques. Selon Corson (1985,1997), l'étendue du vocabulaire serait un puissant prédicteur de la réussite académique.

« To a very large extent, then educational failure or success depends on children having the words, wanting to use them and being able to use them. » (Corson, 1985, p. 3).

Des recherches en lecture en L1 ont démontré l'importance à la fois de la taille et de la profondeur des connaissances lexicales (Beck et al., 1982). En fait, ces deux aspects du savoir lexical seraient étroitement reliés. Et puis, l'étendue et la profondeur des connaissances lexicales ont des effets communs dans l'organisation du lexique en mémoire.

« This parallel development of vocabulary size and depth is particularly pertinent if we adopt a network building perspective on depth, in that vocabulary growth also entails the building of more extensive linkages between items in the mental lexicon. » (Read, 2004, p. 221).

La connaissance lexicale est complexe et multidimensionnelle. C'est un savoir qui se construit et cette nature progressive (*incremental*) explique les différences de performance entre des individus pour une tâche donnée. Il en va de même pour un apprenant d'une L2. Nous aborderons dans ce qui suit les caractéristiques du savoir lexical en L2 ainsi que les besoins des apprenants adultes.

2.2.1 La connaissance lexicale en L2

Poursuivre des études dans une langue seconde demande de posséder des connaissances lexicales solides, tant en étendue qu'en profondeur. En fait, davantage que la maîtrise des structures de la L2, la compétence lexicale est de première importance compte tenu, notamment, de la place qu'occupe la lecture dans l'acquisition de savoirs académiques.

« [...] vocabulary knowledge is the single most important area of second language (L2) competence when learning content through that language is the dependent variable; grammatical accuracy is of little importance to student's immediate academic needs [...] » (Saville-Troike, 1984, p. 199).

La conception hiérarchique du savoir lexical amène certains auteurs à distinguer connaissances et habiletés ou compétences : les connaissances, ce sont les informations apprises sur un mot, elles font partie du savoir déclaratif. Les compétences relèvent du savoir procédural qui permet un traitement automatique (Henriksen, 1999; Jiang 2004) et c'est le but à atteindre pour ce qui est du lexique en L2. Ce savoir-faire expliquerait des différences de performance et de comportement langagier entre les locuteurs natifs et les non natifs. Par exemple, les lecteurs en L2, même les apprenants avancés, seraient beaucoup moins rapides (une différence pouvant aller jusqu'à 30%) que les lecteurs en L1 (Segalowitz et al., 1991; Weber, 1991). Cette différence de performance serait causée par une moins grande vitesse dans le traitement automatique et une précision moindre dans la reconnaissance des mots (Segalowitz et al., 1991). Pour Nagy et Scott (2004), malgré sa dimension déclarative, la connaissance lexicale est surtout un savoir procédural.

« Word knowledge is applied knowledge. A person who knows a word can recognize it, and use it, in novel contexts, and uses knowledge of the word, in combination with other types of knowledge, to construct meaning for a text. » (Nagy et Scott, 2004, p. 579).

Selon Meara (2005), les locuteurs non natifs n'arrivent que rarement à atteindre un degré d'aisance avec les mots comparable à celui des locuteurs natifs. L'apprenant de L2 va éviter d'utiliser certains mots, notamment ceux pour lesquels il n'y a pas d'équivalence dans la L1, va préférer les mots plus fréquents et ceux qui sont morphologiquement plus simples, va parfois surgénéraliser l'usage de certains mots (Meara, 1984). Même lorsqu'ils sont avancés dans la L2, les locuteurs non natifs auraient tendance à utiliser le terme général plutôt que le mot spécifique, à bien maîtriser le sens prototypique, mais à manquer de connaissances sur les frontières sémantiques et ils auraient du mal à tenir compte des contraintes de registre et de connotation (Ijaz, 1986). Comparant des locuteurs natifs et non natifs dans une tâche d'association de mots, Meara (1984) puis Read (1993) ont remarqué que les premiers établissaient des liens surtout sur une base sémantique alors que les seconds produisaient des associations beaucoup moins prévisibles, ce qui traduirait que leurs connaissances lexicales et leur organisation dans le lexique mental, surtout en début d'apprentissage, sont très peu stables. Bien entendu, les différences de comportement entre locuteurs natifs et non natifs sont plus marquées lorsqu'il s'agit d'apprenants débutants; elles vont s'estomper peu à peu, au fur et à mesure que se développe l'habileté dans la L2. Nous verrons plus en détails quelles sont les conséquences d'un bagage lexical incomplet ou insuffisant pour la compétence linguistique en L2 et pour la réussite académique.

Tout d'abord, il faut avoir atteint un certain niveau, en termes de taille du vocabulaire, pour que les habiletés acquises en L1 puissent être transférées en L2 (Laufer, 1997; Cobb, 1999; Tschirner, 2004). L'étendue du vocabulaire est un préalable absolu pour le développement de l'habileté dans l'usage de la langue cible (Nation, 1993).

« [...] if learners do not develop a sufficiently large vocabulary then skill in language use and the knowledge that comes from that skill will not develop. »
(Nation, 1993, p. 132).

En lecture L2, par exemple, c'est davantage le niveau de compétence lexicale que les habiletés (dont la lecture en L1) qui détermine le succès (Laufer, 1992) et les difficultés du lecteur seraient liées aux limitations au niveau lexical plutôt que syntaxique (Nagy et al., 1993). Même de bonnes connaissances grammaticales de la langue cible ne suffisent pas à assurer une aisance en lecture (Weber, 1991).

Quelle est donc la taille du vocabulaire nécessaire pour assurer une bonne compréhension ? Selon plusieurs auteurs, un vocabulaire d'environ 5,000 des mots les plus fréquents ferait en sorte que le lecteur puisse comprendre environ 95% d'un texte (Laufer, 1992, Nation et Newton, 1997). Cette estimation est valable pour plusieurs langues, notamment l'anglais, le français, le russe et le hollandais (Hazenberg et Hulstijn, 1996). Ce serait là la base, le niveau seuil, pour la compréhension et pour le transfert des stratégies de lecture acquises en L1 (Coady 1997, Huckin et Coady, 1999). Liu et Nation (1985) ont comparé les résultats en lecture chez des sujets qui avaient des connaissances lexicales leur permettant de comprendre 90% et 95% des mots d'un texte et ont découvert que ceux qui avaient 95% réussissaient beaucoup mieux à déduire le sens des mots inconnus par le contexte. Qian (1999) explique que le fait d'avoir un vocabulaire étendu fournit à l'apprenant un ensemble de données nécessaires pour déduire le sens des mots.

Beaucoup d'apprenants auraient tendance à surévaluer leurs propres connaissances lexicales en L2. Selon Laufer et Yano (2001), identifier ce qui est inconnu peut être assez simple pour un débutant ou pour un apprenant très avancé. Cependant, les apprenants qui se situent entre ces deux pôles peuvent être moins certains de ce qu'ils ne savent pas bien. Dans leur étude réalisée auprès d'étudiants universitaires de niveau avancé en L2, les sujets ayant obtenu les meilleurs scores à un test de vocabulaire étaient aussi ceux qui avaient estimé le plus justement leurs connaissances lexicales.

« A high vocabulary coverage of the text, i.e. a large percentage of known vocabulary of the text, is claimed to be essential for successful reading comprehension. One of the reasons for this may lie in the positive effect a high coverage has on learners' ability to understand what they don't understand. » (Laufer et Yano, 2001, p. 560).

Des chercheurs ont estimé qu'il faudrait connaître davantage que 95% des mots pour comprendre un texte : ce serait plutôt de l'ordre de 97% à 98% (Hu et Nation 2000; Nation, 2001). En fait, selon un simple calcul mathématique, si un lecteur connaît 90% des mots d'un texte, alors il y aura dix mots inconnus pour chaque 100 mots connus, et si chaque ligne compte dix mots, il rencontrera un mot inconnu à chaque ligne de texte (Hirsch et Nation 1992); si le locuteur connaît 95% des mots du texte, le nombre de mots inconnus passera à un pour 20. Un gain de pourcentage au delà du 95% serait d'importance, surtout lorsqu'il s'agit de textes non simplifiés (Hirsch et Nation, 1992). Il faudrait donc connaître quelque 5,000 familles de mots pour atteindre les 97 à 98% de mots connus dans un texte (Hirsh et Nation 1992).

Dans une étude ayant pour sujets des étudiants universitaires de 1^{ère} année, Laufer (1992) a démontré que si l'apprenant possède un vocabulaire de moins de 3,000 familles de mots, ses habiletés académiques générales ne pourront compenser ses connaissances lexicales dans la compréhension d'un texte. L'auteur considère la compétence lexicale comme étant à la base même des habiletés académiques.

« If a good knowledge of foreign vocabulary can compensate for lower general academic ability, then even learners with a mediocre ability can improve considerably in their L2 reading once they have raised their lexical level. » (Laufer, 1992, p. 101).

Le nombre de mots requis pour comprendre des journaux ou des textes scientifiques serait plus élevé que 5,000 familles de mots parce que les termes académiques sont plus nombreux dans des textes informatifs (Waring et Nation, 2004). Et puis, la

compréhension des informations dans ce type de texte tolère mal l'ignorance d'un certain nombre de mots. Nation et al. (1990) ont évalué à 17,000 familles de mots (environ 40,000 mots) la taille du vocabulaire d'un locuteur natif de l'anglais poursuivant des études universitaires. Huckin et Coady (1999) ont établi que la compréhension de textes de niveau universitaire nécessitait de connaître 10,000 familles de mots en L2. Dans une recherche, Hazenberg et Hulstjin (1996) ont trouvé que les étudiants non natifs qui avaient échoué au test d'admission à l'université possédaient un vocabulaire d'environ 9,000 familles de mots. Pour les auteurs, 10,000 familles de mots représentent donc la taille minimale pour aspirer entreprendre et réussir des études universitaires en L2.

Il va de soi que la lecture de textes académiques et la réussite d'études supérieures en L2 nécessitent la maîtrise du vocabulaire académique (Cobb, 1999; Nation, 1993). Des recherches ont permis d'établir une relation entre l'étendue du vocabulaire et les performances académiques. Morris et Cobb (2004) ont mesuré le niveau de vocabulaire de leurs sujets et l'ont comparé avec les résultats à des cours de grammaire de leur programme de baccalauréat en enseignement de l'anglais L2. Les auteurs ont trouvé une corrélation significative entre le niveau de richesse lexicale et les notes obtenues aux cours de grammaire, la corrélation la plus forte étant pour les mots de niveau académique. Laufer et Goldstein (2004) ont soumis quelques centaines d'étudiants de différents niveaux académiques (du secondaire jusqu'à l'université) à des tests mesurant l'étendue du vocabulaire. Les scores obtenus ont ensuite été comparés aux performances des sujets en classe de L2. L'application d'un test de régression a permis d'établir que plus de 40% de la variance totale s'expliquait par la connaissance lexicale des sujets.

S'il a été démontré que le nombre de mots connus en L2 est important, il ne faut pas oublier le rôle de la profondeur du savoir dans la compréhension. Ainsi, le fait d'avoir une connaissance approfondie du lexique va donner de meilleurs résultats quand il

faut déduire le sens d'un mot inconnu (Qian 2002). De fait, tout comme pour la L1, la taille et de l'étendue du vocabulaire en L2 sont interreliées et agissent de façon interactive et interdépendante (Qian, 2002). Qian (1999) a trouvé une forte corrélation entre la taille du vocabulaire, la profondeur des connaissances lexicales et la compréhension en lecture. Les résultats au test administré à des étudiants universitaires avancés en L2 par Greidanus et al. (2004) ont mené à la même observation: étendue et profondeur lexicales se superposent. Il en va de même pour la recherche de Nurweni et Read (1999) dans laquelle les sujets ayant des connaissances lexicales limitées sur le plan du nombre de mots obtenaient aussi un score faible aux mesures de profondeur. La profondeur des connaissances lexicales évoluerait cependant moins rapidement que l'étendue du vocabulaire (Greidanus et al., 2004).

Le locuteur qui poursuit des études dans une langue qui n'est pas sa langue maternelle doit disposer d'un vaste vocabulaire et de connaissances lexicales relativement sophistiquées. Ceci nous amène à considérer les facteurs qui interviennent dans l'acquisition du vocabulaire en L2. Nous verrons tout d'abord les processus de base dans l'acquisition du vocabulaire dans une langue seconde et quels sont les défis particuliers de l'apprenant adulte qui possède déjà un système conceptuel dans sa langue maternelle. Il sera ensuite question des conditions favorables à l'atteinte d'une compétence lexicale suffisante en L2 et des résultats de recherches l'acquisition du vocabulaire.

2.3 L'acquisition du vocabulaire en L2

L'appariement (*mapping*) forme-sens est un des processus de base dans l'acquisition du vocabulaire en L2 (Jiang, 2004; Laufer et Goldstein, 2004; Hulstijn et al., 1996; Nikolova, 2004). L'enfant qui apprend des mots de sa langue maternelle acquiert en même temps des concepts; conséquemment, forme et sens sont

étroitement associés, sinon indissociables. Nous savons que l'adulte apprenant une L2 possède déjà un système conceptuel bien établi. Un des aspects spécifiques dans l'acquisition du vocabulaire L2 chez l'adulte c'est la séparation entre développement lexical et sémantique. À la différence de l'acquisition du vocabulaire en L1, le développement du lexique en L2 n'est pas nécessairement accompagné d'un développement sémantique. La plupart des mots de la L2 que l'apprenant adulte rencontre ont une correspondance conceptuelle et une traduction en L1: l'apprenant adulte peut donc s'appuyer sur son système sémantique pré-établi (Jiang, 2004). Cependant, vue la différence possible entre L1 et L2 dans la façon d'apparier forme et sens, l'apprenant devra procéder à une restructuration sémantique afin d'utiliser correctement la L2 (Jiang, 2004 b).

Selon la conception de Jiang (2004, 2004b), l'apprenant de la L2 va, dans un premier temps, mémoriser la forme d'un mot en L2 en l'associant à un concept ou à une signification en L1: le sens est compris à travers le système sémantique de la langue maternelle. Cette juxtaposition sur ce qui est existant en L1, ou transfert sémantique, peut se produire dans la mesure où il y a effectivement un mot ou un concept semblable au sens du mot en L2. Ce procédé peut permettre l'utilisation correcte d'un mot dans la L2, cependant, l'utilisation exacte, idiomatique, exige que le mot de L2 ait sa propre identité sémantique. Lorsque l'apprenant prend conscience qu'il n'y a pas de correspondance exacte et qu'il sait quelle est la différence, il doit opérer une restructuration sémantique.

« Restructuring will successfully take place only when the conflict between the transferred semantic understanding and the understanding obliged by the context is powerful enough to override the transferred meaning, sometimes accompanied by a communication break down. However, such contexts may be available for some words, but not others. » (Jiang, 2004b, p. 105).

Cette prise de conscience amènera l'apprenant à développer la profondeur de sa connaissance et à intégrer des aspects plus périphériques et raffinés, tels que le

sens figuré, la connotation et pour ce faire, il faudra donc que le mot de L2 acquière son autonomie. Mais le développement sémantique est un processus lent et parfois incomplet chez l'apprenant adulte et même les apprenants avancés continueraient de s'appuyer sur leur système sémantique de la L1 dans l'usage de la L2 (Ijaz, 1986; Jiang, 2004; Jiang, 2004b). Pour tout apprenant de L2, l'acquisition du sens spécifique des items lexicaux représente un défi majeur et même des apprenants avancés de L2 vont se distinguer des locuteurs natifs à cet égard. Le défaut d'opérer rapidement une restructuration sémantique peut mener à une connaissance incomplète des items lexicaux. Jiang (2004b) est d'avis que plusieurs mots n'atteignent jamais le stade d'autonomie sémantique. Il explique qu'il arrive fréquemment que des apprenants de L2 utilisent un mot de la L2 avec le sens de la L1, ou encore, que lorsque deux mots de L2 partagent la même traduction en L1, l'apprenant utilise ces deux mots indistinctement. Selon lui, il y a risque de fossilisation si la restructuration sémantique ne s'opère pas assez rapidement. Schmitt (1998) estime que si un apprenant ne connaît pas les sens particuliers d'un mot dès le début, dans 72% des cas, cette connaissance sémantique ne va pas changer avec le temps. Pour Jiang (2004), l'autonomie lexicale est une partie très importante de la compétence en L2.

« The ultimate goal of L2 learning is to develop a linguistic and semantic means that can function as an autonomous system. Autonomy is important because idiomatic, accurate, and fluent expression of one's ideas in another language cannot be achieved until one has developed a semantic system that is uniquely coded in and by that language, a system that is shared by native speakers of that language. » (Jiang, 2004, p. 425).

Nous ne savons pas si les sujets de notre recherche ont appris la L2 dans un contexte formel ou informel, ni pendant combien de temps. Ils ont cependant été exposés à la L2 d'une façon ou d'une autre et suffisamment pour avoir développé une certaine aisance à l'oral. Quel que soit le mode d'apprentissage du lexique, la nature de l'input et le style personnel de l'apprenant, tous ces facteurs ont une influence sur le nombre d'items lexicaux acquis et sur la qualité de ce qui est acquis.

Le facteur qui joue un rôle majeur dans l'acquisition du lexique, c'est la fréquence d'exposition. Nous traiterons dans la section suivante de tous ces éléments qui constituent les conditions d'acquisition du lexique en L2 et nous ferons état de quelques recherches sur le sujet.

2.3.1 Les conditions d'acquisition

On distingue généralement deux grandes modes d'apprentissage du vocabulaire : de façon incidente, c'est-à-dire, qui n'est pas explicitement axée sur le vocabulaire; de façon intentionnelle, alors que but est de mémoriser des informations lexicales (Laufer, 2003). Une autre distinction est faite entre apprentissage implicite et explicite. Dans le premier cas, il n'y a pas d'opération consciente, alors que dans le second, on porte attention, de façon sélective à l'input (Laufer, 2003). La forme la plus pure d'apprentissage incident est l'acquisition par la lecture et pour certains, comme Krashen (1989), par exemple, c'est la meilleure méthode. L'hypothèse basée sur l'input fait valoir qu'il serait impossible d'apprendre de façon explicite tous les mots nécessaires en L1 et que la seule façon plausible, serait par l'exposition, l'input. Il est vrai que la rencontre des mots de vocabulaire en contexte peut fournir l'environnement riche nécessaire pour développer des connaissances approfondies.

« Well-elaborated semantic knowledge, which includes developing knowledge of usage, collocations and other lexico-grammatical characteristics, is primarily gained through meeting words in context rather than through explicit instruction. » (Hunt et Beglar, 2005, p. 28).

Certains estiment cependant que le taux d'apprentissage du lexique par la lecture seule peut être aussi faible que 4% (Waring et Nation 2004). Comme le fait observer Qian (2002), il faut bien distinguer le type de traitement textuel selon l'objectif poursuivi. Ainsi, lire pour une compréhension générale demande de comprendre des idées et des sous-ensembles reliés au thème principal, alors que lire pour apprendre exige un traitement plus complexe. Une des stratégies de compréhension largement

utilisées est l'inférence, définie comme étant le traitement qu'il faut appliquer lorsqu'on ne connaît pas la signification d'un mot (Parel, 2004). Selon plusieurs auteurs, cette stratégie aurait de sérieuses limites. L'inférence est un processus lent qui ne mène pas nécessairement à la rétention parce que ce qu'il faut pour deviner le sens d'un mot inconnu n'est pas nécessairement ce qu'il faut pour le retenir en mémoire, probablement parce que le besoin, comprendre, a été répondu (Sökmen, 1997). Selon Gu (2003), les recherches récentes ont démontré que l'apprentissage incident donnerait de meilleurs résultats avec les locuteurs natifs et les apprenants intermédiaires et avancés en L2. Il demeure cependant que l'acquisition du lexique par la lecture est un processus par petits pas (Hulstijn et al., 1996). La méthode est longue et ardue, il faudrait lire énormément en L2, ce que la plupart des apprenants ne font pas (Cobb et Horst, 2004).

Waring et Nation (2004) expliquent que la grande variance dans les résultats des recherches ayant porté sur l'acquisition de vocabulaire par la lecture tiendrait au nombre d'occurrences des items ciblés et de leur fréquence. Nagy, Herman et Anderson (1985) ont évalué que les chances qu'un mot soit appris après une seule exposition sont d'environ 15%. En effet, l'apprentissage du vocabulaire demande plusieurs expositions (Hulstijn et al., 1996; Nagy, 1997; Hulstijn 2002; Rott et William, 2003; Waring et Tagaki, 2003, Horst et Meara, 1999) : c'est ce que Laufer (2003) appelle le « gain cumulatif » (notre traduction). De rencontrer à plusieurs reprises les mêmes mots a pour effet qu'à chaque fois le lecteur peut raffiner ou modifier le sens attribué (Rott et Williams, 2003). Le nombre nécessaire d'expositions varie selon les auteurs et il va de cinq à 16 (Nation, 1990), 100 (Meara, 1997), huit (Horst et al., 1998). Selon Waring et Nation (2004), c'est probablement une vingtaine d'expositions à un mot qu'il faudrait pour qu'il y ait rétention si l'on ne s'en tient qu'à l'apprentissage incident par la lecture. Pour Laufer (2003), pour qu'un mot soit retenu, il faut non seulement que le locuteur y ait été exposé à plusieurs reprises, mais aussi que le mot soit revu avant qu'il ne soit oublié. Les gains

potentiels sur le plan lexical peuvent donc varier énormément selon l'input auquel est exposé l'apprenant.

La reconnaissance et la production de mots de vocabulaire découlent de leur fréquence d'occurrence dans la langue et ceci est vrai à l'écrit et à l'oral, en compréhension comme en production « [...] language processing is intimately tuned to input frequency. » (Ellis, 2002, p. 143). Cependant, les mots ne se retrouvent pas avec la même fréquence dans les textes oraux et écrits: l'input oral contient normalement moins de mots difficiles ou de basse fréquence que l'input écrit (Nagy et al., 1985). En anglais, par exemple, les 1,000 mots les plus fréquents sont très répandus dans la langue orale et comptent pour 75% des mots d'usage courant dans des textes écrits, alors que les 2,000 mots les plus fréquents sont moins fréquents à l'oral, mais sont régulièrement présents à l'écrit. (Cobb et Horst, 2004). Selon Laufer et Paribakht (1998), les mots plus fréquents passeraient plus facilement de l'état passif à actif que les mots moins fréquents, parce qu'ils sont essentiels pour communiquer et qu'ils doivent donc être pratiqués, ce qui, probablement, aide à la rétention à long terme. L'acquisition du vocabulaire se fait, grosso modo, par l'ordre de fréquence, que ce soit en milieu naturel ou scolaire (Cobb et Horst, 2004), avec parfois comme conséquence, le risque de plafonnement.

« The natural frequency distribution of words will incline many L2 learners to plateau when they have acquired the resources needed for basic spoken interaction [...] » (Cobb et Horst, 2004, p. 4).

La situation d'exposition à l'input en L2 peut être très différente de ce qui se passe en L1 alors que les locuteurs ont constamment la possibilité d'être exposés à du vocabulaire varié dans des contextes différents. L'apprenant de L2 ne bénéficie pas nécessairement de cet environnement (Ghadirian, 2002). La plupart des apprenants ne seront pas en contact avec un vocabulaire très étendu et la majorité du lexique se

retrouve dans des textes écrits. Ainsi, par exemple, un apprentissage de la L2 axé sur l'aisance à l'oral, en milieu naturel, ne fournit pas toujours l'input lexical suffisant pour acquérir un grand nombre de mots du lexique. Le but est différent, l'input aussi (Cobb et Horst, 2004). Même en présence d'un input contextualisé et riche, la compétence en L2 peut cesser de se développer et mener à une fossilisation lexicale si l'apprenant n'a pas les habiletés suffisantes pour utiliser les informations disponibles dans l'environnement (Jiang, 2000).

Laufer (1991) souligne l'importance de la motivation dans l'acquisition du vocabulaire en L2. Dans sa recherche réalisée auprès d'étudiants avancés en L2, elle a trouvé que même dans un environnement les exposant à un vocabulaire riche, les apprenants habiles en L2 qui arrivent à bien communiquer à l'oral ne ressentent pas nécessairement le besoin d'améliorer leurs connaissances lexicales. Il est donc tout à fait possible que des locuteurs semblent très à l'aise pour s'exprimer oralement dans la langue seconde, mais qu'ils ne maîtrisent bien qu'un nombre relativement limité d'items lexicaux, ou encore qu'ils n'aient que des connaissances incomplètes pour des items lexicaux.

Acquérir du vocabulaire, cela signifie encoder puis enregistrer des mots en mémoire, les retenir (Jiang, 2004). Selon plusieurs auteurs, pour qu'il y ait acquisition, il faut un investissement attentionnel de la part de l'apprenant (Ellis, 1994; Huckin et Coady, 1999; Rott et Williams, 2003). L'apprenant doit porter son attention non seulement sur le texte (oral ou écrit) de façon générale, mais spécifiquement sur le vocabulaire (Ellis, 2002). Une attention centrée sur les liens entre la forme et le sens mènerait à une meilleure rétention (Vidal, 2003) et plus il y a traitement cognitif en profondeur, plus les chances que le mot sera retenu sont grandes (Laufer et Hulstijn, 2001).

Enfin, un autre facteur qui relève des caractéristiques individuelles de l'apprenant viendrait expliquer des différences dans l'acquisition du vocabulaire en L2. Les stratégies pour l'apprentissage et la rétention du vocabulaire peuvent différer selon la L1 des apprenants. (Meara, 1984) a constaté, par exemple, que les apprenants chinois de L2 ont énormément de difficulté avec les mots longs et qu'ils portent beaucoup plus d'attention à la fin des mots que les locuteurs natifs de l'anglais; de leur côté, les hispanophones attachent beaucoup d'importance à la syllabe dans la représentation des mots en mémoire en langue maternelle. Dans une recherche comparant des apprenants de l'anglais L2, Wang et al. (2003) ont trouvé aussi des différences dans le comportement de leurs sujets par rapport aux aspects linguistiques selon leur langue maternelle (LM). Ainsi, pour une même tâche donnée, les difficultés peuvent différer selon la L1 des sujets.

De nombreux facteurs peuvent influencer l'acquisition et la rétention du vocabulaire, que ce soit en termes d'étendue ou de profondeur. Plusieurs de ces facteurs relèvent de l'input, alors que d'autres sont liés aux caractéristiques personnelles et aux stratégies de l'apprenant. Lorsque la langue cible est morphologiquement riche comme c'est le cas du français, savoir comment sont formés les mots est assurément une clé pour l'acquisition du vocabulaire. Dans la section suivante, nous décrirons les principes généraux dans la construction des mots en français et nous expliquerons en quoi des connaissances en morphologie dérivationnelle contribuent à la compétence lexicale.

2.4 La morphologie dérivationnelle: description

En français, on ne peut traiter de la question du lexique sans parler de la morphologie. Le lexique est constitué de mots simples, comme *main* et de mots complexes ou dérivés, comme *chansonnier*. Rey-Debove (1984) a établi que 80% des mots du français sont des dérivés. C'est dire l'importance de la dérivation dans

le lexique français et l'attention que nous devons lui accorder. Au plan formel, un autre facteur vaut la peine qu'on s'intéresse aux mots dérivés: contrairement aux mots simples, comme *main*, qui ne s'analysent pas et sont complètement arbitraires, les mots construits sont décomposables et peuvent être analysés (Lehman et Martin-Berthet, 1998).

La morphologie lexicale a pour objet la formation des mots par composition (par exemple, *tire-bouchon*) et par dérivation (par exemple, *chansonnier*), plus particulièrement, elle traite « [...] de la forme des mots, dans leurs différents emplois et constructions, et de la part d'interprétation liée à cette forme même. » (Huot, 2001, p. 9). Dans notre recherche, nous nous intéresserons uniquement à la construction par dérivation.

L'unité de base du mot, c'est le morphème, défini comme étant « [...] un segment préconstruit associant forme et sens, mais qui ne peut plus se décomposer en segments de ce même type. » (Riegel et al., 1994, p. 533). Dans le lexique, on trouve deux types d'unités : les morphèmes libres et les morphèmes liés. Ainsi, *stupid* est un morphème libre, qui peut fonctionner seul, alors que dans *stupidement*, *-ment* est un morphème lié qui ne peut exister par lui-même, qu'on ne peut trouver qu'à l'intérieur d'un mot. Dans le mot *stupidement*, *-ment*, morphème lié et affixe, est accolé à un morphème libre, *stupid-*, qui en est la base ou la racine. La dérivation consiste à créer des mots par combinaison de morphèmes libres et liés, par ajout d'un ou de plusieurs affixes à une base. L'ajout d'un affixe a un impact sémantique et peut aussi avoir des conséquences sur la catégorie grammaticale du nouveau mot ainsi formé par rapport à la base de départ. On parle de familles de mots, de familles morphosémantiques ou de paradigmes pour désigner les groupes de mots dont la base est identique et les morphèmes liés variables (*stupidité*, *stupidement*). Les items d'une famille partagent le sens de l'élément en commun et il y a cohésion sémantique dans un même paradigme (Lehman et Martin-Berthet,

1998), même si cela est parfois subtil à détecter. Être conscient de cette cohésion est d'importance puisque lorsqu'on connaît le sens de la base, il est possible de déduire au moins le sens général d'un dérivé complexe.

En français, on compte deux types d'affixes : les préfixes, placés en tête de mot (*dépister*), les suffixes, en fin de mot (*pisteur*). Les préfixes ont un fonctionnement assez souple, le même préfixe pouvant être accolé à une base verbale, nominale ou adjectivale, comme par exemple pour le préfixe *super-* : *superposer*, *superstructure*, *supersonique*. Les suffixes ont un fonctionnement plus limité en français et ne peuvent être attachés qu'à une ou deux classes grammaticales, leur contenu sémantique sélectionne les bases auxquelles ils peuvent se joindre (Reigel et al., 1994). Dans le système dérivationnel, les suffixes seraient beaucoup plus faciles à maîtriser que les préfixes selon Slobin (1973). Sur le plan du décodage linéaire, le début de la forme dérivée correspond exactement au début du morphème de base, alors que pour les mots préfixés, on accède d'abord au préfixe, puis à la forme de base. La forme dérivée aurait donc un avantage sur le plan de la perception. Mentionnons que dans cette recherche, nous ne traiterons que des suffixes dérivationnels.

Le suffixe peut jouer deux fonctions dans la dérivation : une fonction translative, il peut changer la catégorie grammaticale (ex : *jouer/joueur*); une fonction non translative : il ne change pas la catégorie grammaticale, mais ajoute une nuance sémantique ou stylistique, (ex. : *rêver/rêvasser*). Les suffixes qui ne modifient pas la classe grammaticale sont beaucoup moins fréquents en français (Leclerc, 1989). La morphologie dérivationnelle obéit à certaines règles et la compétence en cette matière implique la connaissance de trois grands principes : le principe relationnel : les mots ont une structure interne et ils peuvent avoir des morphèmes en commun; le principe syntaxique : la dérivation attribue la catégorie grammaticale; le principe

distributionnel, lequel définit les contraintes dans l'affixation, certains affixes ne s'attachant qu'à telle catégorie grammaticale (Tyler et Nagy, 1989).

La compétence dérivationnelle est donc le domaine des règles qui définissent ce qui est possible et impossible linguistiquement, le critère d'exclusion étant la violation d'une ou plusieurs propriétés de la règle appliquée. (Corbin, 1987, p. 6).

La connaissance de la morphologie dérivationnelle implique la capacité à déduire les systématismes de forme liées au sens dans le lexique (Rey-Debove, 1984). Elle implique aussi de savoir le rôle des suffixes dans la formation des mots de diverses catégories et à les reconnaître comme étant des unités qui distinguent les noms, les adjectifs, les verbes et les adverbes (Singson et al., 2000). Par exemple, le suffixe *-ance*, accolé à un verbe, forme un nom d'action: *vengeance*. Cette connaissance permet au locuteur qui rencontre pour la première fois le mot isolé *modélisation*, d'identifier qu'il s'agit d'un nom et non pas d'un adjectif. La connaissance des règles de dérivation explique aussi que le locuteur natif à qui l'on demanderait le nom de l'arbre qui porte des pamplemousses répondrait très probablement *pamplemoussier*, même si ce mot lui était préalablement inconnu, et non pas **pamplemousseur*, par exemple. Selon ce qu'a trouvé Carlisle (2000), c'est la capacité de dérivation qui contribue le plus dans la définition des mots complexes puisqu'il faut savoir le sens et la fonction des affixes et non seulement la signification de la base.

Les règles de dérivation comprennent certaines subtilités, voire des aspects complexes, qui sont pourtant généralement maîtrisés par les locuteurs natifs. Ainsi, fréquemment, des affixes synonymes sont en concurrence dans un mode de formation donnée et ils ne peuvent être substitués. Si nous avons, par exemple, *gavage* et *séparation*, des dérivés formés sur une base verbale, les formes **gavation* et **séparage* sont impossibles. Et puis, certains affixes ont des variantes ou allomorphes et la forme change selon la base à laquelle ils sont accolés : parallèlement, nous avons *-aison* (*crevaision*,) et *-ation* (*tractation*), *-eur* (*batteur*) et *-ateur* (*adaptateur*) (Riegel, et al., 1994; Grévisse et Goosse, 1989).

Parce qu'il y a récurrence de formes et de sens dans les mots dérivés, le locuteur peut voir que le mot *ami* est présent dans *amical*, *amitié*, tout comme il devrait être en mesure de décoder le sens – au moins approximatif - d'un mot inconnu en le décomposant en ses parties signifiantes. Analyser la structure interne d'un mot suppose l'identification des éléments dont il est formé (Lehman et Martin-Berthet, 1998). La morphologie permet donc de trouver un indice à l'intérieur du mot et elle peut aider à se souvenir d'un mot nouveau (Pressley et al., 1987). La compétence morphologique implique donc non seulement la prise de conscience de la construction des mots, mais aussi la capacité de manipuler ces structures (Carlisle, 1995). C'est dire que les locuteurs qui tirent profit de l'information morphologique détiennent un outil puissant pour accroître leur vocabulaire (White et al., 1989).

On sait que les aspects de fréquence et de productivité sont importants pour la compréhension et l'acquisition des formes complexes chez les enfants en langue maternelle (Carlisle, 2000). White et al. (1989) ont formulé une hypothèse de généralisation morphologique selon laquelle les mots affixés très fréquents, qui peuvent avoir été acquis de façon incidente, jouent un rôle dans le développement du vocabulaire et de la compétence morphologique. Tous les affixes ne sont pas également productifs en français: certains l'ont été autrefois, mais ne sont plus utilisés aujourd'hui pour la création de nouveaux mots (Grevisse et Goosse, 1989; Riegel et al., 1994). En français, des affixes, comme *-ion*, sont présents dans plus de 2,000 mots, alors que d'autres, comme *-esse*, ont servi à former à peine un centaine de dérivés (Huot, 2001). Certains suffixes sont particulièrement productifs dans un champ lexical spécifique et plus restreint, comme les sciences ou la technologie, par exemple. La connaissance des suffixes les plus productifs est donc une clé qui ouvre la porte à la compréhension d'un grand nombre de mots.

Apprendre la morphologie, c'est trouver des similitudes sur le plan phonologique et aussi sur le plan sémantique entre des mots comme dans la paire *ami/amitié* (Freyd

et Baron, 1982). En français, il arrive que la base soit différente dans le mot simple et dans le mot dérivé, soit pour des raisons historiques (modifications phonologiques, comme pour *sel-saler*), soit parce que le dérivé est construit sur une base latine, *bœuf/bovin*, par exemple (Grévisse et Goosse, 1989). Il est question ici du degré de transparence dans la dérivation.

La vitalité des suffixes est fondée sur deux caractères : la *productivité*, c'est-à-dire, l'aptitude à produire des dérivés, et la *motivation*, c'est-à-dire, le fait que les usagers perçoivent les dérivés comme contenant une base et un suffixe. (Grévisse et Goosse, 1993, p. 200).

Dans la forme régulière, le morphème de base conserve sa forme orthographique et le dérivé est donc transparent : *peur/peureux*, par exemple. Alors que dans les formations irrégulières, les changements orthographiques ou phonétiques rendent la forme dérivée plus opaque (Koda, 2000). Le degré de transparence peut aussi se situer au plan sémantique (Derwing, 1976; Colé et al., 2003). Longtin et al., (2003) donnent l'exemple des couples *fille/fillette* et *fauve/fauvette*, alors que pour le deuxième couple, le lien sémantique avec la base n'est pas clairement établi. La transparence a de l'importance dans l'acquisition, car elle suscite chez le locuteur la construction de paradigmes qui rendent explicites les régularités présentes dans le lexique (Clark, 1993). L'enfant s'appuierait beaucoup sur la transparence phonologique de la base, alors que l'adulte serait plus sensible à la transparence sémantique et serait moins gêné par des changements formels (Derwing, 1976). Selon Hancin-Bhatt et Nagy (1994), l'habileté à déceler des relations morphologiques entre des mots malgré des changements sur le plan orthographique ou phonologique dépendrait de la sensibilité métalinguistique. Cette sensibilité serait déjà présente, quoique en développement, chez des enfants du primaire en langue maternelle.

La connaissance de la morphologie et des processus de dérivation aide à la compréhension des items lexicaux et elle apporte aussi une contribution significative

dans les tâches de lecture. Ce sont principalement des recherches faites auprès d'enfants du primaire qui ont démontré l'apport significatif de l'analyse morphologique dans la compréhension de textes, notamment celles de Carlisle (2000), Singson et al. (2000) et Casalis et Louis-Alexandre (2000). L'analyse morphologique contribue à la compréhension en lecture dans la mesure où le lecteur est conscient des composantes morphologiques des mots, de leur sens et de leur rôle grammatical (Carlisle, 2000). Lecoq et al. (1986) font remarquer qu'il y a davantage de mots dérivés dans des textes écrits que dans la langue orale et que le caractère séquentiel de l'écrit permet de repérer les affixes. Pour Nagy et al. (1989), une des caractéristiques du bon lecteur, c'est sa capacité à reconnaître rapidement et efficacement les mots, ce qui est encore plus vrai pour les mots longs. Ainsi, parce que les mots longs sont des mots construits et plus complexes sur le plan morphologique, un manque de connaissance en morphologie affecte la lecture. Le lecteur habile, face à des mots longs qui sont peu fréquents, est capable de les décomposer en des parties qui elles, sont fréquentes et connues. L'utilisation du contexte et l'analyse morphologique sont, dans les faits, les principales sources d'information dont dispose le lecteur lorsqu'il rencontre un mot inconnu (Nagy et Scott, 2004). Enfin, la compréhension de mots complexes moins fréquents est facilitée par la connaissance de ses composantes morphémiques qui elles, sont fréquentes (Koda, 2000).

Les études en psychologie de l'apprentissage démontrent que ce ne sont pas les facultés comme la mémoire ou l'intelligence qui distinguent les experts des novices. Ce serait plutôt que les experts ont développé des connaissances approfondies qui influencent leur façon de porter attention et à leur façon d'organiser, de se représenter et d'interpréter les informations présentes dans l'environnement. Tout cela a pour effet d'influencer leurs habiletés à mémoriser, raisonner et solutionner des problèmes (Bransford et al., 1999). L'expert morcelle le problème, le découpe en morceaux qui représentent des patterns familiers, cela réduit la charge sur la mémoire de travail. Développer une compétence en morphologie dérivationnelle

suppose de reconnaître des régularités dans la langue, d'appliquer des stratégies d'analyse et de morcellement des mots complexes pour en déduire le sens. Comment la connaissance en morphologie dérivationnelle peut-elle contribuer à la compréhension et à l'acquisition d'items lexicaux en L2 et dans quelle mesure les apprenants l'appliquent-ils ? C'est ce dont il sera question dans la section qui suit.

2.4.1 Morphologie dérivationnelle et L2

On reconnaît maintenant un rôle à la morphologie dans la reconnaissance des mots en L2 (Frost et Grainger, 2000). Tout comme ce qui se passe pour la lecture en L1, face à un mot inconnu, l'apprenant qui connaît les structures morphologiques de la L2 peut utiliser ces connaissances pour en découvrir le sens (Durgunoglu, 1997). L'habileté à utiliser les relations morphologiques entre les mots procure donc un net avantage au lecteur face à un mot qui ne lui est pas familier (Morin, 2003). Nassaji (2004) a trouvé que le niveau de profondeur des connaissances lexicales, dont font partie les connaissances morphologiques, contribuait nettement au succès dans la déduction du sens de mots inconnus et ce de façon plus significative que les stratégies d'inférence textuelle. Dans la recherche de Parel (2004) auprès d'élèves du secondaire en anglais langue seconde, les problèmes de compréhension en lecture découlaient d'un manque d'habileté dans l'application de stratégies d'inférence lexicale. Parel (2004) a démontré que la compétence en lecture L2 peut être améliorée par l'apprentissage des suffixes les plus productifs ainsi qu'avec l'enseignement des principes de l'analyse morphologique.

Haastrup (1991) a établi que le type de stratégies d'inférence choisi par l'apprenant de L2 est à mettre en lien avec sa compétence dans la langue cible : les lecteurs plus compétents utilisent d'avantage d'indices linguistiques comme les préfixes et les suffixes, alors que les lecteurs moins avancés font plus souvent appel à des stratégies interlinguales et s'en tiennent à l'utilisation des morphèmes de base. Car

s'il est important d'apprendre beaucoup de mots, il faut aussi apprendre comment déduire le sens des mots inconnus, notamment par l'inférence lexicale (Haastrup, 1991).

Les sujets de la recherche de Parel (2004) mentionnée précédemment ne connaissaient pas tous les 2,000 mots les plus fréquents de la L2 et selon l'analyse de l'auteur, il faut avoir atteint un niveau seuil en termes d'étendue de vocabulaire pour pouvoir appliquer des stratégies lexicales, comme l'analyse morphologique.

« [...] accurate morphological analysis, even in conjunction with the context, is not possible in the absence of a threshold level of knowledge of core vocabulary. » (Parel, 2004, p. 849).

Pour le locuteur non natif, le problème de reconnaissance des mots serait surtout lié à la question de la transparence dans la construction (Bogaards, 1994). Cette préférence pour la stabilité formelle se voit chez les enfants apprenant une L2: ils s'en tiennent d'abord à des racines connues pour former des mots nouveaux, sans changements formels dans la base : ex : *poissonner (pour pêcher) (Clark, 1993, p. 120).

Il y aurait une corrélation entre la connaissance des affixes et la taille du vocabulaire (Schmitt et Meara, 1997; Mochizuki et Aizawa, 2000). Le vocabulaire, en s'accroissant, aide à la compréhension des affixes et en même temps, la connaissance des affixes aide à accroître le vocabulaire (Mochizuki et Aizawa 2000). Pressley et al. (1987) ont trouvé que les connaissances en morphologie contribuent non seulement à développer le vocabulaire mais aussi à consolider son apprentissage.

« Such findings suggest that remembering new vocabulary can indeed be facilitated through students' capitalization on previously learned word parts. »
(Pressely et al., 1987, p. 120).

La familiarité et la fréquence du mot, mais aussi la fréquence de ses composantes morphologiques jouent un rôle dans la reconnaissance des items lexicaux (Frost et Grainger 2000). Cependant, il semble que l'habileté à reconnaître et à manipuler les affixes soit une compétence plus longue à développer (Mochizuki et Aizama 2000).

Hancin-Bhatt et Nagy (1994) ont conduit une étude sur la connaissance morphologique chez des élèves bilingues espagnol-anglais, de 4^e, 6^e et 8^e année. Ils ont voulu vérifier l'habileté des locuteurs à reconnaître un congénère (*cognate*) à l'intérieur d'un mot suffixé anglais ainsi que leur connaissance d'une relation systématique entre suffixes anglais et espagnols. Ils ont émis l'hypothèse que pour deux langues assez apparentées, comme l'anglais et l'espagnol, les transferts entre L1 et L2 pourraient permettre non seulement de comprendre des mots individuels, mais aussi faciliter l'apprentissage de la morphologie dérivationnelle de la L2. Leurs résultats ont démontré là aussi un aspect développemental, à savoir que les sujets étaient capables de discerner une racine familière dans des mots dérivés, mais qu'ils n'étaient pas encore en mesure de voir le rôle syntaxique des suffixes apparentés L1-L2 dans les mots dérivés.

Le niveau d'aisance acquis en L2 n'est pas nécessairement une assurance que les notions de morphologie sont bien maîtrisées ou appliquées. Dans une recherche menée auprès de 70 étudiants universitaires finissants en traduction anglaise, Mirhassani et Toosi (2000) ont démontré que même des apprenants de ce niveau pouvaient améliorer leurs compétences lexicales et en compréhension de texte par l'apprentissage des principes de formation des mots dans la L2. Schmitt (1998), dans une recherche longitudinale auprès d'apprenants adultes avancés en anglais L2 a trouvé que, si les sujets obtenaient de bons résultats en orthographe, ils avaient

toutefois bien peu de connaissances en morphologie dérivationnelle. Les résultats obtenus par Meara et Schmitt (1997) dans leur étude réalisée auprès d'adultes japonais apprenants l'anglais L2 de niveau intermédiaire vont dans le même sens : entre le début et à la fin de la session, les apprenants avaient fait des gains très modestes sur le plan des connaissances morphologiques.

On peut supposer qu'il y aura des variations de comportement selon la L2 et selon la L1 de l'apprenant, car toutes les langues ne sont pas semblables quant à l'importance de la morphologie et quant à la façon de construire les mots et les morphèmes (Durgunoglu et Öney, 1999). En mandarin, par exemple, la forme des mots ne change pas, alors que dans plusieurs langues indo-européennes, les mots se modifient par l'ajout d'affixes : le turc, par exemple est très riche sur le plan morphologique (Clark, 1993). La distance typologique de la L1 et de la L2 pourrait expliquer le niveau de performance des apprenants. Cependant, dans une recherche comparant des Coréens et des Chinois apprenants de L2, Koda (1999) a trouvé des différences de performance entre les deux groupes, mais aussi beaucoup de différences entre les sujets d'un même groupe linguistique. Selon son analyse, ces différences individuelles s'expliquent par une plus grande expérience dans le traitement linguistique tant en L1 qu'en L2, ce qui contribue à développer une conscience de la structure interne des mots.

Lorsque deux langues sont riches sur le plan morphologique, alors la morphologie pourrait être un chemin commun entre les deux puisqu'elles demandent aux locuteurs une sensibilité à la construction des mots (Morcoso del Prado et al., 2004). Ainsi, pour deux langues apparentées et assez riches en morphologie, comme par exemple le français et l'espagnol, l'apprenant pourra utiliser des stratégies de transfert de la L1 vers la L2 et il est possible de le sensibiliser aux correspondances dans les formes morphémiques de L1 et L2 (Hancin-Bhatt et Nagy, 1994; Meara, 1993). Cependant, le fait que la langue maternelle et la langue cible soient

typologiquement proches n'assure pas que les apprenants utiliseront l'analyse morphologique pour décoder le sens de mots inconnus. C'est ce qu'ont trouvé Osburne et Mulling (2001) dans une étude portant sur des étudiants adultes hispanophones apprenant l'anglais L2. Pour les auteurs, c'est probablement parce que l'analyse morphologique est exigeante sur le plan cognitif que les apprenants n'y ont pas recours de façon systématique. Enfin, lorsque la L1 et la L2 sont proches, on pourra aussi constater des surgénéralisations et des formes erronées produites à partir des hypothèses sur les correspondances L1-L2 que s'est construites l'apprenant (Swan, 1997).

Malgré le rôle significatif que joue la morphologie dérivationnelle pour la compréhension et l'acquisition du lexique, nous venons de voir que plusieurs conditions doivent être réunies pour que l'apprenant de L2 utilise des stratégies lexicales. Même lorsque les conditions sont favorables, il n'est pas garanti qu'il aura soit la sensibilité métalinguistique nécessaire, soit la capacité ou la volonté de fournir l'effort cognitif requis.

2.5 Résumé

Nous résumons ici les questions principales ayant guidé notre recherche sur les connaissances lexicales et en morphologie dérivationnelle chez des locuteurs de L2 adultes. Tout d'abord, le lexique mental est une base de données complexe, constituée de toutes les unités lexicales que nous connaissons et de tout ce que nous connaissons sur ces unités. Chacune de ces unités est encodée avec ses caractéristiques, enregistrée, intégrée et reliée à divers réseaux, ce qui nous permet de la retrouver en mémoire et de l'utiliser. De la qualité de l'encodage des items dans les processus d'enregistrement et d'intégration vont dépendre les performances d'un sujet dans une tâche donnée. La morphologie fait partie des informations encodées dans l'entrée lexicale.

Si le cerveau humain a la capacité de gérer le lexique mental de plus d'une langue, cependant, l'acquisition du vocabulaire d'une L2 pose des défis, particulièrement chez l'apprenant adulte, qui possède déjà un lexique et un système conceptuel bien établis.

Les écrits et les rapports de recherche que nous avons consultés ont mis en évidence que les connaissances lexicales sont une partie importante de la compétence dans une L2. Les connaissances lexicales sont généralement décrites selon deux grandes dimensions interreliées, l'étendue et la profondeur. Plutôt que l'habituelle distinction entre vocabulaire passif et vocabulaire actif, il faut considérer les connaissances lexicales sur un continuum allant de la capacité à reconnaître un mot à celle de savoir le produire en situation. Le savoir lexical est complexe et il se construit progressivement, chaque exposition offrant une opportunité d'intégrer des éléments nouveaux et d'approfondir ses connaissances. Il demeure toutefois que les connaissances lexicales peuvent être limitées, incomplètes, voire fossilisées chez un locuteur.

Des recherches ont établi qu'il y a une base minimale quant au nombre de mots connus pour assurer une bonne compréhension dans la lecture d'un texte et pour que les habiletés acquises dans la L1 puissent être transférées dans la L2. De plus, tout locuteur qui aspire entreprendre des études supérieures a besoin de posséder des connaissances lexicales solides et étendues.

L'acquisition du vocabulaire dans une L2 est un long processus et la plupart des apprenants n'y investissent pas le temps nécessaire ou n'y appliquent pas l'attention requise. L'acquisition des items lexicaux se fait selon leur ordre de fréquence et il faut plusieurs expositions à un mot pour qu'il soit retenu.

La compétence en morphologie dérivationnelle suppose la capacité à déceler les régularités dans la forme et le sens des mots. La racine d'un mot constitue l'unité sémantique dans une famille morphologique et le suffixe fournit des informations sur la catégorie grammaticale. En français, une partie importante du lexique est constituée de mots construits par dérivation. On comprend alors que les connaissances en morphologie peuvent contribuer à accroître le vocabulaire d'un locuteur en français et qu'elles sont très utiles pour déduire le sens des mots inconnus. L'analyse morphologique demande cependant un effort cognitif et ce ne sont pas tous les apprenants de L2, même ceux qui sont avancés, qui l'utilisent. Pour le locuteur non natif, il serait plus facile de reconnaître la parenté morphologique lorsque le mot dérivé n'a pas subi de transformation formelle.

Dans le prochain chapitre, nous décrivons la démarche méthodologique que nous avons suivie pour évaluer les connaissances lexicales et en morphologie dérivationnelle de nos sujets. Nous présenterons d'abord les participants, puis les instruments de mesure et nous décrivons l'administration des épreuves utilisées. Nous terminerons ce chapitre avec la description du codage des données et des analyses statistiques.

CHAPITRE III

MÉTHODOLOGIE

3.1 Introduction

Notre objectif était de réaliser une recherche sur les connaissances lexicales et en morphologie dérivationnelle d'adultes LNN avancés en français et de les comparer à celles d'adultes locuteurs natifs.

Nous détaillerons dans ce chapitre la démarche méthodologique que nous avons appliquée pour réaliser notre recherche ainsi que les motifs expliquant ce choix. Nous ferons d'abord le portrait des participants à la section 3.2. À la section suivante, nous présenterons les instruments de mesure des connaissances lexicales et morphologie dérivationnelle que nous avons utilisés. À la section 3.3, nous décrirons les conditions dans lesquelles s'est faite la collecte des données. Enfin, nous traiterons de la compilation des données et des outils d'analyse.

3.2 Les participants

Les étudiants dont la langue maternelle n'est pas le français sont de plus en plus nombreux dans les universités francophones québécoises. Plusieurs de ces étudiants échouent au test d'admission en français et doivent suivre un cours de

rattrapage ou de mise à niveau afin de combler leur manque de connaissances du code écrit. Il y a aussi beaucoup d'étudiants francophones qui n'arrivent pas à satisfaire les exigences universitaires en matière de qualité de la langue et doivent eux aussi s'inscrire à des cours pour améliorer leur maîtrise de la grammaire du français. Il nous est donc apparu intéressant de mener notre recherche auprès de groupes d'étudiants d'un cours de mise à niveau en grammaire et de comparer les connaissances des étudiants non francophones avec celles des étudiants francophones faibles en français.

Les participants à notre recherche sont tous des étudiants inscrits à un cours de mise à niveau en grammaire du français écrit à une université francophone à la session d'automne 2005. Sur l'ensemble des étudiants du cours, 149 ont complété les épreuves à la fois en début et en fin de session et notre échantillon compte 78 étudiants locuteurs natifs du français (LN) et 71 locuteurs non natifs (LNN) du français.

Sept participants qui avaient déclaré avoir le français comme langue maternelle ont finalement été reclassés comme non francophones. Trois raisons nous ont motivée à les considérer comme tels : 1) ils avaient obtenu des scores très faibles (moins de 60%) dans les deux versions de la tâche de reconnaissance lexicale (Meara, 1992), conçue, pour des locuteurs non natifs; 2) Ils étaient nés à l'extérieur du Québec; 3) ils avaient déclaré maîtriser une autre langue que le français. Il faut ici mentionner qu'au Québec, le terme « francophone » est parfois utilisé pour décrire plusieurs réalités qui sont pourtant différentes: en principe, le terme « francophone » sert à désigner les Québécois dont la langue maternelle est le français et permet de les distinguer des autres Québécois. Cependant, il arrive que des personnes qui peuvent s'exprimer en français s'estiment francophones. Nous pensons que les sept sujets de notre recherche reclassés comme locuteurs non natifs se considéraient

probablement comme francophones à cause de leur capacité à s'exprimer en français.

La moyenne d'âge pour l'ensemble des sujets est de 27,17 ans. Mesurée selon les sous-groupes linguistiques, la moyenne d'âge est de 26,18 ans pour les francophones et de 29,85 ans pour les non francophones. Nous présentons dans le tableau 4.1 qui suit la répartition de tous les sujets selon le sexe et selon le groupe linguistique, locuteurs natifs (LN) et non natifs du français (LNN). Au tableau 3.1 ci-dessous, nous présentons la répartition des sujets selon le sexe et le groupe linguistique.

Tableau 3.1
Portrait de l'ensemble des sujets selon le sexe et le groupe linguistique

Sexe	Tous		LN		LNN	
	Nbre	%	Nbre	%	Nbre	%
F	86	57,72	55	70,51	31	43,66
M	63	42,28	23	29,49	40	56,34
Total	149	100,00	78	100,00	71	100,00

Pour l'ensemble des 149 sujets, il y a une plus grande représentation des sujets de sexe féminin. Cependant, chez les LNN, ce sont les sujets masculins qui sont les plus nombreux.

Le profil linguistique des sujets LNN est assez diversifié, avec toutefois la présence plus importante de locuteurs arabophones, comme nous le voyons dans le tableau 3.2.

Tableau 3.2
Langue maternelle déclarée des sujets LNN

LM	Nombre
Arabe	24
Anglais	6
Créole haïtien	5
Roumain	5
Espagnol	4
Berbère	4
Chinois	2
Albanais	2
Autres	19
Total	71

3.3 Le choix et l'élaboration des instruments de mesure

À notre connaissance, il existe peu d'instruments conçus pour évaluer les connaissances lexicales et en morphologie dérivationnelle en français langue seconde chez des sujets adultes avancés. Il nous a donc fallu, pour la plupart des tâches de notre épreuve, soit adapter des instruments existant pour l'anglais langue seconde, soit nous en inspirer.

Compte tenu des contraintes d'administration de l'épreuve, il fallait que celle-ci soit relativement rapide à compléter pour les sujets, tout en nous permettant de recueillir suffisamment de données, et ce, sur plusieurs aspects des connaissances. Le nombre de tâches de l'épreuve a été restreint et nous avons aussi limité le nombre d'items pour chaque tâche.

Nous avons élaboré deux versions équivalentes de notre épreuve et nous avons procédé à deux collectes de données. Cette façon de procéder en deux temps nous a permis de vérifier s'il y avait amélioration des connaissances de nos sujets LNN

entre le début et la fin de leur cours de grammaire. D'autre part, une deuxième collecte nous protégeait en quelque sorte contre les imprévus d'une cueillette qui devait se faire en très peu de temps, auprès de nombreux participants et par l'entremise de plus d'un administrateur.

Chaque épreuve comptait cinq tâches, deux visant les connaissances lexicales (tâches 1 et 2) et trois, les connaissances en morphologie dérivationnelle (tâches 3, 4 et 5). Toutes les tâches étaient de type papier-crayon. Chacune des tâches est décrite dans les sous-sections qui suivent.

La tâche 1 : connaissances lexicales réceptives

Cette tâche a pour objectif de mesurer les connaissances lexicales réceptives. Nous avons choisi l'épreuve de reconnaissance lexicale de Meara (1992), conçue pour le français langue seconde. Le test consiste en une liste de 60 items lexicaux, soit 40 mots réels et 20 pseudo-mots. Les pseudo-mots sont des mots possibles, selon la construction phonologique de la langue, mais ils sont dénués de sens. Dans cette tâche, le sujet doit simplement indiquer si oui ou non un item est un mot connu.

L'épreuve de Meara est construite avec 2,600 mots tirés du *Français fondamental* de G. Gougenheim (1956) et est constituée de plusieurs tests classés en deux niveaux. Le niveau 1 (1,070 items lexicaux) s'adresse aux débutants. Comme nos sujets étaient des apprenants avancés en français L2, nous avons choisi deux tests parmi ceux du niveau 2. Voici un exemple du test de Meara et de la consigne donnée aux participants:

- Cet exercice porte sur la **signification** des mots.

Lisez attentivement les mots de la liste.

Si vous connaissez la signification du mot, inscrivez O (pour oui) dans les parenthèses.

Si vous ne connaissez pas la signification du mot ou si vous n'êtes pas certain de sa signification, inscrivez N (pour non) dans les parenthèses.

Vous devez faire cet exercice rapidement (pas plus de 5 minutes).

Attention! Pour s'assurer que vous ne répondez pas au hasard, il y a aussi des mots imaginaires, qui n'existent pas en français, dans la liste.

- | | | |
|--------------|---------------|----------------|
| 1 [] choint | 2 [] délicat | 3 [] urgence |
| 4 [] avenir | 5 [] climat | 6 [] plassant |

Le test de Meara a l'avantage d'être d'application facile et rapide à compléter. Il permet aussi de cerner l'étendue du vocabulaire des sujets puisqu'il est conçu avec des items lexicaux d'une grande fréquence.

La tâche 2 : connaissances lexicales productives

Cette tâche a pour objet les connaissances lexicales productives. La tâche que nous avons utilisée s'apparente à un test de closure partiel, sauf que le contexte est une phrase et non pas un paragraphe ou un texte. Seule la première syllabe du mot est fournie, mais les phrases contiennent des indices qui permettent de déduire le sens du mot recherché. Nous avons construit cette tâche en nous inspirant du test de Laufer et Nation (1999). Exemples tirés du test conçu par Laufer et Nation (1999).

*"He has a successful ca_____ as a lawyer.
Many people are inj_____ in car accidents every year."*

Un des aspects que souhaitions évaluer dans notre recherche, c'était l'étendue du vocabulaire des sujets. Une façon répandue de mesurer l'étendue du vocabulaire est d'utiliser les rangs de fréquence des items lexicaux. Le test de Laufer et Nation (1999) est conçu en fonction de la fréquence des items lexicaux (2,000, 3,000, etc.). Comme les sujets de notre recherche étaient soit des étudiants locuteurs natifs du français, soit des locuteurs non natifs, apprenants de niveau avancé en français, à l'aise avec la L2 et non pas des apprenants débutants en langue seconde, les items de notre test ont été construits avec les 2,000 à 3,000 mots les plus fréquents, une base minimale pour un étudiant universitaire (Laufer, 1992). Les items lexicaux ont été sélectionnés dans les listes de fréquence disponibles sur le site internet de T. Cobb, professeur en linguistique à l'UQAM, ainsi que dans l'échelle Dubois-Buyse. L'échelle Dubois-Buyse regroupe 3,787 mots d'usage courant, choisis pour leur très grande fréquence dans des écrits spontanés.

Il nous fallait aussi contrôler la fréquence des contextes phrastiques dans lesquels étaient insérés les items lexicaux. Les phrases ont été construites à l'aide de concordanciers, tels que ceux disponibles dans la base de données textuelles *Lexiquum* de l'université de Montréal.

La tâche de production lexicale comprenait dix phrases avec un mot à compléter. Nous avons pris soin d'équilibrer et de varier les catégories grammaticales et, sur les dix items demandés, il y avait des adjectifs, des noms, des adverbes et des verbes. Pour ce qui est des verbes, la structure de la phrase ne demandait que des formes non conjuguées. Le contexte facilitait systématiquement la production du bon mot en orientant sémantiquement et grammaticalement les sujets vers la réponse recherchée. Voici quelques exemples de la tâche 2:

Tout était tranquille lorsque su _____, sans qu'on s'y attende, le chien s'est mis à gronder.
L'unique bou _____ qu'on avait allumée éclairait faiblement la pièce.

Dans une tâche de production libre, le sujet peut éviter d'utiliser des mots qu'il ne maîtrise pas bien (Laufer et Nation, 1999; Laufer et Paribakht, 1998), alors que dans une tâche de production contrôlée comme la tâche 2, le sujet n'a pas le choix, il n'y a qu'une réponse possible. Ce type de tâche offre donc la possibilité d'avoir un portrait assez précis et fidèle des connaissances des sujets. Cependant, cela ne fournit pas d'information sur les stratégies utilisées par les sujets, mais ne permet que d'évaluer ses connaissances à partir des résultats.

La tâche 3 : connaissances morphologiques productives

Dans cette tâche, nous souhaitons mesurer les connaissances productives en morphologie dérivationnelle. Nation (2001) propose une épreuve dans laquelle le sujet doit construire l'item ciblé avec le suffixe approprié:

« I went to the doctor for a _____ (consult). »

Afin de réduire le temps nécessaire à la réalisation de la tâche, nous avons choisi d'utiliser, essentiellement, le même type de construction que pour la tâche lexicale productive contrôlée. Au lieu de fournir la première syllabe du mot, c'est la base du mot que nous avons donnée et le sujet devait trouver le suffixe approprié. Nous avons contrôlé la fréquence des items lexicaux et des contextes phrastiques de la même manière qu'à la tâche 2. Nous avons de plus choisi les suffixes parmi les plus fréquents en français selon la classification de Huot (2001). Ici encore, nous avons fait attention à ce que les différentes catégories grammaticales soient présentes de façon équilibrée et à concevoir des phrases demandant un verbe à l'infinitif lorsque c'était l'item ciblé. Voici quelques exemples de la tâche 3:

Le chef du parti s'est engagé à atteindre plus de justice sans sacrifier _____ la liberté.

Il faut maintenir un taux d'humidité _____ d'environ 40% durant l'hiver.

La tâche 3 est une tâche de production contrôlée qui oblige le sujet à produire l'item ciblé. Elle a pour but d'évaluer si les sujets connaissent les suffixes fréquents en français et s'ils savent que certains suffixes servent à former des mots d'une catégorie grammaticale précise. Comme nous l'avons mentionné pour la tâche 2, ce type d'instrument nous permet d'évaluer les connaissances seulement à partir de ce que les sujets ont produit.

La tâche 4 : mots apparentés

La tâche 4 ciblait les connaissances en morphologie dérivationnelle sur le plan de la compréhension. Plus spécifiquement, nous voulions vérifier à quel point les participants étaient capables d'identifier la parenté de sens entre des mots d'une même famille. Pour ce faire, il faut pouvoir reconnaître la base du mot. La parenté de sens peut être plus difficile à identifier selon que la dérivation est plus ou moins transparente, soit sur le plan formel, soit sur le plan sémantique. Nous avons construit une tâche dans laquelle sont présentées 15 paires de mots, soit cinq paires de mots dont la dérivation est transparente, cinq paires de mots dont la dérivation n'est pas transparente et cinq paires de mots semblables au plan de la forme, mais sans parenté morphologique. Voici quelques exemples de la tâche 4:

Cette activité porte sur la signification des mots. Pour chaque paire de mots, vous devez indiquer s'il y a un lien, une ressemblance dans la signification.

Il y a deux réponses possibles :

0) **Il n'y a pas de ressemblance** de signification entre les deux mots

1) **C'est certain**, il y a une ressemblance de signification entre les deux mots

Cochez ✓

	0	1
parti-partisan		
sembler-semblable		
ouvrir-ouvrier		

La tâche 4 a l'avantage d'être rapide à compléter. Il y a cependant le risque que des participants répondent au hasard.

La tâche 5 : production morphologique

La tâche 5 avait pour objet les connaissances productives en morphologie dérivationnelle. Plus exactement, nous voulions évaluer les capacités des participants à produire des dérivés d'une même famille. Dans la tâche proposée par Nation (2001) ci-après, le sujet doit démontrer sa connaissance des familles morphologiques, produire des dérivés et savoir à quelle classe grammaticale ils appartiennent.

<i>Noun</i>	<i>Verb</i>	<i>Adjective</i>	<i>Adverb</i>
<i>argument</i>			
	<i>evaluate</i>		
		<i>distinct</i>	
			<i>normally</i>

Nous avons adapté la tâche de Nation (2001) et nous avons ajouté un exemple aussi explicite que possible afin de guider les sujets, supposant que certains d'entre eux ne maîtrisaient peut-être pas très bien le métalangage. Voici un exemple de ce qui était demandé dans cette tâche:

Complétez le tableau avec les mots de la même famille.

Exemple :

NOM	VERBE	ADJECTIF	ADVERBE
amour	aimer	aimable	aimablement

NOM	VERBE	ADJECTIF	ADVERBE
facilité			
	doubler		
		libéral	
			formellement

Les sujets avaient donc à produire douze mots en tout, soit trois formes d'une même famille pour chacun des quatre items. Les suffixes ciblés avaient été choisis parmi les plus fréquents selon la classification de Huot (2001).

La tâche 5 avait pour but d'évaluer la capacité à produire des dérivés d'une même famille et demandait de la part des sujets de connaître les suffixes servant à former des mots dans chacune des classes grammaticales. Pour cette tâche de production, les sujets ne bénéficiaient pas d'un contexte. La tâche 5 mesurait plusieurs choses à la fois avec le risque que cela comporte. Par exemple, il existe la possibilité qu'un participant sache comment un item lexical est formé sans toutefois savoir dans quelle catégorie grammaticale le situer.

Nous savons qu'il n'est pas possible de mesurer tous les aspects des connaissances lexicales dans une seule épreuve. Laufer et Nation (1999) considèrent que, pour avoir un bon portrait du savoir lexical, il faut différents types de mesures. Nous avons utilisé des tâches réceptives et productives ainsi que des tâches productives de type contrôlé. Notre épreuve comprenait des items lexicaux en contexte phrastique et sans contexte pour mesurer les aspects lexicaux et morphologiques. Nous présentons dans le tableau ci-après une synthèse des tâches

de notre épreuve indiquant les connaissances ciblées, le type de tâche et s'il s'agissait d'items lexicaux en contexte.

Tableau 3.3
Sommaire des tâches de l'épreuve

Tâche	Cible	Production	Compréhension	Items en contexte
1	Lexique		X	
2	Lexique	X		X
3	Morphologie	X		X
4	Morphologie		X	
5	Morphologie	X		

3.4 L'administration des épreuves

Nous avons exposé les objectifs de notre recherche à la responsable des cours de grammaire du français écrit de l'université, laquelle nous a autorisé à réaliser notre étude auprès des étudiants inscrits au cours *Connaissances de base en grammaire du français écrit*, un cours d'appoint en grammaire. Notre projet a ensuite été présenté aux professeures et chargées de cours qui ont accepté que les épreuves soient administrées dans leur groupe d'étudiants. Cependant, comme le programme de ce cours est très chargé et son plan très serré, la durée d'administration de l'épreuve devait être courte, c'est-à-dire, pas plus d'une vingtaine de minutes. Cette contrainte de temps a donc grandement influencé le choix des instruments de mesure. Les moments d'administration des épreuves, soit la 2^e et la 14^e semaine de la session, ont aussi été déterminés avec les enseignantes afin de perturber le moins possible le programme.

Les épreuves ont d'abord été pilotées auprès de dix étudiants de l'université (deux francophones, huit allophones) au mois d'août 2005, ce qui nous a permis par la

suite d'y apporter quelques petits ajustements et de vérifier le temps requis pour y répondre.

Les épreuves ont été administrées dans huit groupes d'étudiants inscrits au cours d'appoint en grammaire à la session d'automne 2005. En septembre, la moitié des groupes a complété la version 1 de l'épreuve, l'autre moitié la version 2. Les versions ont été alternées en décembre. À la fin de la session, nous avons eu accès aux notes obtenues au cours et nous avons pu les analyser en fonction des résultats de nos épreuves.

Les étudiants étaient libres de participer à l'étude et les volontaires devaient signer un formulaire de consentement éclairé, validé par les autorités universitaires. Le but de la recherche a été présenté comme étant de trouver des moyens d'aider les étudiants en français et d'améliorer les cours. Nous avons assuré les étudiants que les résultats au test ne comptaient pas pour la note du cours de grammaire et qu'ils devaient le compléter au meilleur de leurs connaissances. Soulignons que nous avons obtenu une très bonne participation des sujets pour qui une recherche sur l'enseignement du français était importante et près de la moitié des étudiants inscrits au cours de grammaire ont complété le test aux deux moments d'administration.

La passation du test s'est faite durant la période de cours et les étudiants avaient été avisés du moment de notre présence. Étant donné qu'il y avait plusieurs groupes et que certains cours avaient lieu au même moment, la collecte s'est faite soit par la chercheuse, soit par une étudiante ou une professeure en linguistique. Nous avons produit un guide écrit des consignes à observer pour l'administration de l'épreuve afin de nous assurer que tous les sujets aient la même information et que la collecte de données se réalise le plus possible dans les mêmes conditions dans tous les groupes.

Les sujets disposaient d'une trentaine de minutes en tout (incluant la lecture et la complétion du formulaire de consentement éclairé ainsi que des questions d'identification personnelle) pour réaliser l'épreuve. Tous ont été en mesure de compléter l'épreuve dans le temps prescrit. Les ouvrages de références, tels que dictionnaire et grammaire étaient interdits. La première tâche du test (connaissance lexicale) était minutée, les étudiants ayant un maximum de cinq minutes pour la compléter.

3.5 Codage des données

Lors de la première administration de l'épreuve, nous avons attribué un numéro à chaque sujet qui l'avait complété. À la deuxième administration, nous n'avons conservé que les épreuves des sujets qui avaient complété les deux versions du test. Les épreuves incomplètes ont été éliminées. Toutes les données ont été compilées dans le logiciel Excel. Pour chaque participant, nous avons aussi entré les informations concernant son sexe, son âge et sa langue maternelle. Les données compilées dans Excel ont ensuite été classées en deux groupes linguistiques, c'est-à-dire, francophone et non francophone.

3.5.1 La tâche 1: connaissances lexicales réceptives

Les résultats globaux sur 100 ont été calculés avec la grille de Meara (1992) qui accompagne ce test. Nous avons de plus compilé les erreurs pour les vrais mots et pour les pseudo-mots. Nous avons calculé le score pour chaque sujet, la moyenne pour chacun des groupes linguistique ainsi que l'écart type.

3.5.2 La tâche 2 : connaissances lexicales productives

Toutes les productions des participants ont été compilées. Après leur analyse, nous avons établi une codification, ce qui a donné ce qui suit :

0= pas de réponse

1= mot inexistant

2 = mauvaise catégorie et erreur de sens

3= mauvaise catégorie

4= bonne catégorie, mais erreur de sens

5= bonne réponse

Nous avons d'abord calculé le taux de réussite à cette tâche pour chacun des participants, puis fait la moyenne et calculé l'écart type pour l'ensemble des sujets selon le groupe linguistique. Nous avons ensuite établi les scores selon notre codification pour chacun des sujets, puis fait la moyenne par groupe linguistique.

3.5.3 La tâche 3 : connaissances morphologiques productives

Toutes les productions ont été compilées et nous avons ensuite établi la codification suivante :

0= pas de réponse

1= erreur dans la suffixation

2= erreur de catégorie

3= erreur lexicale

5= bonne réponse

Nous avons calculé le taux de réussite pour chacun des participants, puis établi la moyenne selon le groupe linguistique ainsi que l'écart type. Nous avons calculé les

scores selon notre codification pour chacun des sujets, puis fait la moyenne pour tous les sujets par groupe linguistique.

3.5.4 La tâche 4 : mots apparentés

Nous avons calculé d'abord le score total pour chacun des sujets. Nous avons ensuite classé les erreurs selon qu'il s'agissait d'une paire réelle de mots ou d'une pseudo-paire de mots. Nous avons calculé l'écart-type pour chacun des groupes linguistiques et établi la moyenne tant pour le score global que pour le taux d'erreur.

3.5.5 La tâche 5 : production morphologique

Nous avons d'abord compilé toutes les productions et nous avons ensuite établi la codification suivante pour distinguer le type d'erreur commise:

- 2= mot inexistant
- 3= erreur de catégorie
- 4= erreur lexicale

Nous avons calculé le score total pour chacun des participants puis la moyenne par groupe linguistique. Nous avons ensuite établi le taux d'erreur pour chacun des sujets et fait la moyenne pour chacun des groupes.

3.6 Outils d'analyse

Nous avons utilisé le logiciel *Statistical Package for the Social Sciences* (SPSS) pour procéder aux analyses statistiques de nos résultats. Comme notre objectif était de comparer les performances des sujets LN et des sujets LNN, nous avons appliqué des analyses de variance afin de vérifier si les écarts étaient significatifs, et ce, pour toutes les tâches de l'épreuve de même que pour les résultats au cours d'appoint en

grammaire. Nous avons calculé l'écart type pour les meilleurs résultats à toutes les tâches pour le groupe des LNN et celui des LN.

Nous avons appliqué un test t pour vérifier l'existence de différences significatives entre les résultats à l'épreuve aux deux moments d'administration pour chacun des groupes linguistiques. De plus, un test ANOVA à deux facteurs a été appliqué afin de mesurer les effets possibles des versions du test et du moment de leur administration.

Nous avons vérifié s'il y avait des différences significatives entre les performances des deux groupes linguistiques selon le type de connaissances ciblées (lexique, morphologie) et le type de tâche (production, compréhension) et nous avons appliqué un test de Pearson afin d'évaluer l'existence de corrélations entre les tâches visant le même type de connaissance. Enfin, nous avons vérifié à l'aide d'un test de Pearson l'existence de corrélations entre le score obtenu pour chacune des tâches et la note finale au cours d'appoint en grammaire.

CHAPITRE IV

PRÉSENTATION DES RÉSULTATS

La présentation des résultats obtenus dans le cadre de notre recherche sera l'objet de ce chapitre. Notre travail était motivé, à la base, par le besoin de connaître l'état des connaissances lexicales et en morphologie dérivationnelle chez des sujets adultes locuteurs non natifs du français désirant entreprendre des études dans la L2. Nous rappelons ici les hypothèses que nous avons formulées lors de notre analyse de la recherche dans les domaines pertinents :

- 1) Les sujets locuteurs non natifs seront plus faibles que les locuteurs natifs sur le plan des connaissances lexicales et en morphologie dérivationnelle;
- 2) Il n'y aura pas de différence significative de leurs performances entre les deux temps de passation du test;
- 3) Il y aura une corrélation entre les résultats aux épreuves et les résultats obtenus au cours de rattrapage en grammaire du français écrit.

À la section 4.1, nous brosserons un premier portrait des connaissances des sujets LN et LNN avec la présentation des résultats globaux pour l'ensemble des tâches. Nous présenterons aussi les résultats de tests statistiques que nous avons appliqués.

À la section 4.2, nous ferons une analyse plus détaillée de chacune des tâches afin de comparer de façon plus fine les comportements des LN et des LNN et dans le but d'identifier à quoi peuvent être attribuables les différences dans leurs performances. Puis nous présenterons les résultats obtenus par les sujets au cours d'appoint en grammaire du français écrit à la section 4.3.

4.1 Résultats globaux

Il faut tout d'abord rappeler que nous avons utilisé deux versions de chaque tâche. Une version a été administrée à un groupe de sujets en septembre, alors que l'autre groupe avait à compléter la deuxième version. Lors de la passation de décembre, les versions ont été inversées. Ainsi, quatre groupes ont reçu la version 1 en septembre (temps 1) et la version 2 en décembre (temps 2), alors que, pour les quatre autres groupes, ce fut la version 2 au temps 1 et la version 1 au temps 2.

Nous présentons, au tableau 4.1, les résultats moyens (M) en pourcentage pour chaque version du test (version 1 et version 2), en pré (T1) et en post test (T2) et l'écart type (é.t.), ainsi que la moyenne pour les deux versions pour chacun des groupes linguistiques.

Tableau 4.1
 Comparaison des résultats, temps 1 et temps 2

Version	LN (n=78)				LNN (n=71)			
	T1		T2		T1		T2	
	M	é.t.	M	é.t.	M	é.t.	M	é.t.
Tâche 1								
V1T1 – V2T2	85,23	10,80	83,43	9,53	79,00	18,68	65,42	21,59
V2T1 – V1T2	78,00	15,39	85,74	12,89	68,50	17,86	73,70	20,29
<i>Moyenne</i>	<i>81,62</i>	<i>13,10</i>	<i>84,59</i>	<i>11,21</i>	<i>73,75</i>	<i>18,27</i>	<i>69,56</i>	<i>20,94</i>
Tâche 2								
V1T1– V2T2	86,29	12,39	71,14	15,86	59,39	25,26	42,12	24,40
V2T1 – V1T2	72,56	12,44	81,40	14,07	43,74	21,86	56,83	26,00
<i>Moyenne</i>	<i>79,43</i>	<i>12,42</i>	<i>76,27</i>	<i>14,97</i>	<i>51,57</i>	<i>23,56</i>	<i>49,48</i>	<i>25,20</i>
Tâche 3								
V1T1 – V2T2	90,86	9,81	77,71	12,15	71,21	20,48	59,39	17,87
V2T1 – V1T2	80,00	9,76	93,49	8,13	61,62	18,68	68,40	25,63
<i>Moyenne</i>	<i>85,43</i>	<i>9,79</i>	<i>85,60</i>	<i>10,14</i>	<i>66,42</i>	<i>19,58</i>	<i>63,90</i>	<i>21,75</i>
Tâche 4								
V1T1 – V2T2t	89,14	18,16	91,62	6,73	80,40	18,15	81,62	11,99
V2T1 – V1T2	91,78	7,10	94,57	6,71	82,10	12,98	80,24	21,84
<i>Moyenne</i>	<i>90,46</i>	<i>12,63</i>	<i>93,10</i>	<i>6,72</i>	<i>81,25</i>	<i>15,57</i>	<i>80,93</i>	<i>16,92</i>
Tâche 5								
V1T1– V2T2	61,84	29,67	71,67	20,73	68,43	17,21	69,44	22,22
V2T1 – V1T2	68,80	18,55	71,71	19,26	70,20	14,58	59,00	23,60
<i>Moyenne</i>	<i>65,32</i>	<i>24,11</i>	<i>71,69</i>	<i>20,00</i>	<i>69,32</i>	<i>15,90</i>	<i>64,22</i>	<i>22,91</i>

La tâche 1 en était une de connaissance lexicale, la tâche 2, de production lexicale. La tâche 3 ciblait les connaissances en morphologie dérivationnelle (production), alors que dans la quatrième tâche, il s'agissait de reconnaître la parenté sémantique entre deux mots d'une même famille morphologique. Enfin, dans la tâche 5, les sujets devaient produire des mots d'une même famille morphologique.

Il y a présence de trois facteurs dans le tableau 4.1: le groupe linguistique, la version du test et le moment de passation. Exception faite de la tâche 5, nous pouvons constater que les résultats des LNN sont inférieurs à ceux des LN pour toutes les tâches et quel que soit le moment de passation de l'épreuve.

Un test t comparant les résultats au temps 1 et au temps 2 pour chacun des groupes linguistiques n'a révélé aucune différence significative pour les LNN et, pour les LN, une différence significative ($p < ,001$) uniquement pour la tâche 5.

Étant donné que deux versions du test ont été utilisées, mais à des moments différents, nous avons vérifié si l'ordre de passation et la version auraient pu avoir un effet sur les performances des sujets. À cette fin, pour chaque tâche, nous avons comparé les moyennes et appliqué un test d'analyse de variance (ANOVA) à deux facteurs (groupe et version) afin de vérifier l'effet possible de chacun. De plus, nous avons mesuré l'effet possible de la combinaison groupe et version. Les résultats sont présentés dans le tableau 4.2.

Tableau 4.2
Analyse de variance des facteurs groupe et version

	PRÉ		POST	
	F(1, 145)	p-valeur	F(1, 145)	p-valeur
Tâche 1				
Version	11.33	0.0010	3.74	0.055
Groupe	8.91	0.0033	30.24	<.0001
Version*Groupe	0.38	0.5382	1.18	0.2788
Tâche 2				
Version	22.49	<.0001	13.98	0.0003
Groupe	80.73	<.0001	64.35	<.0001
Version*Groupe	0.10	0.7499	0.45	0.5045
Tâche 3				
Version	17.15	<.0001	19.36	<.0001
Groupe	59.19	<.0001	59.24	<.0001
Version*Groupe	0.06	0.8050	1.43	0.2333
Tâche 4				
Version	0.87	0.3526	0.12	0.7281
Groupe	15.65	0.0001	31.52	<.0001
Version*Groupe	0.04	0.8404	1.02	0.3135
Tâche 5				
Version	1.65	0.2010	2.22	0.1384
Groupe	1.38	0.2413	4.57	0.0343
Version*Groupe	0.59	0.4426	2.25	0.1355

Nous voyons que les différences de performance entre les deux groupes linguistiques sont significatives pour toutes les tâches, sauf la tâche 5. Autrement dit, l'avance des LN sur les LNN que nous avons observée dans nos résultats bruts est significative pour quatre des cinq tâches. De plus, cette avance n'est pas attribuable à la version du test, que ce soit au temps 1 ou au temps 2, car il n'y a pas d'interaction significative entre les deux facteurs. Nous sommes donc en mesure de confirmer la supériorité des LN sur les LNN dans toutes les tâches, à l'exception de celle demandant la production de mots d'une même famille morphologique (la tâche 5).

Par contre, il y a des différences significatives lorsque l'on combine version et groupe pour certaines tâches et dans certaines versions: les tâches 1, 2 et 3 au temps 1, les tâches 2 et 3 au temps 2. Ainsi, il y a eu des différences de performance soit parce qu'une des versions était plus difficile, soit parce que des sujets ont contre performé pour certaines tâches, dans l'une ou l'autre version du test. Autrement dit, les versions que nous croyions équivalentes ne l'étaient pas pour tous les participants et pour toutes les tâches.

Afin d'explorer les différences entre les deux versions en plus de détails, nous avons compilé les scores de tous les sujets selon les versions du test. Le tableau 4.3 qui suit présente le bilan des résultats pour chaque groupe linguistique et pour chacune des versions ainsi que la moyenne des deux versions.

Tableau 4.3
Comparaison des résultats selon le groupe et selon la version du test

	v1		v2		M v1-v2	
	LN	LNN	LN	LNN	LN	LNN
Tâche 1	85,45	76,35	80,70	66,95	83,08	71,65
Tâche 2	84,35	58,10	71,85	42,90	78,10	50,50
Tâche 3	92,20	69,80	78,85	60,50	85,53	65,15
Tâche 4	91,85	80,30	91,70	81,85	91,78	81,08
Tâche 5	66,75	63,70	70,25	69,80	68,50	66,75
M ensemble des tâches	84,12	69,65	78,67	64,40	81,40	67,03

À la lecture des données du tableau 4.3, nous pouvons voir que la version 2 a généralement été moins bien réussie par les deux groupes, ce qui suggère que celle-ci était plus difficile que la version 1, du moins, pour certains individus et pour certaines tâches. Plus problématique est le fait que nous avons relevé des écarts importants entre les résultats chez les mêmes sujets, même pour des tests normalisés comme le test de connaissances lexicales de Meara (1992). Par

exemple, un participant a obtenu des scores de bonnes réponses de 81% à une version et de 66% à l'autre. D'autres sont passés de 95 à 66% et de 38 à 82% à la même tâche. Nous avons relevé des scores qui sont passés de 100 à 74%, de 71 à 94%, de 50 à 87%, pour la tâche 4. Comme les tâches que nous avons administrées ne comportaient pas de difficultés extraordinaires et que la plupart des participants ont obtenu des scores comparables aux deux versions du test, on peut supposer que des sujets qui ont obtenu deux résultats extrêmement différents n'ont pas complété le test avec la même application aux deux moments de passation et que, dès lors, leurs résultats ne reflètent pas avec exactitude leurs compétences et leurs connaissances. Nous aurions pu éliminer tous ces participants afin de régler le problème, mais ils étaient nombreux et en les écartant, nous aurions introduit d'autres problèmes. Nous avons finalement choisi de ne conserver que le meilleur résultat obtenu par chaque participant à chaque tâche. Notre volonté était de préserver un critère commun de sélection tout en éliminant les contre-performances qui ne pouvaient pas vraiment être attribuables à l'incapacité de faire la tâche. Ce sont ces meilleurs résultats pour chacun des groupes linguistiques que nous présentons au tableau 4.4 ci-dessous.

Tableau 4.4
Comparaison des meilleurs résultats pour toutes les tâches

	Tâche 1		Tâche 2		Tâche 3		Tâche 4		Tâche 5	
	LN	LNN								
n	78	71	78	71	78	71	78	71	78	71
<i>M</i>	88,99	80,93	85,51	60,15	92,31	74,51	95,47	87,04	79,74	77,58
écart type	8,16	14,09	11,47	23,75	8,82	18,11	5,93	8,94	14,68	14,65

Nous voyons que même après avoir exclus les contre-performances, des écarts, parfois importants, subsistent entre les deux groupes. Selon l'analyse de variance, les différences de résultats sont significatives pour les tâches 1 ($F= 18,649$, $dl = 148$, $p < ,001$), 2 ($F= 70,870$, $dl = 148$, $p < ,001$), 3 ($F= 59,817$, $dl = 148$, $p < ,001$) et 4 ($F= 46,750$, $dl = 148$, $p < ,001$), mais pas pour la tâche 5.

4.2 Analyses détaillées

Nous présenterons maintenant les performances des sujets de façon détaillée pour chacune des tâches de l'épreuve. Notre objectif était d'identifier de façon plus fine ce qui a posé problème aux sujets ainsi que ce qu'ils ont mieux réussi.

4.2.1 Tâche 1 : connaissances lexicales réceptives

Dans la tâche 1, un test de reconnaissance lexicale conçue par Meara (1992) pour des apprenants de L2, le sujet devait répondre par oui ou non s'il connaissait les items qui lui étaient soumis. Sur les 60 items lexicaux, 40 étaient des mots réels et 20 des pseudo-mots. Nous avons utilisé la grille de notation de Meara (1992) pour établir le score de chacun des sujets. Dans cette notation, chaque erreur soustrait des points au total des bonnes réponses.

Meara (1992) estime que les résultats à ce test doivent être interprétés en termes de pourcentage des mots qui sont connus : il s'agit donc ici d'une évaluation de la taille du vocabulaire. Le score des LNN indiquerait donc qu'ils connaissent, en moyenne, un peu plus de 80% des mots de haute fréquence, tirés du *Français fondamental*, utilisés pour bâtir ce test.

Dans la tâche 1, il y avait quatre types de réponses possibles: deux types de bonnes réponses lorsque le sujet identifiait correctement un mot réel ou un pseudo-mot; deux types d'erreurs lorsque le sujet ne reconnaissait pas un item comme étant un mot réel ou un pseudo-mot. Nous avons relevé pour chaque groupe linguistique le nombre moyen d'erreurs par type (erreurs de jugement pour les vrais et les pseudo-mots). Les résultats sont présentés dans le tableau 4.5 ci-dessous.

Tableau 4.5

Tâche 1 : comparaison du nombre moyen d'erreurs selon le type

	Vrais mots (n=40)		Pseudo-mots (n=20)	
	LN (n=78)	LNN (n=71)	LN (n=78)	LNN (n=71)
<i>M</i>	0,23	0,51	2,08	3,46
Écart type	0,60	0,94	1,53	2,82

Les LNN ont commis deux fois plus de fautes que les LN dans l'identification des mots réels, mais leur pourcentage d'erreur n'est pas très élevé et la différence entre les deux groupes n'est pas significative. Tant les LN que les LNN ont commis, en moyenne, plus d'erreurs pour les pseudo-mots que pour les mots réels. Dans le cas des LNN, une moyenne d'erreur de 3,46 sur 20 pseudos mots signifie que dans un peu plus de 17% des cas, ils ont faussement identifié des items lexicaux comme étant français. Ainsi, le score total plus faible obtenu par les LNN s'expliquerait surtout par leurs erreurs pour les pseudo-mots. Un test ANOVA nous a permis d'établir la signifiante de la différence de performance entre les deux groupes dans le cas des pseudo-mots ($F = 2,044$, $df = 148$, $p = 0,029$).

4.2.2 La tâche 2: connaissances lexicales productives

Nous avons vu que dans la tâche de production lexicale, la performance des LNN a été nettement inférieure à celle des sujets francophones et que l'écart entre les deux groupes est significatif. Dans cette tâche de production contrôlée, il n'y avait qu'une seule bonne réponse possible et il s'agissait toujours d'un item lexical fréquent et à la portée d'un étudiant universitaire. Il nous reste à tenter de savoir ce qui a causé des difficultés aux sujets non natifs: problème de compréhension de la phrase ou absence de connaissance du mot requis, ou encore, connaissance limitée sur le plan sémantique. Et puis, certains items lexicaux ont-ils été moins bien réussis que d'autres ?

Dans la tâche 2, qui est un test de closure partielle, les sujets devaient compléter un mot dans le contexte d'une phrase. Seule la première syllabe du mot était donnée, mais les phrases fournissaient des indices sur l'item ciblé. Rappelons que les dix phrases avaient été construites à l'aide de concordanciers afin de s'assurer de leur fréquence d'apparition dans ces contextes et dans la langue en général.

Pour le traitement des données, nous avons d'abord compilé toutes les productions de tous les sujets afin d'établir une grille de codification et de classement des erreurs. Nous avons relevé des productions dans lesquelles des sujets ont commis des erreurs dans la forme du mot recherché, c'est-à-dire, lorsqu'ils ont fourni un mot formé avec la base du mot cible, mais dans la mauvaise catégorie grammaticale. Il est arrivé aussi que des sujets produisent un mot qui n'était pas celui recherché et que de plus, il soit de la mauvaise catégorie grammaticale. D'autre part, certaines fois, la forme produite respectait la grammaire, mais non le sens et, enfin, des sujets ont produit des mots qui, même si l'orthographe n'était pas en prise en compte, ne correspondaient à rien d'existant en français. Voici un rappel de la codification que nous avons utilisée :

0= pas de réponse

1= mot inexistant

2 = mauvaise catégorie et erreur de sens

4= bonne catégorie, mais erreur de sens

5= bonne réponse

Pour la tâche 2, plus de 75% des LN et des LNN ont obtenu de meilleurs résultats dans la version 1 du test et 25%, dans la version 2. Dans le tableau 4.6 qui suit, nous comparons les performances de chacun des groupes de sujets par type de réponse.

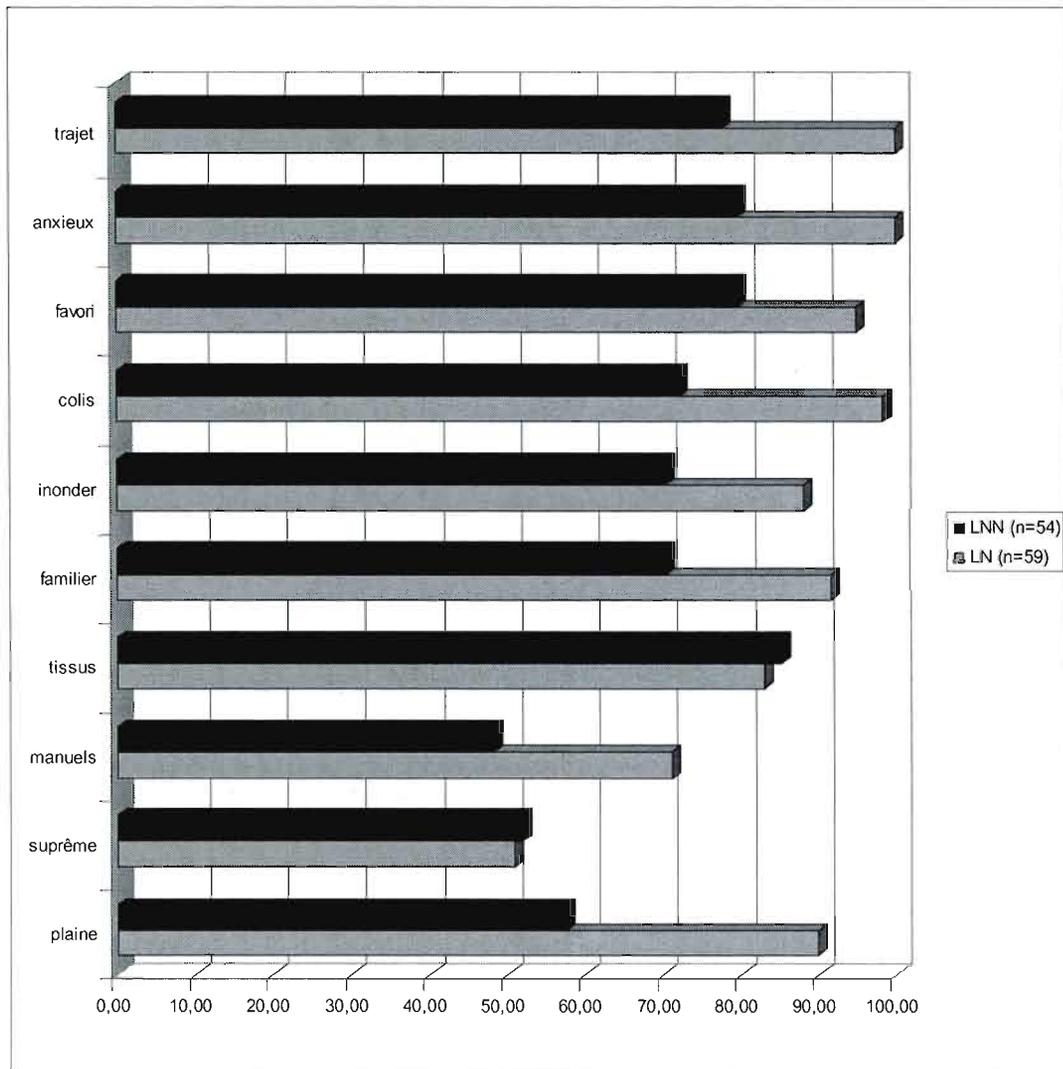
Tableau 4.6
Tâche 2, comparaison des résultats selon le type de réponse

Réponse	Version1		Version 2	
	LN (n=59)	LNN (n=54)	LN (n=19)	LNN (n=17)
0	3,90	15,00	5,79	31,74
1	0,68	3,70	0,53	7,06
2	1,36	1,67	3,16	2,94
4	7,29	10,37	5,79	7,42
5	86,78	69,24	84,74	51,02

Nous voyons que les LNN ont été plus fréquemment que les LN dans l'impossibilité de fournir une réponse quelle qu'elle soit. Ces derniers n'ont pas été en mesure de tirer parti du contexte de la phrase ni de l'indice fourni par la première syllabe du mot dans plus de 25% des cas en moyenne, ce qui est assez élevé. De plus, si l'on additionne les réponses de type 2 et 4, dans plus de 10% des cas, la forme fournie n'était pas celle que le sens de la phrase demandait. Enfin, il faut observer dans la version 2, un taux relativement important de production de mots qui n'existent pas en français. Les LN n'ont pas connu, comme principale difficulté, l'absence totale d'intuition ou de connaissance pour le mot ciblé comme cela semble avoir été fréquemment le cas chez les LNN.

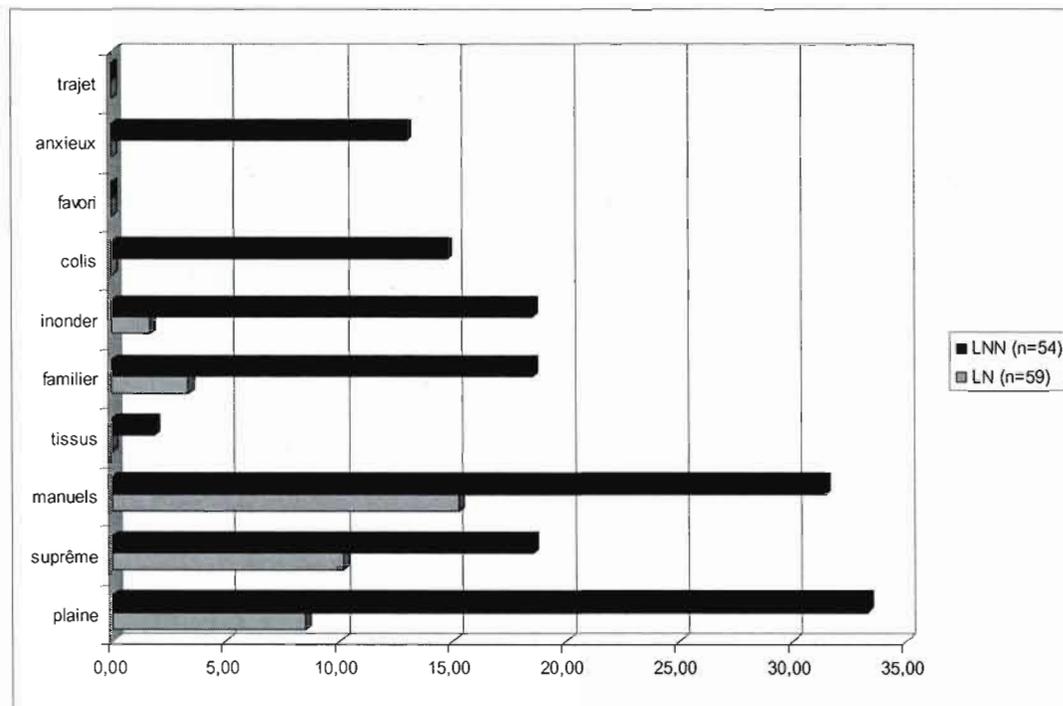
Dans ce qui suit, nous verrons quels sont les items lexicaux qui semblent avoir posé le plus de difficulté aux LNN comparativement aux LN. Tout d'abord, nous présentons à la figure 4.1 le taux de réussite (code « 5 ») pour chacun des items à la version 1.

Figure 4.1
Tâche 2, version 1 : comparaison du taux de réussite par item



Les données de la figure 4.1 nous font voir que pour tous les items, sauf « suprême », le taux de réussite est supérieur chez les LN. Nous avons vu, au tableau 4.6, qu'en moyenne, les LNN ont été plus nombreux à ne fournir aucune réponse. Quel a été le comportement des LNN pour chacun des items à cet égard? À la figure 4.2 ci-dessous, nous comparons les pourcentages d'absence de réponse (code « 0 ») chez les deux groupes pour chacun des items de la version 1.

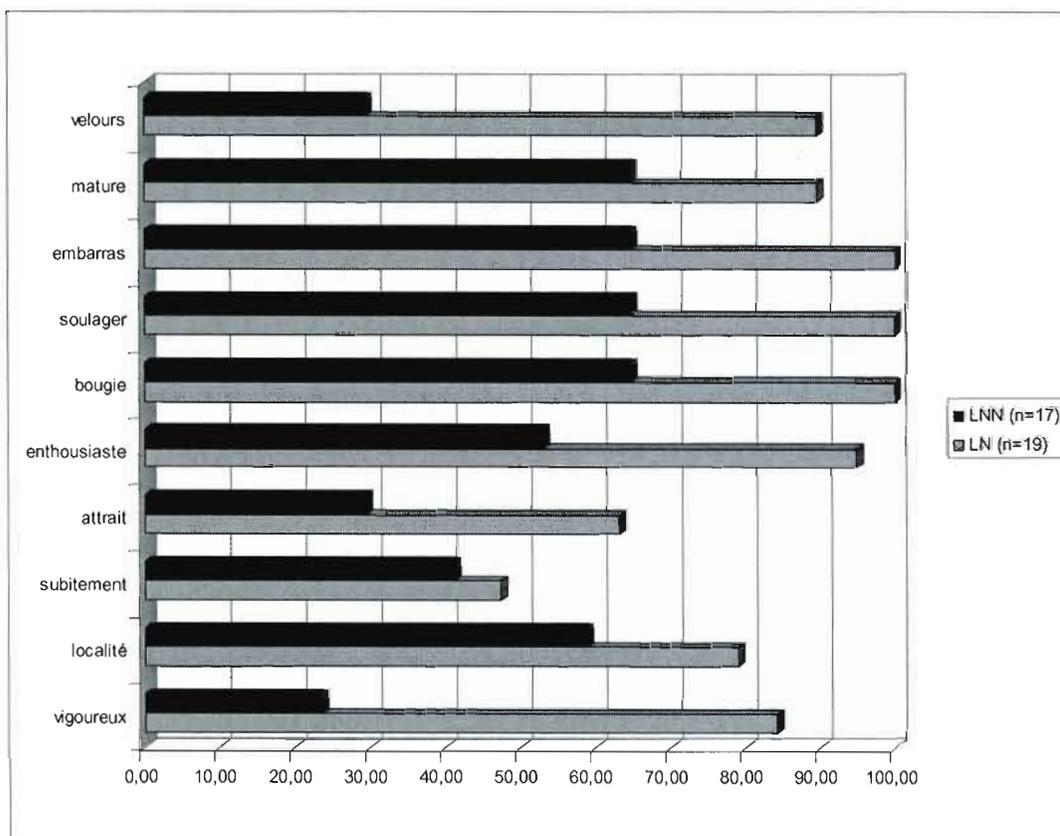
Figure 4.2
Tâche 2, version 1 : comparaison du taux d'absence de réponse par item



Nous pouvons constater que le taux d'absence de réponse chez les LNN est supérieur à celui des LN pour tous les items. Ce taux dépasse 25% pour « manuels » et « plaine ».

Nous présentons maintenant les résultats pour la version 2 de la tâche. Dans la figure 4.3 qui suit, nous comparons le taux de réussite (code « 5 ») pour chacun des items.

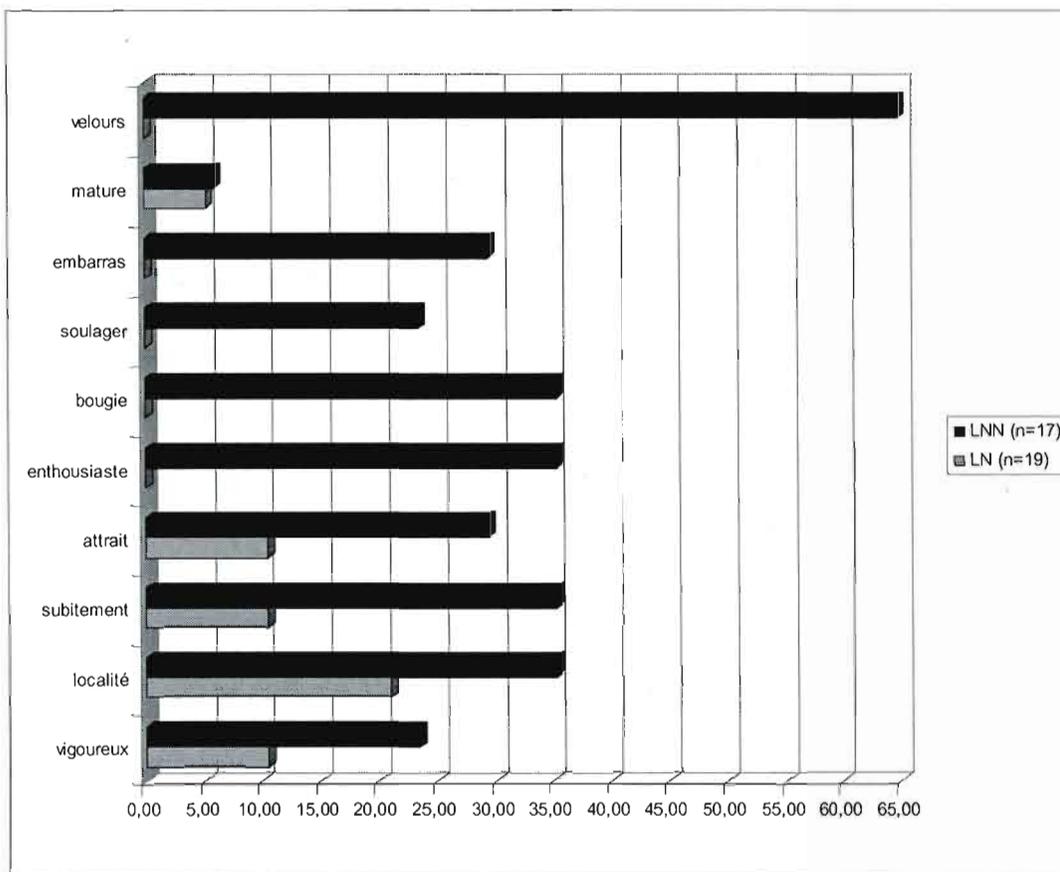
Figure 4.3
Tâche 2, version 2 : comparaison du taux de réussite par item



Les graphiques de la figure ci-dessus font voir que dans la version 2 aussi, les LNN ont obtenu de moins bons résultats que les LN, et ce, pour tous les items. Voyons maintenant, à la figure 4.4 ci-dessous, quel a été le taux d'absence de réponse (code « 0 ») pour chacun des groupes linguistiques et pour chacun des items de la version 2.

Figure 4.4

Tâche 2, version 2: comparaison du taux d'absence de réponse par item



Dans la version 2, les LNN ont produit un taux d'absence de réponse supérieur à celui des LN pour tous les items. Pour sept des dix items de la tâche, ce taux excède 25%.

Comme nous l'avons déjà mentionné, la tâche 2 contenait des mots des différentes catégories grammaticales, et demandait, lorsque c'était le cas, des formes verbales non conjuguées. Nous pouvons constater que seulement un item atteint le cap du 80% de réussite pour les LNN (« tissus ») et que cinq sur vingt sont sous le 50%, soit les adjectifs « manuels » et « vigoureux », les noms « velours » et « attrait »

ainsi que l'adverbe « subitement ». Mentionnons que la contre performance des locuteurs natifs pour l'item « enthousiaste » est due dans la majorité des cas, à la production de la forme adjectivale « enthousiasme », une confusion avec la forme nominale. Quant à l'item « attrait », le faible taux de réussite s'explique par la fréquente production de la forme « attraction », forme qui n'a pas été acceptée.

Pour tous les items lexicaux; le taux de non réponse est plus élevé chez les LNN. Ce taux atteint ou dépasse 30% pour sept des 20 items, soit : « plaine », « subitement », « bougie », « enthousiaste », « localité », « velours » et « manuels ». Le taux élevé de non réponse pour le mot cible « velours » est étonnant puisqu'il s'agit d'une expression très courante en français : « douce comme du velours ». Dans l'ensemble, nous pouvons constater que c'est l'absence de réponse qui est la principale explication du faible score des LNN pour ces items lexicaux de la tâche 2.

Nous avons soumis la liste des items lexicaux ciblés dans la tâche 2 à une base de données lexicales afin de vérifier si la fréquence pouvait expliquer les performances des LNN. La base informatisée *Lexique 2* (New et al., 2001) contient environ 130,000 items lexicaux ayant une forme orthographique distincte. Elle a été constituée à partir de 3,200 textes littéraires.

Dans le tableau 4.7 ci-dessous, nous présentons les résultats de notre requête. La fréquence est donnée par million d'occurrences pour la base *Lexique*. Nous avons identifié par un X dans la colonne de droite les items lexicaux pour lesquels le taux de réussite des LNN a été de 50% ou moins.

Tableau 4.7
Tâche 2 : fréquence des items lexicaux .

Item	Fréquence	Moins de 50%
mature	0,0	
inonder	1,0	
localité	3,19	
soulager	4,13	
enthousiaste	4,23	
favori	4,45	
anxieux	4,97	
colis	6,10	
vigoureux	7,52	X
attait	8,32	X
embarras	8,39	
subitement	9,13	X
bougie	10,65	
manuels	12,94	X
familier	14,35	
trajet	14,48	
tissu	16,84	
velours	20,55	X
suprême	24,52	
plaine	28,61	

Afin de mieux comprendre ces chiffres, voici comment se répartissent les 130,000 items lexicaux de la base *Lexique* (fréquences d'occurrence par millions) :

Entre 0 et 1 de fréquence : 98,000 mots
 1 et 3 : 15,000 mots
 3 et 10 : 9,000 mots
 10 et 31 : 4,000 mots
 31 et 100 : 1,600 mots
 100 et plus : 700 mots

Nous pouvons constater que ce n'est pas seulement pour des items lexicaux moins fréquents que les sujets LNN ont éprouvé des difficultés. En effet, le taux de réussite a été assez faible pour les items «velours» et «manuel» qui, selon *Lexique*, sont plus fréquents que « favori », par exemple. En ce qui concerne l'item « mature », pour lequel *Lexique* ne donne aucune indication de fréquence, il s'agit pourtant d'un mot assez usité dans la langue française parlée et écrite au Québec. D'autre part, deux dictionnaires d'usage courant, le *Larousse* et le *Robert*, lui attribuent le sens d'une personne qui a atteint une maturité psychologique, sens qui était celui du contexte dans lequel le mot « mature » était inséré.

4.2.3 La tâche 3: connaissances morphologiques productives

La tâche 3 ciblait les connaissances en morphologie dérivationnelle. Dans cette épreuve de production contrôlée, les sujets devaient aussi compléter un mot, mais dans ce cas, il leur fallait trouver le suffixe approprié, la racine du mot ciblé étant donnée. Le contexte de la phrase fournissait des indices sur le sens du mot, mais aussi sur sa catégorie grammaticale. La tâche 3 comptait dix phrases.

Nous n'avons pas tenu compte des fautes d'orthographe dans la production des sujets et nous avons accepté les formes qui, au plan phonologique, correspondaient à l'item ciblé. À partir de la compilation de toutes les productions, nous avons établi une grille de codification. Nous la rappelons ici :

0= pas de réponse

1= erreur dans la suffixation

2= erreur de catégorie

3= erreur lexicale

5= bonne réponse

Pour la tâche 3, la presque totalité des participants ont obtenu de meilleurs résultats dans la version 1. Un seul sujet LN a mieux performé dans la version 2 et du côté des LNN, douze ont mieux réussi dans cette version. Ce petit nombre de LNN et surtout de LN ne nous offrant pas de possibilités sérieuses de comparaison entre les groupes linguistiques, nous ne présenterons donc que les résultats pour la version 1.

Nous présenterons tout d'abord au tableau 4.8 la moyenne des résultats en pourcentage selon le type de réponse produit pour chacun des groupes linguistiques.

Tableau 4.8

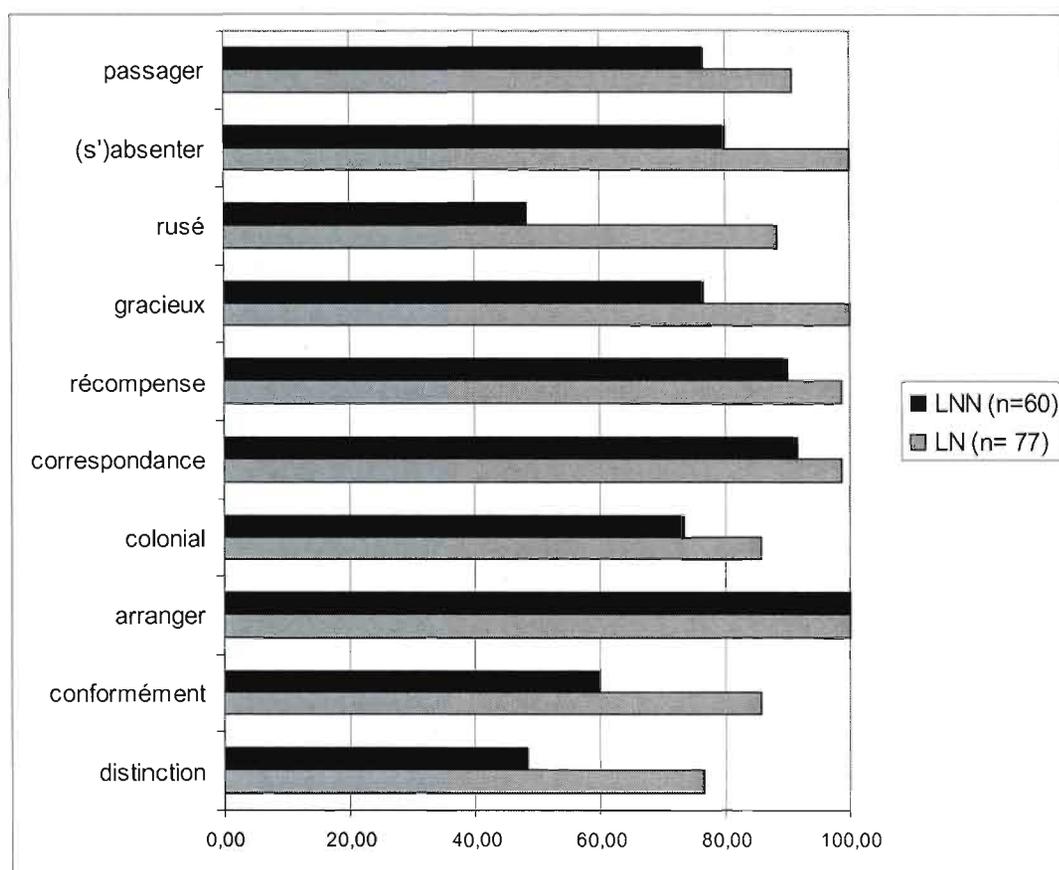
Tâche 3, version 1 : Comparaison des résultats selon le type de réponse

Réponse	LN n=77	LNN n=60
0	0,91	6,17
1	2,08	7,50
2	3,77	10,83
3	0,78	1,00
5	92,47	74,50

Nous pouvons observer dans les données de ce tableau que la proportion d'absence de réponse chez les LNN, quoique moins élevée que ce que nous avons vu pour la tâche 2, demeure supérieure à celle des LN. Il faut rappeler que, dans la tâche 3, une plus grande partie du mot ciblé était fournie aux sujets et qu'ils n'avaient à produire que le segment final, le suffixe. Les LNN ont aussi été plus nombreux que les LN à commettre des erreurs dans la formation des items: en additionnant les réponses de type 1 et 2, nous obtenons un taux moyen d'erreur dans la construction des dérivés de plus de 18%.

Nous passons maintenant aux performances des sujets pour chacun des items de la tâche 3. La figure 4.5 ci-dessous présente les taux de réussite (code « 5 ») par item pour chacun des groupes linguistiques.

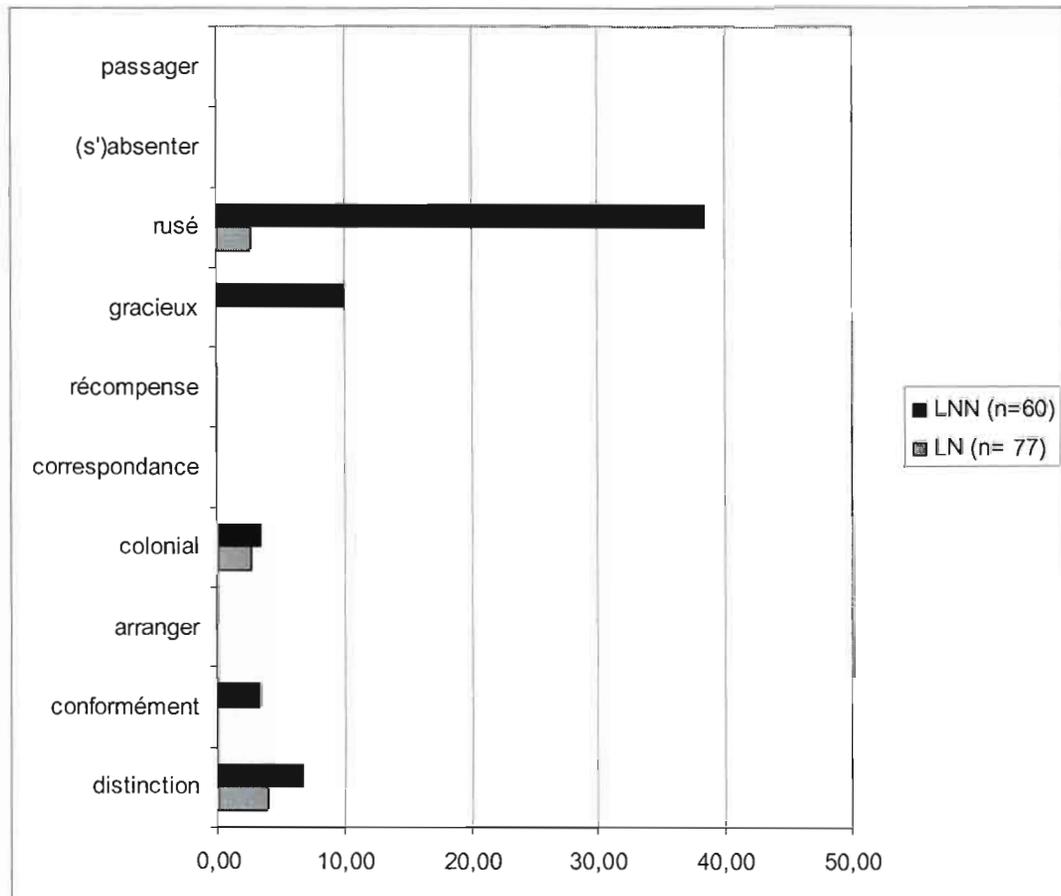
Figure 4.5
Tâche 3 version 1: comparaison du taux de réussite par item



Nous voyons que le taux de réussite est supérieur chez les LN pour tous les items, sauf pour «arranger». Le taux de réussite atteint ou dépasse 90% pour cinq des dix items chez les LN et trois de ceux-ci ont été réussis par 100% des sujets LN. À la figure 4.6 qui suit, nous comparons le taux d'absence de réponse (code « 0 ») des deux groupes linguistiques.

Figure 4.6

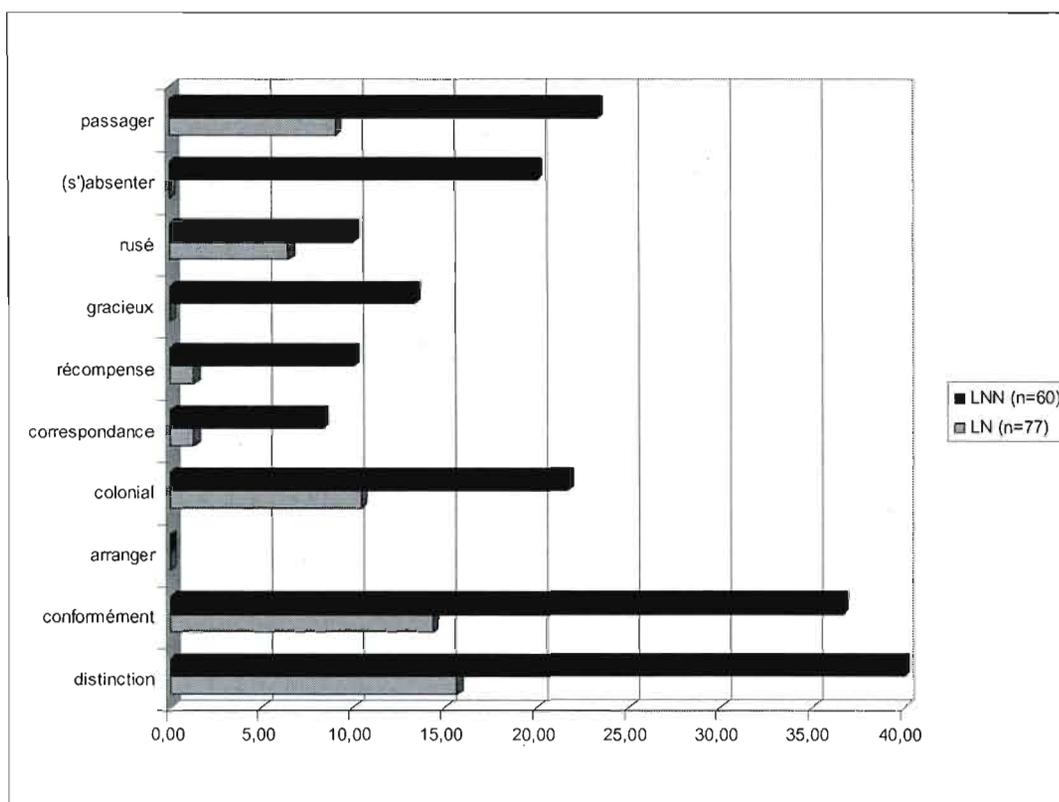
Tâche 3, version 1 : comparaison du taux d'absence de réponse par item



Les graphiques de la figure 4.6 illustrent que l'écart entre LN et LNN en ce qui a trait au taux d'absence de réponse est important dans le cas des adjectifs « rusé » et « gracieux ». On peut penser que, pour ces adjectifs, plusieurs sujets n'avaient aucune forme à proposer et qu'ils ignoraient peut-être tout à fait le mot ciblé. Cette hypothèse se trouve renforcée par le fait que le pourcentage de bonne réponse est assez bas : moins de 50% dans le cas de « rusé ».

Nous verrons maintenant dans quelle mesure les erreurs dans la formation des dérivés peuvent expliquer les performances pour chacun des items de la tâche. Nous avons considéré ensemble les erreurs de suffixation et de dérivation (codes « 1 » et « 2 ») et à la figure 4.7 ci-après, nous comparons les performances des LN et des LNN.

Figure 4.7
Tâche 3, version 1 : comparaison du taux d'erreur par item



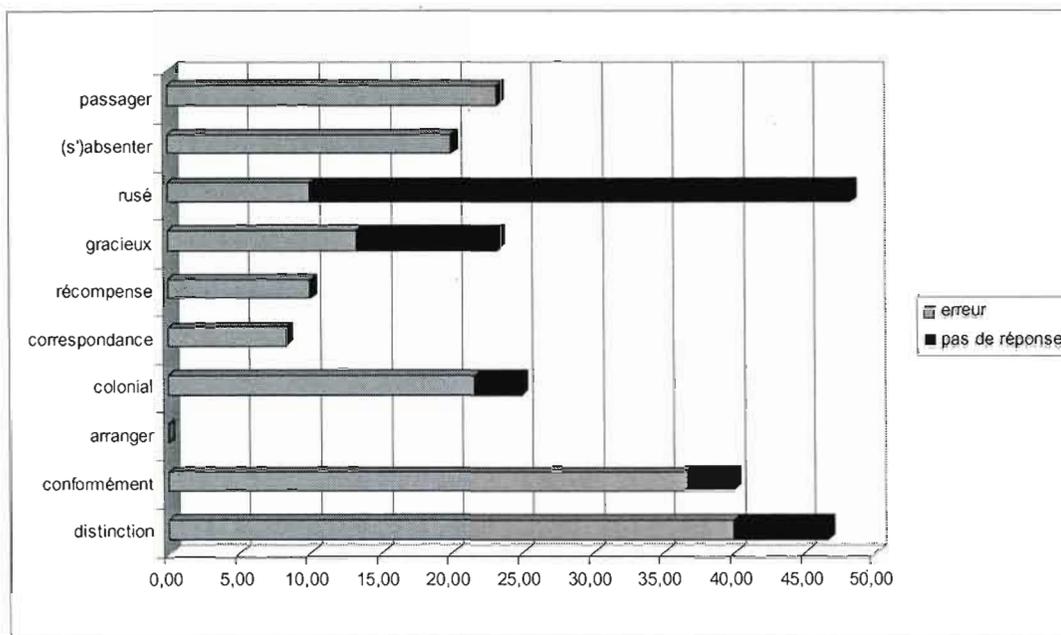
La figure 4.7 nous fait voir que les sujets LNN ont éprouvé beaucoup de difficultés avec certains dérivés. C'est le cas de « distinction », par exemple, pour lequel plusieurs des sujets ont produit l'adjectif, « distinct », plutôt que la forme nominale, alors qu'il s'agissait, dans la phase donnée, « de la plus haute distinction ». Dans le

cas de l'item « conformément » le taux d'erreur est surprenant : la construction des adverbes en français est assez simple et, dans la tâche, il s'agissait de l'expression courante « conformément à ... ». Plusieurs sujets ont produit la forme adjectivale, « conforme », alors que dans d'autres cas, le mot était formé avec un suffixe qui donnait un mot inexistant en français tel « *conformation ». On peut se demander si certains des sujets savaient qu'il fallait absolument un adverbe dans le contexte de la phrase. Ou encore, peut-être le savaient-ils, mais étaient incapables de le former. On peut aussi faire l'hypothèse que certains des LNN ne connaissaient pas l'expression, courante en français, « conformément à ... ».

L'explication pour la faible réussite des adjectifs « gracieux » et « rusé » est différente. Le pourcentage important de non réponse que nous avons vu à la figure 4.6 semble démontrer que plusieurs sujets ignoraient tout simplement ces mots. Quant à « passager », pour lequel le taux de non réponse a été faible, la difficulté rencontrée se trouverait plutôt être dans la construction de la forme nominale ou, sinon, de l'absence de connaissance de la forme « passager », car plusieurs des sujets ont produit la forme « passage » pour cet item.

Les données du tableau 4.8 indiquaient que les scores des LNN ont été, en moyenne, davantage affectés par des erreurs de catégorie ou dans la suffixation des items ciblés que par l'absence de réponse. Nous avons additionné le taux d'erreur dans la dérivation et nous l'avons comparé au taux de non réponse pour chacun des items. Les résultats sont présentés à la figure 4.8 qui suit.

Figure 4.8
Tâche 3: taux d'erreur et d'absence de réponse pour chaque item, LNN



Les données de la figure ci-dessus suggèrent une explication nuancée aux performances des LNN pour la tâche 3. Si, en moyenne, le taux d'absence de réponse est inférieur au taux d'erreur, cette affirmation n'est pas vraie pour tous les items de la tâche. Le faible taux de réussite pour les items « gracieux », mais surtout « rusé », s'expliquerait davantage par la proportion assez élevée d'absence de réponse. Par contre, il faut noter le pourcentage élevé d'erreurs dans la formation des items « conformément », « distinction » et « passager ».

Les résultats détaillés que nous venons de présenter semblent démontrer que les LNN ont de faibles connaissances en morphologie dérivationnelle. Leurs performances dans les tâches du test ont été inférieures à celles des LN. Nous ne pouvons cependant pas écarter l'hypothèse que leurs difficultés aient pour base des connaissances lexicales faibles ou insuffisantes. Nous avons donc soumis la liste

des items ciblés dans la tâche 3 à la base de données *Lexique* pour établir leur fréquence. C'est ce que nous présentons au tableau 4.9. Nous avons identifié à l'aide d'un X dans la colonne de droite les items pour lesquels le taux de réussite des sujets LNN était de 50% ou moins.

Tableau 4.9
Tâche 3 : fréquence des items lexicaux

Item	Fréquence	Moins de 50%
(s')absenter	0,0	
rusé	3,19	X
colonial	4,61	
passager	4,71	
gracieux	5,26	
récompense	7,32	
arranger	10,58	
conformément	11,35	
correspondance	19,74	
distinction	28,43	X

La fréquence moins grande de l'adjectif « rusé » pourrait expliquer la faible performance des LNN pour cet item, notamment, le taux d'absence de réponse. Le cas du nom « distinction » est cependant différent : alors qu'il est de plus haute fréquence, le taux d'erreur dans la construction du mot est tout de même élevé.

4.2.4 La tâche 4 : mots apparentés

Dans la tâche 4, qui ciblait les connaissances en morphologie dérivationnelle, les sujets devaient juger si des mots présentés en paires partageaient ou non une racine commune, d'où un sens en commun. Dix paires étaient composées de mots dérivés, comme « sembler/semblable », et cinq de mots qui n'avaient aucun lien de parenté sémantique, mais une ressemblance sur le plan de la forme, comme

« défi/définir ». Les sujets pouvaient donc identifier à tort une parenté sémantique ou, au contraire, ne pas reconnaître entre les deux items un lien de signification. Nous avons donc classifié le type d'erreur commis dans la tâche afin de vérifier s'il y avait des différences entre le comportement des sujets locuteurs natifs et les sujets locuteurs non natifs.

Un peu plus de 60% des sujets LN ont obtenu de meilleurs résultats à la version 1 de la tâche 4; cette proportion est d'environ 52% chez les LNN. Nous présentons au tableau 4.10 ci-dessous les erreurs produites en moyenne par chacun des deux groupes linguistiques pour chacune des versions.

Tableau 4.10
Tâche 4: comparaison des erreurs selon le type

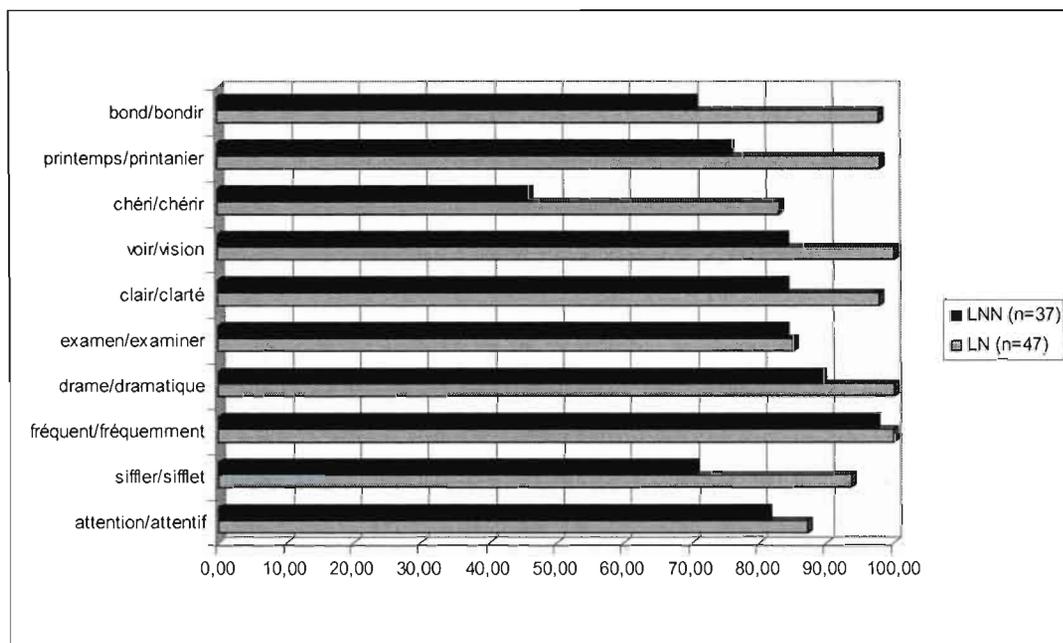
Version1	Vraies paires (n=10)		Pseudo-paires (n=5)	
	LN (n=47)	LNN (n= 37)	LN (n=47)	LNN (n= 37)
Moyenne	0,57	2,16	0,07	0,56
Écart t.	0,80	1,74	0,25	1,13
Version 2	LN (n=31)	LNN (n= 34)	LN (n=31)	LNN (n= 34)
Moyenne	1,03	2,44	0,25	0,53
Écart t.	1,06	2,40	0,51	0,79

Nous voyons que les LNN ont eu plus de difficulté que les LN à identifier la présence ou l'absence de liens sémantiques entre les paires de mots. Dans le cas des vraies paires de mots, les LNN n'ont pas reconnu, en moyenne, dans plus de deux des dix paires réelles, des liens sémantiques. La différence de performance entre les deux groupes linguistiques pour l'ensemble des résultats de la tâche 4, vérifiée à l'aide d'un ANOVA, est par ailleurs significative, tant pour les vraies paires ($F = 6,383$, $dl = 148$, $p < ,001$), que pour les pseudo-paires de mots ($F = 3,221$, $dl = 148$, $p = ,014$).

Nous verrons maintenant en détail les résultats pour l'ensemble des items de la tâche 4. Nous présentons à la figure 4.9 les performances détaillées des deux groupes pour les vraies paires de mots de la version 1.

Figure 4.9

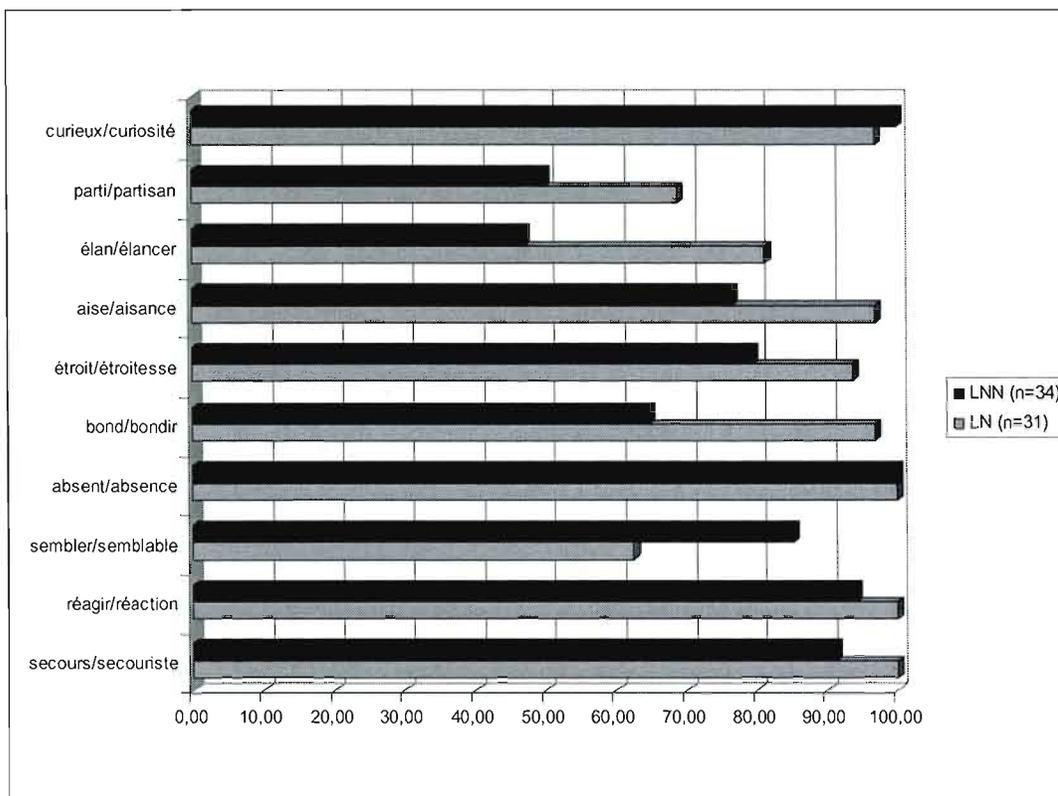
Tâche 4, version 1: comparaison du taux de réussite pour les vraies paires de mots



Voici maintenant, à la figure 4.10 le taux de réussite dans l'identification des vraies paires de mots de la version 2.

Figure 4.10

Tâche 4, version 2: comparaison du taux de réussite pour les vraies paires de mots



Dans la tâche 4, les LNN ont obtenu de meilleurs scores que les LN pour quelques items. Nous pouvons constater que le degré de transparence dans la dérivation ne semble pas avoir influencé les performances des LNN : ainsi, le taux de réussite pour les paires « printemps-printanier », « voir-vision », « clair-clarté », « fréquent-fréquemment », « réagir-réaction », « absent-absence » et « curieux-curiosité » est supérieur à celui des paires « chéri-chérir », « bond-bondir », « siffler-sifflet », « parti-partisan » et « élan-élancer » dont la dérivation est transparente. Il n'y a que la paire « drame-dramatique » qui échappe à cette constatation avec un pourcentage de réussite de plus de 90%. Il est à remarquer que la paire « bond-bondir » était présente dans les deux versions du test. Voilà pourquoi il y a deux résultats dans la figure ci-dessus.

Nous avons fait la recherche de fréquence des items lexicaux des vraies paires de mots de la tâche 4 dans la base de données de *Lexique*. Nous présentons les résultats au tableau 4.11 ci-dessous. Le X dans la colonne de droite indique les paires de mots pour lesquelles le taux de réussite a été de 50% et moins chez les sujets LNN.

Tableau 4.11
Tâche 4: fréquence des items des vraies paires de mots

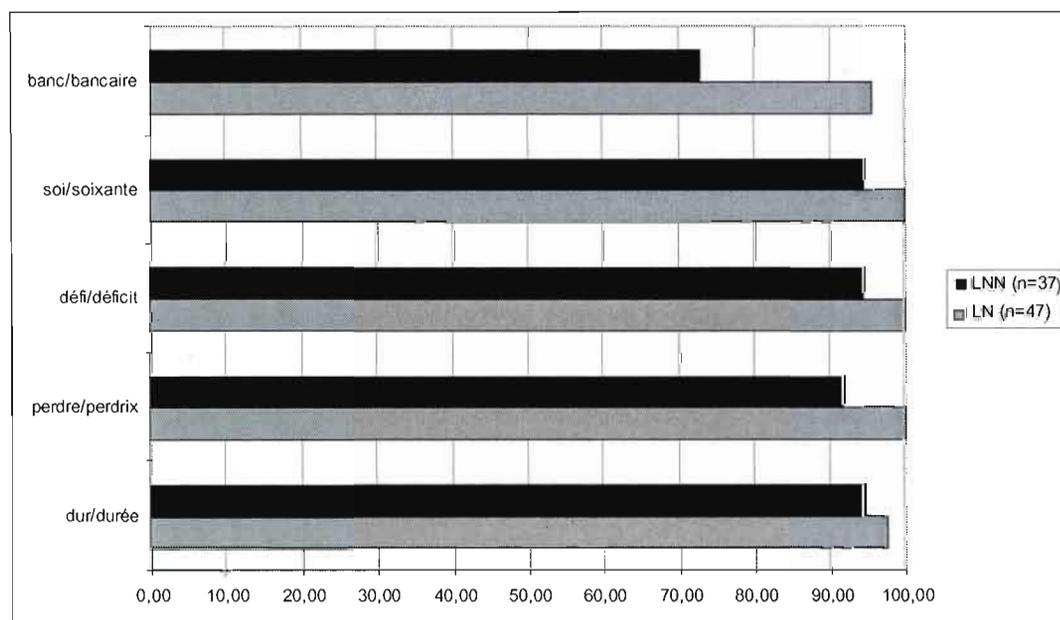
Item	Fréquence	Item	Fréquence	Moins de 50%
siffler	4,32	sifflet	7,13	
sembler	5,29	semblable	31,71	
réagir	7,77	réaction	43,94	
fréquent	10,00	fréquemment	25,61	
bond	13,58	bondir	5,29	
absent	13,77	absence	73,74	
chéri	14,48	chérir	1,00	X
élan	29,81	élancer	2,42	X
secours	31,35	secouriste	0,10	
aise	31,45	aisance	12,26	
drame	32,26	dramatique	16,16	
curieux	39,74	curiosité	36,26	
printemps	43,39	printanier	1,55	
examen	58,74	examiner	19,94	
clair	81,13	clarté	25,19	
attention	111,03	attentif	12,90	
parti	112,71	partisan	3,65	X
voir	479,94	vision	30,90	

Il est difficile de fournir une explication définitive pour la performance des LNN dans l'identification de la parenté sémantique des items de la tâche 4 à la lumière de la fréquence des items. En effet, on constate à la lecture du tableau ci-dessus que l'un et l'autre item d'une même paire peuvent avoir des niveaux de fréquence très différents. On peut cependant penser que dans les deux cas où le taux de réussite a été de moins de 50%, c'est peut-être la plus faible fréquence d'un des items qui

expliquerait la difficulté d'un bon nombre de LNN à en identifier le sens et à établir un lien de parenté.

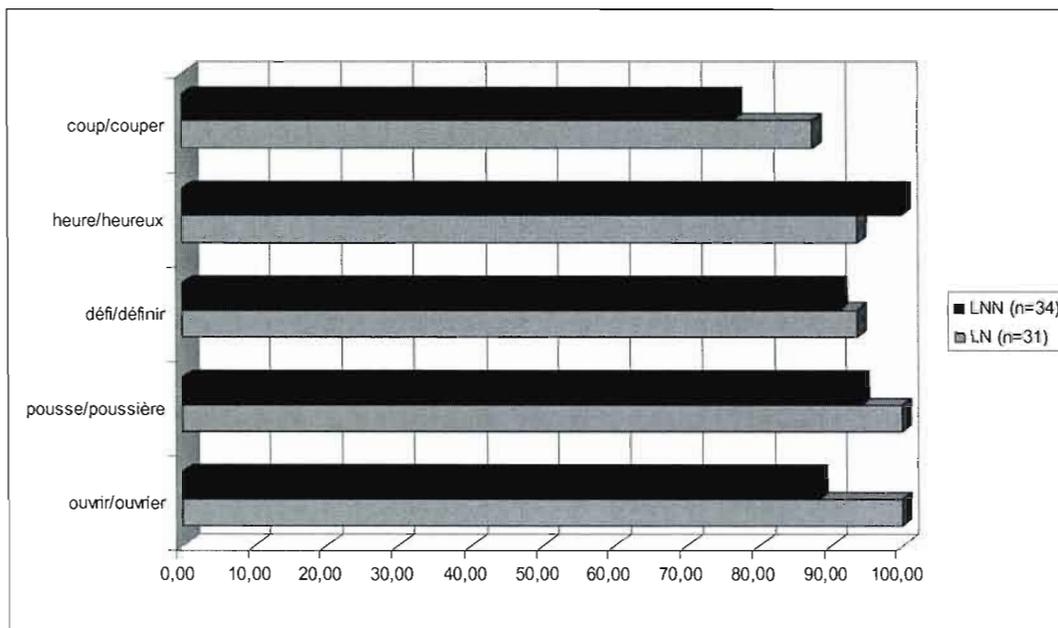
Nous passons maintenant aux performances des LN et des LNN pour les paires de mots n'ayant aucune parenté morphologique. Les données pour la version 1 sont présentées à la figure 4.11

Figure 4.11
Tâche 4, version 1: comparaison du taux de réussite
pour les pseudo-paires de mots



Les performances des deux groupes linguistiques à la version 2 de la tâche 4 sont présentées à la figure 4.12 ci-dessous.

Figure 4.12
Tâche 4, version 2: comparaison du taux de réussite pour les pseudo-paires de mots



Dans l'ensemble, les LNN ont été plus habiles à identifier les pseudo-paires de mots que les vraies paires de mots. Ils se sont toutefois quelque peu distingués des LN pour la paire « banc-bancaire », qui ne semble pourtant pas avoir causé de difficulté de jugement chez les francophones.

Le tableau 4.12 ci-dessous nous donne les informations sur la fréquence des items telle que nous l'avons trouvée dans la banque de données *Lexique*.

Tableau 4.12
Tâche 4 : fréquence des items des pseudo-paires de mots

Item	Fréquence	Item	Fréquence
défi	12,16	déficit	11,84
défi	12,16	définir	38,42
banc	25,71	bancaire	3,58
pousse	37,13	poussière	49,00
dur	56,84	durée	72,55
ouvrir	60,10	ouvrier	27,32
perdre	70,39	perdrix	1,52
soi	116,90	soixante	38,52
heure	305,61	heureux	83,42
coup	360,68	couper	21,68

De façon globale, les items des paires de mots n'ayant pas de parenté morphologique semblent être de plus grande fréquence que ce que nous avons vu pour les vraies paires de mots.

4.2.4 La tâche 5: production morphologique

La cinquième et dernière tâche de notre épreuve a été complétée avec beaucoup moins d'application que les autres. Nous en présentons tout de même les résultats que nous estimons toutefois moins fiables que ceux des quatre tâches précédentes.

Dans cette tâche, les sujets devaient produire des mots d'une même famille morphologique et les inscrire dans la bonne catégorie grammaticale dans une grille. Les sujets ne bénéficiaient ni de l'aide du contexte d'une phrase, ni d'une portion de mot, mais un mot amorce était déjà inscrit dans la grille à compléter. La consigne donnée pour réaliser la tâche ainsi que l'exemple fourni les informaient clairement que les mots demandés faisaient tous partie d'une même famille morphologique. Ici encore, nous n'avons pas considéré les erreurs d'orthographe. Après avoir compilé toutes les formes produites, nous avons consulté les dictionnaires Larousse et Petit

Robert pour juger de ce qui était acceptable. Nous avons accepté : industriel (nom); industriels (adjectif); doublé (nom et adjectif); libéral (nom); formalité (nom); formation (nom); format (nom); formaliser (verbe).

Dans la version 1 de la tâche 5, 35 LN et 21 LNN ont obtenu de meilleurs résultats, alors que dans la version 2, ce sont respectivement 43 LN et 50 LNN qui ont réalisé une meilleure performance. Les sujets devaient produire en tout, douze items lexicaux, soit trois noms, trois verbes, trois adjectifs et trois adverbes. Nous avons fait la moyenne du taux de bonne réponse et d'absence de réponse pour les deux versions du test. Nous présentons au tableau 4.13 ci-dessous, une comparaison des scores des deux groupes.

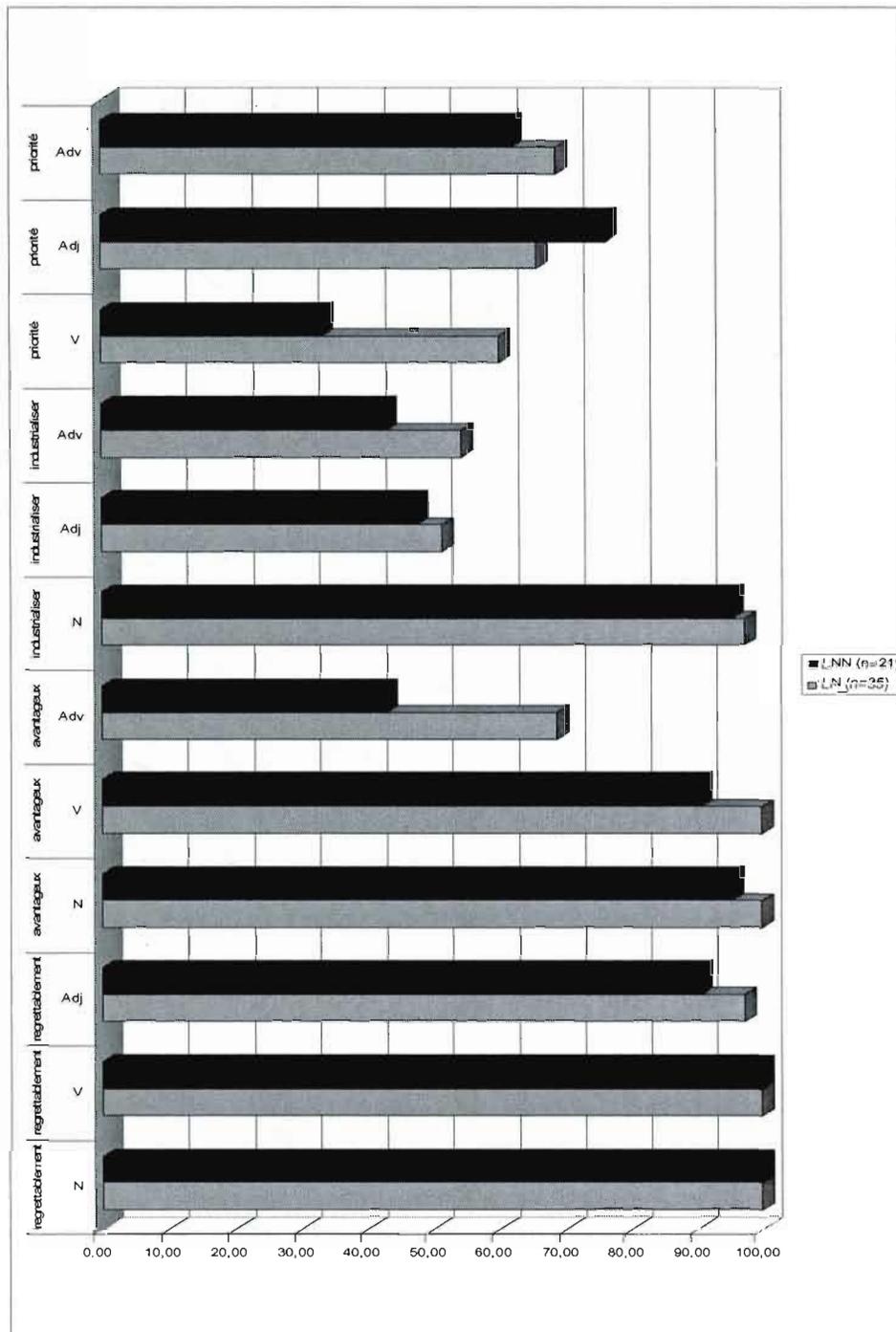
Tableau 4.13
Tâche 5: comparaison du taux de réussite et d'absence de réponse

	Version 1		Version 2	
	LN (n=35)	LNN (n=21)	LN (n=43)	LNN (n=50)
Bonne réponse	77,02	75,03	82,45	80,15
Pas de réponse	12,62	21,43	7,17	8,33

Les scores moyens de réussite des deux groupes sont comparables dans l'une et l'autre version de la tâche 5. Nous avons vu d'ailleurs que la différence entre les performances des deux groupes n'est pas significative. Le taux d'absence de réponse plus élevé chez les LNN, particulièrement à la version 12, suggère que ces derniers ont été plus fréquemment que les LN dans l'impossibilité de produire une réponse.

Nous comparons à la figure 4.13 ci-dessous le taux de réussite pour chacun des items de la version 1. Nous avons indiqué dans l'axe le mot amorce qui était fourni ainsi que la classe grammaticale demandée.

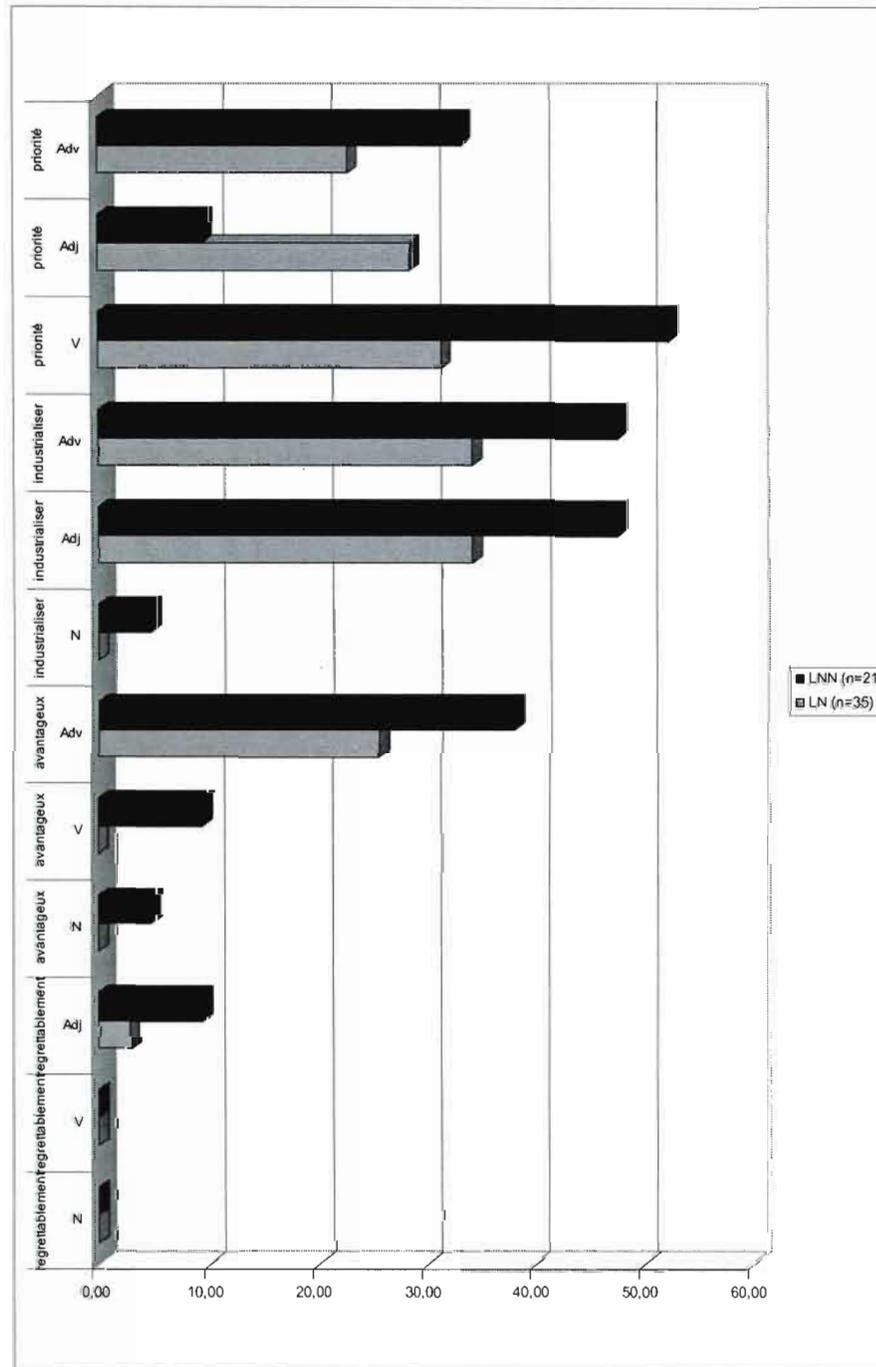
Figure 4.13
Tâche 5, version 1: Comparaison du taux de réussite par item



De façon générale, il n'y a pas de grandes différences entre les performances des deux groupes. Cependant, nous pouvons observer un taux de réussite assez faible pour l'adverbe « avantageusement » chez les LNN, ce qui est un peu surprenant puisque l'adjectif « avantageux », à partir duquel se construit l'adverbe, était fourni comme mot amorce. Le taux de réponse est également assez faible pour l'item « priorité » (forme verbale).

Nous avons vu qu'il était fréquent que les LNN ne produisent aucune réponse à la tâche 5. Dans la figure 4.14, nous comparons le taux d'absence de réponse des LNN et des LN pour chacun des items de la version 1.

Figure 4.14
Tâche 5, version 1: comparaison du taux d'absence de réponse par item

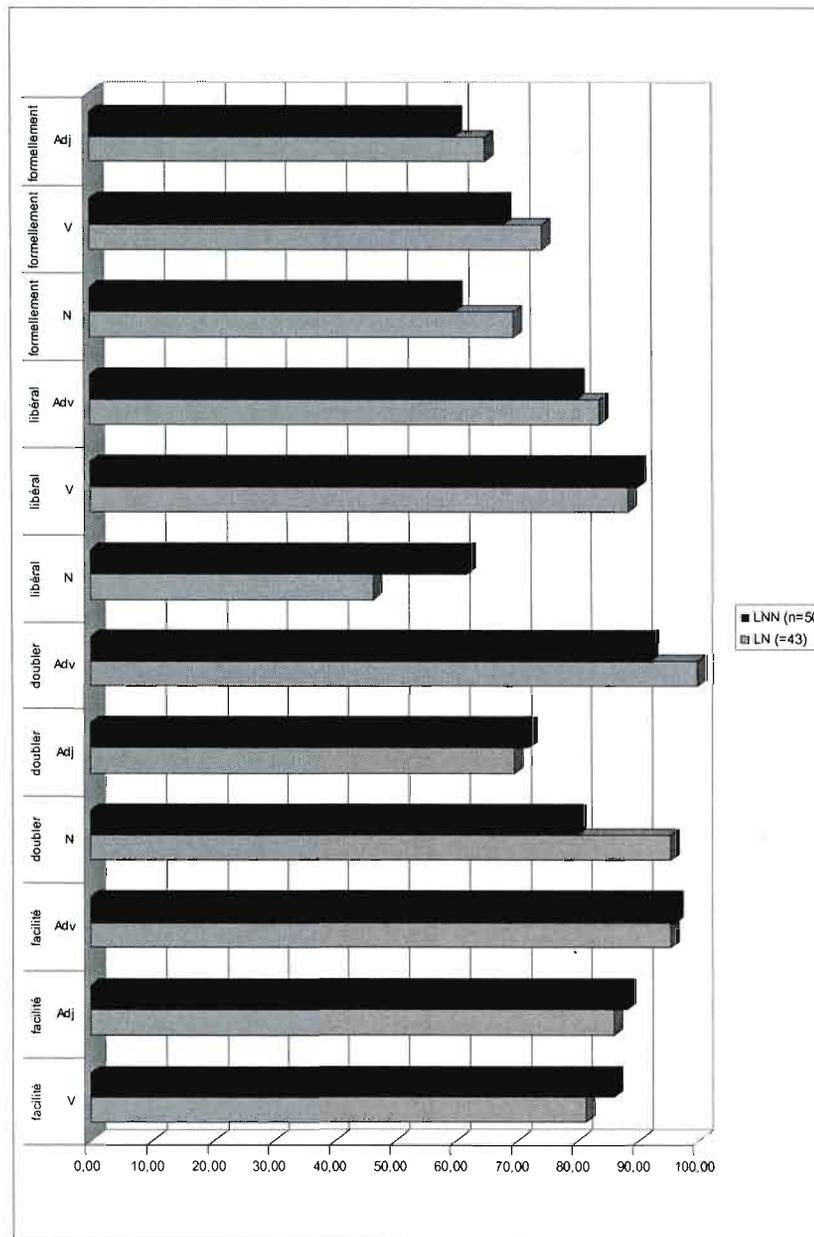


Les données de la figure ci-dessus nous laissent voir que, pour la majorité des items, les LNN ont eu un taux d'absence de réponse plus élevé que les LN. Le taux d'absence de réponse dépasse 40% pour trois des douze items qu'ils avaient à produire. Nous pouvons penser que l'absence de réponse pourrait expliquer, au moins en partie, le score obtenu par les LNN. Il est à remarquer que pour l'item «avantageusement», réussi par à peine 40% des sujets LNN, le taux d'absence de réponse est de plus de 30%.

Nous verrons maintenant les résultats pour la version 2 de la tâche. Nous présentons à la figure 4.15 le taux de réussite pour chaque item chez les deux groupes de sujets.

Figure 4.15

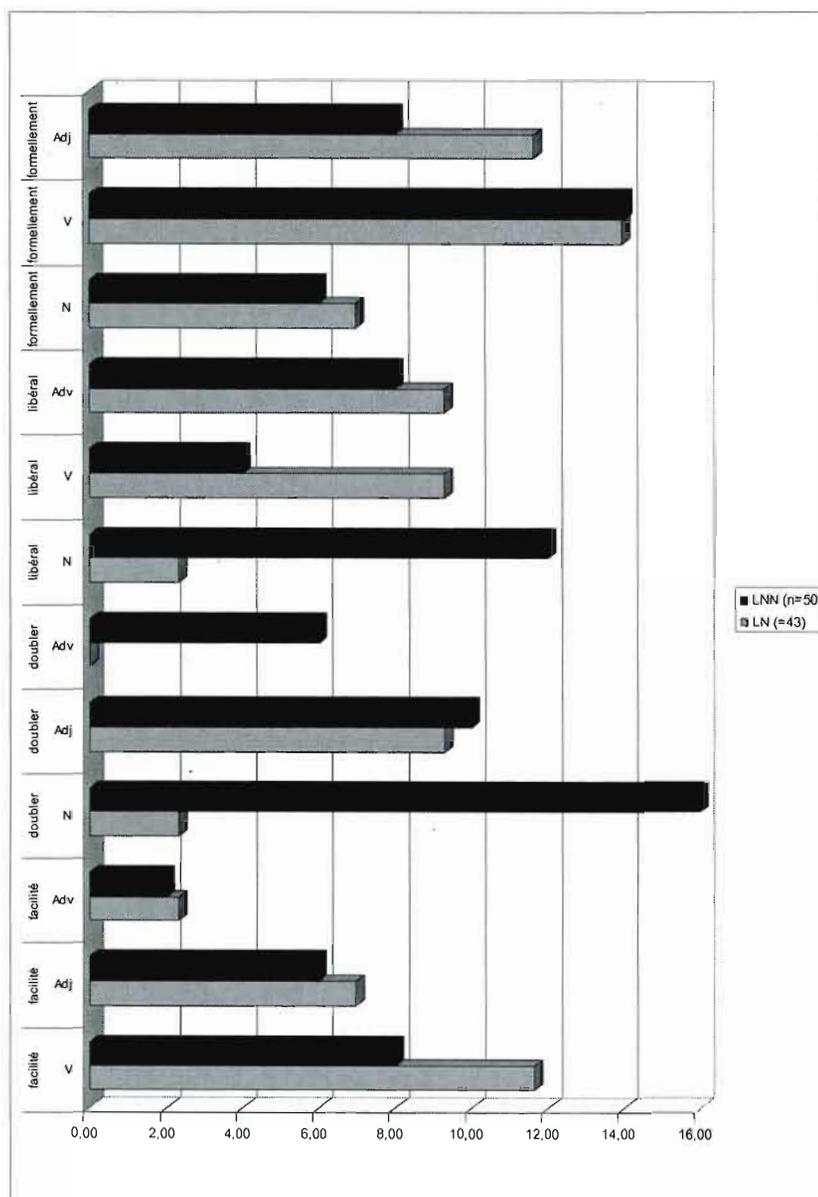
Tâche 5, version 2: comparaison du taux de réussite par item



La version 2 a été mieux réussie par l'ensemble des sujets et leurs performances sont, globalement, comparables. Nous présentons à la figure 4.16 ci-dessous, le taux d'absence de réponse pour chacun des items de la version 2.

Figure 4.16

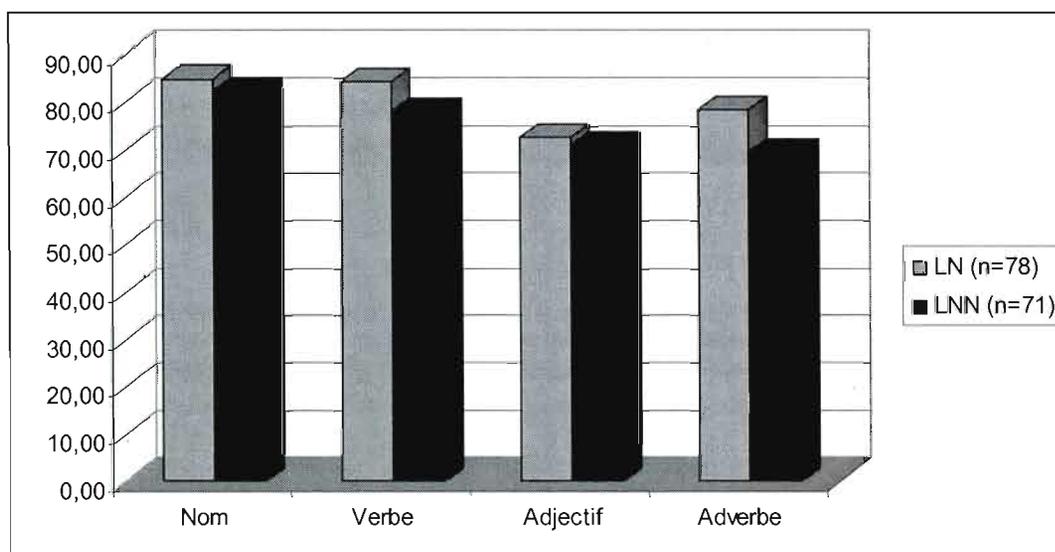
Tâche 5, version 2: comparaison du taux d'absence de réponse par item



Nous pouvons constater que le taux de non-réponse chez les LNN est comparable à celui des LN et qu'il est parfois même inférieur, sauf pour quatre items. Nous avons

comparé les performances des deux groupes de sujets selon la catégorie grammaticale des items à produire dans la tâche 5. Nous présentons à la figure 4.17 qui suit le taux moyen de réussite par catégorie grammaticale, pour l'ensemble des items.

Figure 4.17
Tâche 5: comparaison du taux de réussite par catégorie grammaticale

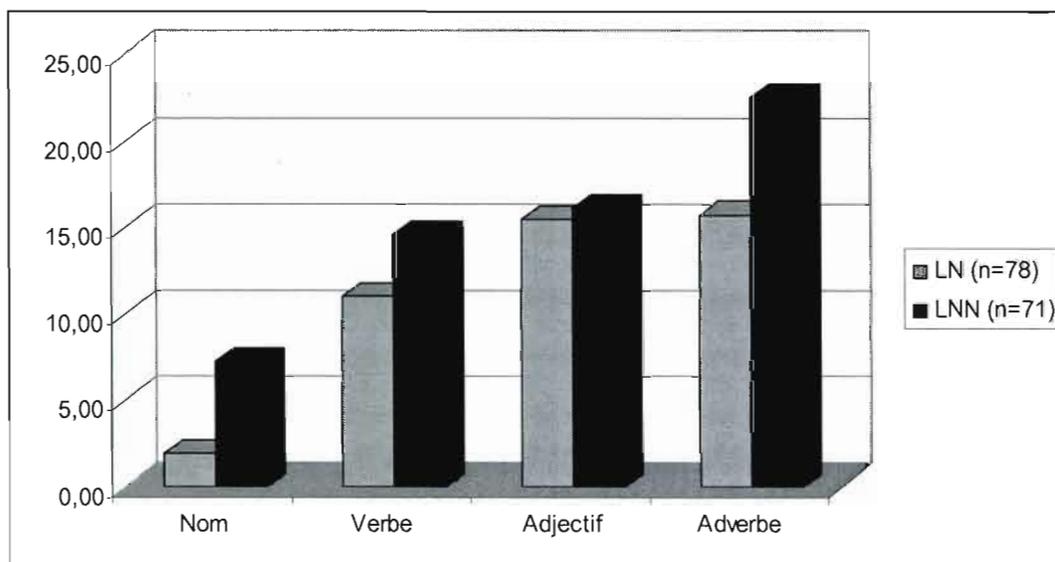


Le taux de réussite suit sensiblement la même courbe chez les deux groupes linguistiques avec une différence à l'avantage des LN pour toutes les catégories grammaticales. Il faut cependant remarquer que l'écart le plus grand (près de 10%) entre les deux groupes se situe dans la production des adverbes.

Nous comparerons, à la figure 4.18, le taux d'absence de réponse pour chacune des catégories grammaticales dans le but d'évaluer dans quelle proportion cela a pu influencer la performance des sujets.

Figure 4.18

Tâche 5: comparaison du taux d'absence de réponse par catégorie grammaticale



La figure ci-dessus nous montre que chez les LNN le taux d'absence de réponse est supérieur dans toutes les catégories grammaticales. Le taux le plus élevé (près de 20%) a été pour les adverbes, catégorie pour laquelle leur score (moins de 60%) avait été plus bas que les LN.

Rappelons que nous avons établi une codification pour les erreurs produites par les sujets dans la tâche 5:

- 2= mot inexistant
- 3= erreur de catégorie
- 4= erreur lexicale

Le nombre d'erreurs lexicales a été dans l'ensemble peu important. Notre objectif dans la tâche 5 était d'évaluer les connaissances en morphologie dérivationnelle. Nous avons donc totalisé les erreurs de type 2 et 3 et nous avons calculé la

moyenne par groupe linguistique pour chacune des versions. Ces résultats en pourcentage sont présentés au tableau 4.14.

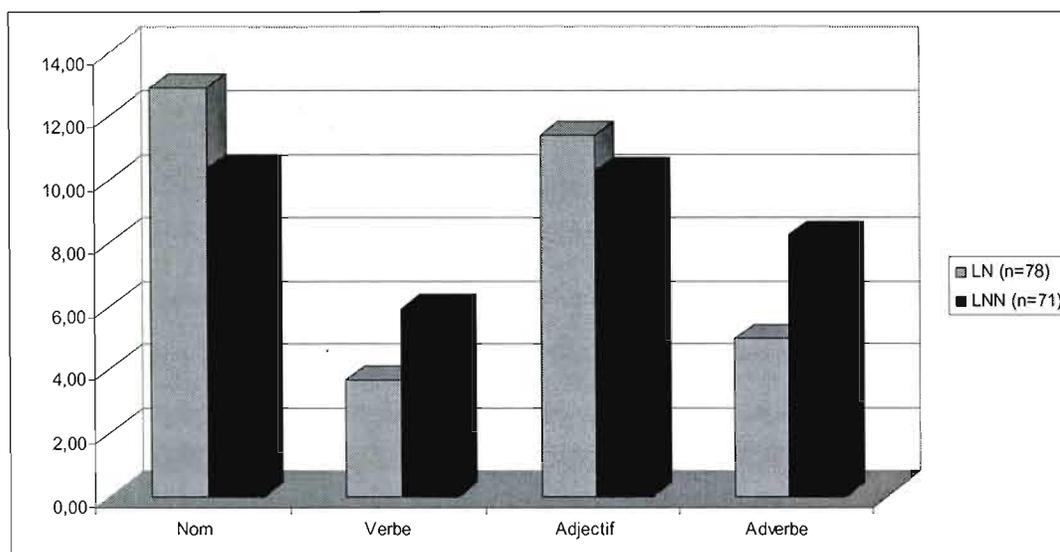
Tableau 4.14
Tâche 5: Comparaison du taux d'erreur

Version 1		Version 2	
LN (n=35) 5,47	LNN (n=21) 4,37	LN (n=43) 12,40	LNN (n=50) 15,00

À la lecture du tableau qui précède, nous voyons que le comportement des deux groupes linguistiques dans la production d'erreurs est assez comparable, ce qui suggère que les sujets ont éprouvé les mêmes difficultés dans la réalisation de la tâche 5.

Afin d'identifier ce qui pouvait avoir posé plus de difficultés chez les deux groupes de sujets, nous avons fait la moyenne des erreurs pour chaque catégorie grammaticale des items demandés. À la figure ci-dessous, nous comparons les performances des deux groupes.

Figure 4.19
Tâche 5: comparaison du taux d'erreur par catégorie grammaticale



Nous constatons à l'examen des graphiques de la figure ci-dessus que bien qu'il y ait des différences entre les deux groupes, et parfois, à l'avantage des LNN, ces différences ne sont que de quelques points de pourcentage. C'est dans la production des adverbes que les LNN se distinguent le plus des LN.

Les données sur le taux d'erreur et d'absence de réponse suggèrent que ce qui les distingue des sujets francophones, c'est davantage le taux élevé d'absence de réponse. Ce taux d'absence de réponse serait responsable de leurs faibles performances particulièrement dans la production des formes adverbiales et adjectivales.

Nous avons vérifié le niveau de fréquence des items lexicaux de la tâche 5 (incluant les formes acceptées) afin de voir si cette variable avait pu jouer un rôle dans les résultats obtenus par les LNN. Les données sont présentées au tableau 4.15 ci-

dessous sont tirées des bases *Lexique*. Un X dans la colonne de droite identifie les items pour lesquels le taux de réussite chez les LNN a été de 50% et moins.

Tableau 4.15
Tâche 5: fréquence des items lexicaux

Item	Fréquence		Item	Fréquence	
prioritairement	0,0		libéral	6,74	
prioriser	0,0	X	regretter	8,52	
avantager	0,16		faciliter	12,74	
formaliser	0,65		libérer	12,74	
libéralement	0,77		regret	18,45	
industriel	0,9		industriel	22,03	X
prioritaire	1,1		former	30,68	
industriellement	1,48	X	avantage	34,23	
avantageusement	1,9	X	facilement	42,35	
doublure	2,13		facile	74,06	
formalité	2,45		industrie	82,71	
regrettable	3,77		double	88,16	
doublement	4,58		formation	90,58	X
formel	5,65	X	liberté	115,26	
format	5,97	X	forme	282,58	X

Le verbe «prioriser», pour lequel *Lexique* ne nous donne pas d'indication de fréquence est cependant répertorié dans le *Grand Dictionnaire terminologique* de l'*Office québécois de la Langue française*. Selon cette source, il s'agit d'un emploi très répandu pour signifier «accorder la priorité à ...» et qui est tout à fait conforme à la langue française. Nous avons consulté la base textuelle de *Lexiquum* et nous avons trouvé des occurrences tirées du quotidien *La Presse*, comme par exemple, « une politique qui devra prioriser les cadres du parti. ».

En ce qui concerne l'adverbe «prioritairement», lui aussi répertorié dans le *Grand Dictionnaire Terminologique*, la base textuelle de *Lexiquum* nous fournit des

contextes d'occurrence tirés du quotidien *La Presse*, comme par exemple, « ... une politique d'habitation centrée prioritairement sur les besoins des démunis ».

Nous avons vu que les LNN ont été faibles dans la production des adverbes. Nous pouvons supposer que leur faible score pour l'adverbe « industriellement », pourrait être dû à leur mauvaise performance pour l'adjectif « industriel » qui, pourtant, est plus fréquent. Les LNN ont obtenu un score très comparable aux LN pour la production de l'adverbe « facilement » qui lui, est un item plus fréquent que les autres adverbes de la tâche 5.

4.2.6 Résultats selon le type de tâche

Toutes les données présentées jusqu'ici indiquent que les LNN sont faibles sur le plan des connaissances lexicales et en morphologie dérivationnelle. Leurs résultats sont, pour toutes les tâches, inférieurs à ceux des LN qui sont eux-mêmes des étudiants faibles en français. Nous avons vu que la différence entre la performance moyenne des LN et des LNN était significative pour toutes les tâches, sauf pour la tâche 5. L'épreuve que nous avons utilisée dans cette recherche comportait des mesures des connaissances réceptives (tâches 1 et 4) et productives (tâches 2, 3 et 5). Bien que ces distinctions dans la nature des connaissances ne soient pas tout à fait dichotomiques et qu'il faille situer plutôt les connaissances lexicales sur un continuum, il demeure que la production demande une plus grande maîtrise et qu'elle fait appel au savoir procédural (Segalowitz, 1991; Weber, 1991). Dans le tableau 4.16 qui suit, nous avons fait un sommaire de l'ensemble des résultats de l'épreuve en distinguant les tâches de nature réceptive et productive.

Tableau 4.16
 Comparaison des résultats selon le type de tâche

	LN (n=78)		LNN (n=71)	
	M	é. t.	M	é. t.
Tâche 1	88,99	8,16	80,93	14,09
Tâche 4	95,47	5,93	87,04	8,94
Moyenne tâches réceptives	92,23	7,045	83,99	11,52
Tâche 2	85,51	11,47	60,15	23,75
Tâche 3	92,31	8,82	74,51	18,11
Tâche 5	79,74	14,68	77,58	14,65
Moyenne tâches productives	85,85	11,66	70,75	18,84

Nous pouvons observer que les résultats sont supérieurs chez tous les sujets dans les tâches réceptives. Les différences de résultats entre LN et LNN sont significatives tant pour les tâches réceptives ($F= 59,164$, $df = 148$, $p <,001$) que pour les tâches productives ($F= 43,751$; $df= 148$, $p <,001$).

Nous verrons maintenant quelles sont les différences dans les performances des sujets selon qu'il s'agit d'une tâche ayant pour cible le savoir lexical (tâches 1 et 2) ou visant les connaissances en morphologie (tâches 3, 4 et 5). Dans le tableau 4.17 ci-dessous, nous avons regroupé les résultats par type de tâche, lexicale ou morphologique. Les écarts types (é.t.) sont inscrits dans la colonne de droite.

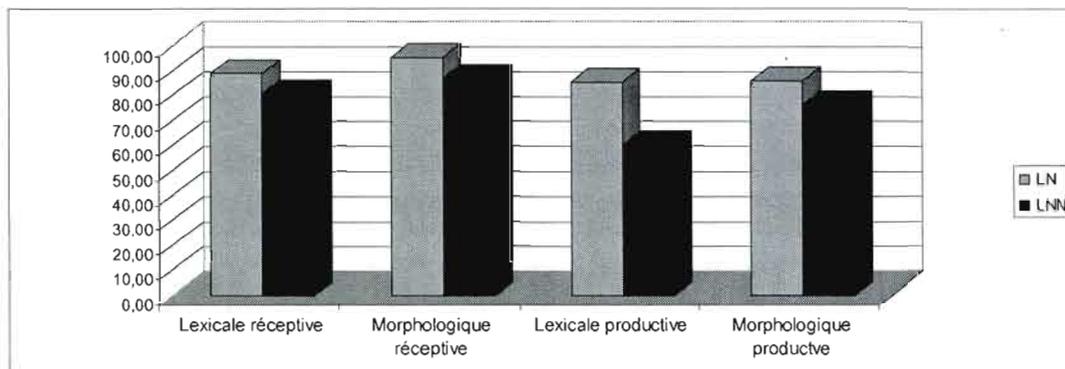
Tableau. 4.17
 Comparaison des résultats selon les connaissances ciblées

	LN (n=78)		LNN (n=71)	
	M	é. t.	M	é. t.
Tâche 1	88,99	8,16	80,93	14,09
Tâche 2	85,51	11,47	60,15	23,75
Moyenne t. lexicales	87,25	9,82	70,54	18,92
Tâche 3	92,31	8,82	74,51	18,11
Tâche 4	95,47	5,93	87,04	8,94
Tâche 5	79,74	14,68	77,58	14,65
Moyenne t. morphologiques	89,17	9,81	79,71	13,90

Nous voyons qu'il y a une différence entre LN et LNN dans les deux types de tâche. Cet écart est significatif pour les tâches lexicales ($F= 87,249$, $dl = 148$, $p <,001$) et celles visant les connaissances en morphologie ($F=25,873$, $dl = 148$, $p <,001$).

La figure 4.20 ci-dessous présente un portrait comparé des performances des deux groupes linguistiques selon le type de tâche et les connaissances ciblées.

Figure 4.20
 Comparaison des performances selon le type de tâche



Les graphiques de la figure 4.20 font ressortir les écarts de performance entre les deux groupes et nous pouvons voir que c'est pour la tâche lexicale productive que les LNN se distinguent le plus des LN.

Afin de vérifier s'il existait des corrélations entre les tâches du test qui visaient le même type de connaissance, nous avons appliqué un test de Pearson dont les résultats sont présentés dans le tableau 4.18.

Tableau 4.18
Corrélations entre les tâches, LN et LNN

Tâches	LN n=78	LNN n=71
Tâches 1 et 2	$r = ,352, p = ,002$	$r = -,091, p = ,448$
Tâches 2 et 3	$r = ,403, p <,001$	$r = ,716, p <,001$
Tâches 3 et 4	$r = ,136, p = ,234$	$r = ,562, p <,001$

Pour les LN, nous avons trouvé une corrélation moyenne entre les tâches lexicales 1 et 2 et entre les tâches de production 2 et 3, mais pas de corrélation entre les tâches 3 et 4 (morphologie). Chez les LNN, nous avons obtenu une corrélation forte entre les tâches 3 et 4 visant les connaissances en morphologie dérivationnelle et entre les tâches de production 2 et 3, mais pas de corrélation entre les tâches 1 et 2 (lexique). Nous n'avons pas trouvé de corrélations significatives entre les tâches 3 et 5 (morphologie), 4 et 5 (morphologie), entre les tâches 1 et 4 (connaissances réceptives) ni entre les tâches 2 et 5 (production) chez l'un et l'autre groupe.

4.3 Les résultats au cours d'appoint en grammaire

Rappelons que nous voulions vérifier l'existence d'une relation entre les performances des sujets à l'épreuve utilisée dans cette recherche et la note finale du

cours, laquelle est un indicateur du niveau de maîtrise du code grammatical en français écrit.

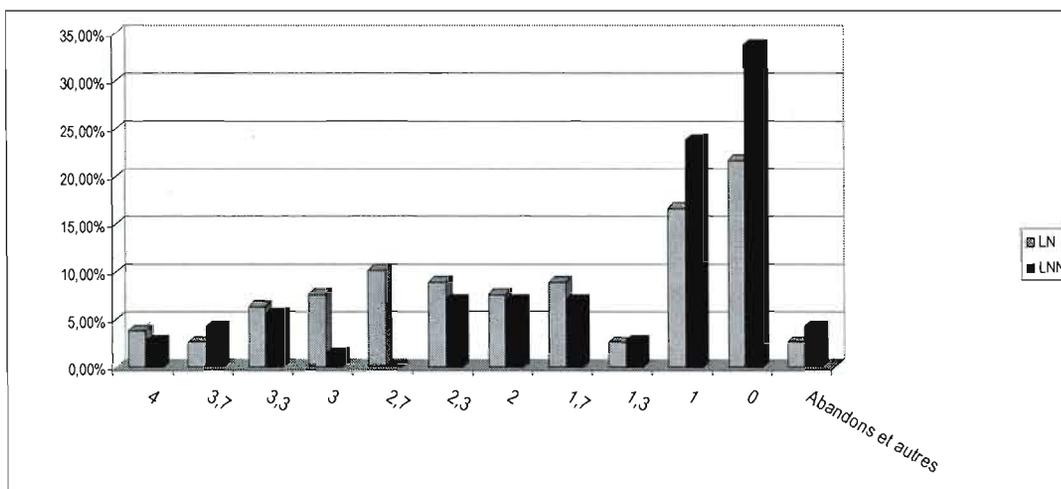
Tout d'abord, le tableau 4.19 présente le barème de notation utilisé dans le cours et ses équivalences en note numérique et exprimée en pourcentage. La notation alphabétique (de A+ à E) correspond à une note numérique (de 4,3 à 0), laquelle traduit une échelle de points en pourcentage. Par exemple, la note B+, correspond à la note numérique 3,3, soit un score se situant entre 78 et 81%. Une note inférieure à 55% entraîne un échec.

Tableau 4.19
Barème de notation du cours de grammaire

Alpha	Numérique	%
A+	4,3	93 et +
A	4	85 à 93
A-	3,7	81 à 85
B+	3,3	78 à 81
B	3	75 à 78
B-	2,7	71 à 75
C+	2,3	68 à 71
C	2	65 à 68
C-	1,7	61 à 65
D+	1,3	58 à 61
D	1	55 à 58
E (échec)	0	moins de 55

À la figure 4.21 ci-dessous, nous comparons les résultats obtenus au cours d'appoint en grammaire par les LN et les LNN.

Figure 4.21
 Comparaison des résultats au cours de grammaire



Nous pouvons voir dans ces données que le taux d'échec est assez élevé : plus de 20% chez les LN, près de 35% chez les LNN. Ces résultats ne sont pas exceptionnels : en effet, selon les données colligées par le Groupe de réflexion sur l'enseignement de la grammaire (GREG) de l'UQAM et présentées dans son rapport en 2005, le taux de réussite au cours d'appoint en grammaire a varié d'un peu plus de 54% à 60% entre les années académiques 1999-2000 à 2003-2004. Bien que quelques LNN aient obtenu une note entre 3,3 et 4, de façon générale, leurs résultats sont plus faibles que ceux des francophones.

Nous trouvons dans le tableau 4.20 ci-après, la moyenne des résultats obtenus par l'ensemble de nos sujets au cours d'appoint en grammaire. Mentionnons que nous n'avons pas de données pour trois des sujets, car leur note finale n'était pas enregistrée au moment de l'analyse.

Tableau 4.20
Comparaison de la moyenne des résultats au cours de grammaire

	LN n=77	LNN n=69
Moyenne	1,70	1,23
Écart type	1,24	1,23

Les LNN ont obtenu une moyenne (1,23) correspondant à une note en pourcentage de moins de 61%. Quant aux LN, leur moyenne numérique (1,7) donne une note en pourcentage se situant entre 61 et 65%. Un test d'analyse de variance a établi que la différence entre les résultats au cours des LN et des LNN est significative ($F=5,334$, $df = 145$, $p = ,022$).

Nous avons vérifié l'existence de corrélations entre les meilleurs résultats obtenus à notre épreuve et la note obtenue pour le cours en calculant les r de Pearson. Les résultats sont différents selon le groupe linguistique. En effet, chez les LN, il y a une corrélation faible, mais significative, entre les résultats pour la tâche 2 ($r = ,244$, $p = ,033$) et une corrélation un peu plus forte pour la tâche 3 ($r = ,316$, $p = ,005$). Nous n'avons trouvé aucune corrélation entre les tâches de notre épreuve et la note finale au cours d'appoint en grammaire pour les LNN.

Nous avons présenté et comparé les résultats des sujets LN et LNN de façon globale et de façon détaillée pour l'ensemble des tâches de notre épreuve. Nous avons tenté d'identifier ce qui pouvait expliquer les performances plus faibles des LNN dans toutes les tâches et de situer où se trouvaient possiblement leurs difficultés.

Dans le chapitre suivant, à partir d'une synthèse des résultats obtenus par les sujets LNN, nous les discuterons en faisant référence à notre cadre conceptuel. Nous établirons aussi dans quelle mesure nos hypothèses de départ ont été ou non validées.

CHAPITRE V

DISCUSSION

Deux questions principales ont été à la base de notre travail: nous voulions savoir quelles étaient les connaissances lexicales et en morphologie dérivationnelle chez des adultes non francophones aspirant entreprendre des études supérieures dans une L2. Plus particulièrement, nous avons comparé les connaissances de LNN relativement avancés en L2 à celles de LN faibles en français écrit, population qui les côtoie dans les cours d'appoint en grammaire. Nous avons tenté d'évaluer l'étendue de leurs connaissances d'items lexicaux fréquents à travers des tâches de compréhension et de production.

Nous avons mesuré les connaissances de nos sujets à deux reprises à l'aide d'un même type de test, en deux versions. Nous avons comme hypothèse de départ qu'il n'y aurait pas d'amélioration significative des connaissances lexicales et en morphologie dérivationnelle entre la première passation du test (2^e semaine du cours d'appoint en grammaire du français écrit) et la deuxième (14^e semaine du cours) chez les sujets LNN. D'après la recherche sur l'acquisition lexicale chez les apprenants avancés en L2, poursuivant des études universitaires et très à l'aise en français, ces étudiants auraient déjà atteint une certaine stabilité lexicale. Les résultats que nous avons obtenus viennent appuyer cette hypothèse. Bien que nous ne soyons pas en mesure d'en identifier les causes, cette stabilité existe. Cette

constatation est importante, car elle nous indique que les différences que nous avons observées entre les populations LN et LNN risquent elles aussi d'être stables et de persister à long terme et ainsi avoir des conséquences tout au long du parcours universitaire des étudiants que nous avons évalués.

Dans toutes les tâches de l'épreuve que nous avons utilisée, les LNN ont réalisé des scores inférieurs à ceux des LN. S'agissant de LNN apprenants assez avancés en L2, nous escomptions des différences moins importantes entre leurs performances et celles des LN (Meara, 1984; Read, 1993), particulièrement parce qu'il s'agissait de LN faibles en français écrit. Dans la tâche de reconnaissance lexicale, les LNN ont tout juste atteint le seuil de 80%, alors que les LN ont obtenu un score moyen de 89%; en production lexicale, les LNN ont à peine réussi un score moyen de 60% alors que chez les LN, il a été de 85,51%; les LNN ont un peu mieux performé en production morphologique avec un score moyen de 74,51%, mais ils sont loin derrière les LN qui ont réussi la tâche à plus de 92%; enfin, si dans la tâche d'identification de parenté morphologique les LNN ont obtenu un score moyen de 87%, les LN les surpassent tout de même avec un taux de réussite de plus de 95%. Sur le plan statistique, les différences entre les performances des deux groupes linguistiques sont significatives, sauf pour la tâche 5.

Le score obtenu à la tâche 1 (reconnaissance lexicale) est un indicateur de l'étendue du vocabulaire réceptif des LNN. Meara (1992) estime que les 2,600 mots sur lesquels sont basés ses tests représentent environ 80% de ce dont un locuteur a besoin dans une conversation de la vie courante. La moyenne des meilleurs résultats à la tâche 1 pour les LNN est d'un peu plus de 80%, mais il y a cependant un écart type important (14,09), ce qui laisse supposer que si certains LNN ont relativement bien performé, d'autres pourraient avoir des connaissances lexicales assez faibles. Nous savons qu'il est possible de reconnaître un mot, mais d'avoir une faible compréhension, alors que l'inverse ne se produit pas (Stanovitch, 1991).

Nous avons vu que la différence de score entre les deux groupes linguistiques était surtout attribuable au taux d'erreur pour les pseudo-mots chez les LNN. Le taux d'erreur pour les pseudo-mots indique dans quelle proportion les sujets ont identifié ces formes inventées comme étant des mots réels. Il est possible que les sujets LNN, apprenants assez avancés en L2, aient de la difficulté à identifier ce qu'ils ne connaissent pas bien (Laufer et Yano, 2001) ou encore que leurs connaissances sur soient incomplètes ou manquent de solidité. Comparativement aux LN qui n'ont pas eu de difficulté à rejeter les mots inventés, les LNN semblent moins confiants dans leurs connaissances du lexique de la L2.

Les items lexicaux ciblés dans la tâche 2 étaient des items fréquents qui devraient faire partie du vocabulaire de la L2 connu et maîtrisé chez des étudiants universitaires. Les contextes phrastiques dans lesquels ils étaient enchâssés avaient été vérifiés à l'aide de concordanciers pour s'assurer de leur fréquence. Or, pour tous les items de cette tâche, les LNN ont obtenu des résultats nettement plus faibles que les LN. De plus, comparativement aux LN, ils n'ont produit aucune réponse dans beaucoup de cas et c'est d'ailleurs ce qui expliquerait en grande partie leur faible score. Les tableaux et graphiques que nous avons présentés pour la tâche 2 démontrent que l'écart d'un peu plus de 25% entre les performances des deux groupes linguistiques serait dû en grande partie à l'incapacité des LNN à fournir une réponse.

Dans la tâche productive ciblant les connaissances en morphologie dérivationnelle (tâche 3), les LNN ont produit un taux assez élevé d'absence de réponse tout comme ils ont commis beaucoup d'erreurs dans la formation des dérivés comparativement aux LN. Les LNN semblent avoir de la difficulté à manipuler les structures des mots, ce qui fait partie des compétences en morphologie dérivationnelle (Carlisle, 1995). La maîtrise des règles de la morphologie dérivationnelle peut être une difficulté pour des apprenants de L2: il y a des affixes

qui ont plusieurs sens et un même sens peut se traduire par des affixes différents (Bogaards, 1994). Il demeure cependant qu'il y a beaucoup de régularités en morphologie dérivationnelle et les LNN de notre recherche ne semblent pas avoir été en mesure d'en tirer profit ou encore, peut-être n'étaient-ils pas sensibles au caractère systématique de la morphologie dérivationnelle en français.

Les erreurs commises dans la production des items de la tâche 3 suggèrent que beaucoup de LNN ne se sont pas servis des indices syntaxiques de la phrase pour déterminer l'item ciblé, ce qui serait un indicateur d'un manque d'habileté métalinguistique. Les erreurs dans la production des items peuvent aussi indiquer un manque de connaissance des morphèmes manquants, même ceux qui sont fréquents en français. Cependant, nous ne pouvons pas écarter complètement une cause de nature lexicale pour expliquer les performances des LNN dans cette tâche. Il est possible en effet qu'ils aient produit la seule forme qu'ils connaissent, parfois plus ou moins bien. Et puis, le taux élevé d'absence de réponse pour certains items pourrait être causé par la faiblesse de leurs connaissances lexicales, comme nous l'avons observé notamment pour la tâche de production lexicale (tâche 2). Cette hypothèse serait cohérente avec Schmitt et Meara (1997) et Mochizuki et Aizawa (2000) qui ont démontré qu'il y a des liens entre la connaissance des suffixes et l'étendue des connaissances lexicales.

Dans la tâche 4, les LNN ne semblent pas avoir perçu dans les paires transparentes les similitudes à la fois phonologiques et sémantiques (Freyd et Baron, 1982). Le savoir lexical, sur le plan de la profondeur, implique la capacité à reconnaître un mot construit, de savoir qu'un mot complexe est constitué de plusieurs parties et que certaines parties peuvent se retrouver dans d'autres mots (Nation 2001). Mais il est difficile, encore une fois, d'attribuer strictement au manque de connaissance en morphologie les résultats obtenus par les LNN. Leurs faibles performances dans l'identification que les mots « chéri » et « chérir », par exemple, partagent une

signification commune peuvent avoir pour cause leurs faibles connaissances lexicales: il est possible qu'ils ignoraient la signification de ces mots ou de l'un de ces mots. Cette hypothèse trouve un écho dans les performances des LNN pour les paires de mots constituées de dérivés non transparents.

Contrairement à ce à quoi on se serait attendu et selon les observations de Bogaards (1994), les sujets LNN ont commis moins d'erreurs pour les dérivations non transparentes, comme par exemple « curieux-curiosité » que pour les paires de dérivés transparents. Si les sujets LNN ont indiqué reconnaître une parenté entre les mots « curieux/curiosité », mais pas dans la paire « chéri/chérir », il se peut que ce ne soit pas leurs connaissances des règles en morphologie dérivationnelle qui aient été impliquées, mais plutôt leurs connaissances lexicales. Nous avons vu, en effet, que dans le cas des mots non apparentés, il s'agissait de mots de plus grande fréquence, comparativement aux mots des paires de dérivés qui eux, étaient moins fréquents.

Les différences de performance entre les deux groupes n'étaient pas significatives pour la tâche 5. Selon nos observations, cette dernière tâche de l'épreuve a été réalisée avec beaucoup moins d'application et donc, les performances de l'un et l'autre groupe ne nous donnent pas d'informations vraiment fiables sur leurs connaissances.

Nous avons vu que les LNN ont globalement réalisé de meilleures performances dans les tâches faisant appel aux connaissances réceptives, ce qui était prévisible d'après ce que nous savons du développement du lexique (Read, 2004). En effet, dans une tâche productive, il faut non seulement retrouver l'item en mémoire, mais aussi sélectionner certains traits caractéristiques (Tréville, 2000), ce qui est plus exigeant que lorsqu'il s'agit de reconnaître une forme lexicale. C'est dans la tâche 2

de production lexicale que les sujets LNN ont obtenu leur plus faible score et qu'ils se distinguent le plus fortement des LN. Il s'agit là d'une mesure du vocabulaire actif contrôlé, c'est-à-dire, de ce que les sujets sont capables de produire dans un contexte spécifique, délimité. Or, leurs faibles résultats ainsi que le fort taux d'absence de réponse laissent à penser qu'il leur manque des connaissances et qu'ils ont peu d'habileté pour utiliser des items lexicaux même fréquents de la L2. L'absence de corrélation entre les tâches 1 (reconnaissance lexicale) et 2 (production lexicale) suggère qu'il y a effectivement une différence entre les connaissances lexicales réceptives et productives chez les LNN, ce qui n'est pas le cas chez les LN.

Globalement, les LNN ont mieux performé dans les tâches visant les connaissances en morphologie que dans les tâches lexicales, alors que chez les LN, les résultats sont comparables. Cela peut signifier qu'ils ont quelques notions en morphologie dérivationnelle, mais qu'elles sont insuffisantes pour combler leurs faiblesses sur le plan du lexique. La forte corrélation entre les résultats pour les tâches 2 (production lexicale) et 3 (production en morphologie) semble démontrer que les connaissances lexicales et en morphologie dérivationnelle chez ces LNN sont reliées.

En ce qui concerne les connaissances en morphologie dérivationnelle, la relation est moins claire. En effet, la corrélation entre les tâches 3 (production en morphologie) et 4 (paires apparentées) suggère qu'il y a un lien entre les connaissances réceptives et productives en morphologie dérivationnelle chez les LNN. Par contre, nous n'avons pas trouvé de corrélation entre les tâches 4 (réceptive) et 5 (productive). Le fait que beaucoup de participants semblent avoir un peu bâclé la tâche 5 est peut-être en cause.

Nous n'avons pas pu établir de corrélations entre les résultats à notre épreuve et la note finale obtenue au cours d'appoint en grammaire du français écrit pour les sujets LNN. Notre hypothèse de départ quant à l'existence d'une corrélation entre les résultats aux épreuves et les résultats obtenus à ce cours se trouve donc invalidée. Il est possible que l'écart type important dans les résultats pour la majorité des tâches de notre épreuve explique cette absence de corrélation pour les LNN. Cependant, nous pouvons supposer un lien possible entre le taux d'échec ainsi que la moyenne des résultats à ce cours et les performances des LNN aux différentes tâches du test mettant à l'épreuve leurs connaissances lexicales: en effet, il faut que le locuteur ait d'abord un bagage lexical suffisant en termes d'étendue pour être en mesure de développer toute autre habileté en L2 (Weber, 1991). Et puis, nous savons que les connaissances lexicales comptent pour beaucoup dans les performances en L2 (Laufer et Goldstein, 2004).

Nous savons qu'il y a plusieurs degrés dans les connaissances lexicales et il se peut que ces connaissances soient limitées à certains aspects sémantiques, à certains contextes (Nation, 2001). Nous pouvons faire des liens avec le modèle hiérarchique de Kroll et Goldstein (2004) selon lequel il y a une différence entre être capable de fournir la forme (actif) et fournir le sens (plus passif). Un mot « passif » peut ne pas être disponible du tout en production ou l'être seulement pour des contextes restreints (Coady et Huckin, 1997). De plus, il est possible de connaître le sens d'un mot dans un contexte, mais être incapable de l'utiliser en production (Laufer, 1997).

L'acquisition du sens spécifique des items lexicaux est un processus lent (Jiang, 2004b) et il semble que les sujets LNN de notre recherche n'aient pas une connaissance complète, au sens de Nation (1990) des items qui étaient ciblés. Leurs connaissances lexicales ne correspondent pas non plus à ce que dit Channel (1988), à savoir qu'un mot est acquis lorsqu'il est compris, que ce soit en contexte ou hors contexte et qu'on peut l'utiliser correctement. La compétence en vocabulaire

est largement de nature procédurale (Nagy et Scott, 2004): au-delà de connaître le sens et la forme d'un mot, il faut savoir l'utiliser et c'est d'ailleurs le but à atteindre dans l'acquisition en L2 (Jiang, 2004; Henriksen, 1999).

Il est établi que la connaissance des suffixes donne accès aux autres mots d'une même famille (Schmitt et Meara, 1997). Cependant, rien ne garantit que l'apprenant qui connaît un mot d'une famille morphologique aura appris les autres dérivés de cette famille (Schmitt, 1998). De plus, certaines classes de mots seraient plus complexes à apprendre que d'autres (Laufer, 1997; Schmitt, 1998). Les performances des sujets LNN, particulièrement dans l'identification des paires de mots dérivés transparents, nous semblent correspondre à cette hypothèse, d'autant plus, nous l'avons vu, qu'ils ont démontré beaucoup de faiblesses lexicales.

La production de formes erronées et le taux élevé d'absence de réponse chez les LNN ayant participé à notre recherche, particulièrement dans les tâches 2 (production lexicale) et 3 (production morphologique), peuvent traduire des comportements différents chez les sujets. Dans le premier cas, le sujet tente une réponse; dans le deuxième, il se peut qu'il choisisse de ne pas prendre de risque. Il demeure cependant que dans les deux situations, même si une partie du mot était fournie et que le contexte de la phrase donnait beaucoup d'indices, plusieurs LNN n'ont pas réussi à compléter la forme lexicale demandée parce qu'ils n'avaient pas des connaissances suffisamment solides pour le faire avec succès.

Nous ne savons pas dans quel contexte les sujets LNN qui ont participé à notre recherche ont appris la langue seconde, ni dans quelle mesure ils ont été en contact avec elle. Il demeure cependant que leurs connaissances lexicales sont faibles, mais stables, ce qui selon nous, devrait soulever des inquiétudes. Ces LNN, inscrits à un programme universitaire, auront à lire beaucoup de textes et à en produire. Nous

avons vu qu'il faut connaître entre 97 et 98% des mots pour bien comprendre un texte (Hirsch et Nation, 1992). Or, les LNN ont obtenu un score moyen de 80% au test de reconnaissance lexicale et les items testés dans notre épreuve étaient des mots fréquents et non pas des items de niveau académique. Il faut craindre que ces étudiants éprouvent des difficultés tout au long de leur parcours universitaire car leurs faiblesses lexicales ne pourront pas être compensées par leurs habiletés globales en L2 (Laufer, 1992) et qu'il semblent loin d'avoir atteint un niveau de connaissance de 10,000 familles de mots nécessaires à la poursuite d'études supérieures tel que l'ont évalué Hanzenberg et Hulstijn (1996).

Sachant que, de façon générale, les apprenants de L2 investissent peu de temps par eux-mêmes dans l'apprentissage du lexique (Cobb, 1999), il nous apparaît important que les milieux d'études qui accueillent des étudiants LNN se préoccupent de savoir quelles sont leurs connaissances lexicales et quelles interventions seraient nécessaires pour les améliorer, le cas échéant.

CONCLUSION

Notre recherche avait pour but d'évaluer les connaissances lexicales et en morphologie dérivationnelle d'un groupe d'étudiants allophones désirant poursuivre des études universitaires en français. Pour ce faire, nous avons conçu deux épreuves de même type, ciblant des items lexicaux fréquents et mesurant les habiletés en compréhension et en production. Nous avons comparé les performances des sujets LNN à celles d'adultes francophones faibles à l'écrit. Les sujets des deux groupes linguistiques étaient inscrits au même cours d'appoint en grammaire du français.

Nous avons posé comme hypothèses de départ que les LNN auraient de faibles connaissances lexicales et en morphologie dérivationnelle et qu'il n'y aurait pas d'amélioration de leurs performances entre la première et la deuxième passation du test. Ces deux hypothèses ont été validées. Selon ce que nous avons mesuré aux deux moments de passation du test, les LNN ont réalisé des performances inférieures à celles des LN, tant sur le plan lexical que morphologique. Les différences de performance entre les deux groupes linguistiques étaient d'ailleurs significatives dans quatre des cinq tâches utilisées dans cette recherche. Particulièrement, nous retenons que les LNN ont démontré de grandes faiblesses en production lexicale et que, pour toutes les tâches, leur taux de non-réponse a été plus élevé que celui des LN. De plus, les sujets LNN ne semblent pas avoir eu recours à des stratégies d'utilisation du contexte phrastique pour compléter les tâches. Nous ne nous attendions pas à trouver des différences aussi importantes entre les deux groupes de sujets. Les LNN, des apprenants avancés en L2, ont été plus faibles que des LN faibles en français écrit. De plus, cette différence ne s'est

pas amoindrie, comme le suggère la stabilité de leurs résultats au temps 1 et au temps 2 d'administration de l'épreuve.

Nous savons que la connaissance lexicale est multidimensionnelle et que le développement du lexique se fait d'une façon progressive, allant de l'absence totale de savoir jusqu'à la capacité de produire un item lexical correctement. Le lexique mental n'est pas quelque chose de statique et les connaissances enregistrées en mémoire ne sont pas homogènes; certains items lexicaux peuvent être bien intégrés, d'autres, moins (Meara, 1999). L'enregistrement en mémoire des caractéristiques des items lexicaux se fait de façon graduelle (Laufer, 1990) et la fréquence d'exposition joue un rôle majeur dans l'acquisition (Ellis, 2002). Nous nous serions attendu à des connaissances plus complètes et plus sophistiquées chez des locuteurs de L2 avancés. Les résultats obtenus dans le cadre de cette recherche tendent à démontrer que les LNN n'ont pas acquis des mots de vocabulaire fréquents en français ou qu'ils ne les maîtrisent qu'imparfaitement.

Il se peut que des sujets LNN aient atteint une certaine aisance à l'oral dans la L2, mais ne maîtrisent bien qu'un nombre relativement restreint d'items lexicaux (Laufer, 1991). Pourtant, l'étendue du vocabulaire est un facteur essentiel au développement des habiletés dans la L2, notamment en lecture (Laufer, 1993). Il est possible que les sujets LNN aient des connaissances limitées, qui manquent de raffinement, même pour des items lexicaux fréquents. Leurs connaissances lexicales insuffisantes, en termes de profondeur, expliqueraient probablement leurs faiblesses en morphologie dérivationnelle. D'autre part, la maîtrise des principes de base en morphologie dérivationnelle leur aurait permis d'améliorer leurs performances dans les tâches faisant appel à leurs connaissances lexicales.

Sur le plan de la morphologie, les sujets LNN ne semblent pas avoir des connaissances solides des suffixes fréquents en français et de leur rôle dans l'attribution de la catégorie grammaticale. Cependant, étant donné leurs faiblesses lexicales, nous ne pouvons écarter qu'elles aient aussi joué un rôle dans leurs performances dans les tâches en morphologie. D'autre part, leurs connaissances semblent être parfois limitées à quelques items, voire un seul, d'une même famille morphologique.

Nous savons que la profondeur des liens entre les items dans le lexique mental est capitale (Read, 2004) et que l'organisation de ce qui est entreposé en mémoire est tout aussi important que ce qu'elle contient pour la compétence lexicale (Meara, 1996b). Bien entendu, les sujets LNN ayant participé à notre recherche possèdent des connaissances lexicales en L2, cependant leurs performances, telles que mesurées dans cette étude, ne nous permettent pas d'affirmer qu'ils ont une connaissance solide du système lexical de la L2.

Contrairement à ce que nous avons supposé, nous n'avons pas trouvé de corrélation entre les résultats à notre épreuve et la note obtenue au cours de mise à niveau en grammaire du français chez les sujets LNN. Cette absence de lien trouve peut-être son explication dans la disparité des niveaux de connaissances chez les sujets.

Bien entendu, nous ne pouvons prétendre avoir cerné toute la question des connaissances lexicales des LNN dans cette recherche. Les mesures d'étendue du vocabulaire ont leurs limites, car il faut situer les connaissances lexicales sur un continuum (Wesche et Paribakht, 1996). Il n'était pas toujours facile de déterminer avec exactitude la cause des difficultés, justement parce que le savoir lexical a beaucoup de composantes et qu'elles sont reliées. Comme le soulignent Wesche et

Paribakht (1996), il est plus facile d'évaluer ce qui est maîtrisé que ce qui l'est moins ou pas du tout.

Des contraintes de temps lors de la passation des tests, notamment, ont limité l'étendue de notre cueillette de données. Il aurait été intéressant de pouvoir mesurer avec plus de précision la taille du vocabulaire des sujets en utilisant plusieurs versions du test de Meara (1992), par exemple, et en utilisant des items de différents niveaux de fréquence afin de déterminer plus exactement l'étendue du vocabulaire des sujets. D'autre part, l'évaluation de la profondeur des connaissances est une tâche complexe et nous n'avons pu qu'en avoir un aperçu dans le cadre des questions portant sur la morphologie dérivationnelle.

Ces dernières années, les universités francophones du Québec ont appliqué des critères d'admissibilité basés sur les compétences en français écrit. Plusieurs, sinon toutes, ont élaboré des cours de rattrapage et de mise à niveau en grammaire pour les étudiants ayant échoué au test de français. Loin de nous l'idée de remettre en question la nécessité de tels cours, mais il est possible que l'insistance sur la maîtrise de cette partie de la compétence linguistique ait occulté aux yeux mêmes des LNN leurs faiblesses lexicales. Nous savons que les apprenants d'une L2 accordent généralement peu d'énergie à l'étude du vocabulaire et il arrive très fréquemment que leurs connaissances plafonnent lorsqu'ils ont atteint un certain degré d'aisance à l'oral (Laufer, 1991). Il est possible que les faiblesses en lexique et en morphologie expliquent en partie le taux d'échec des LNN au cours d'appoint en grammaire, car nombre des sujets de notre recherche ne semblaient pas posséder une base de connaissances lexicales suffisante et solide. Mentionnons qu'une des pistes recommandées dans le rapport du Groupe de recherche en enseignement de la grammaire (GREG) était de mettre sur pied un cours de grammaire mieux adapté aux besoins spécifiques des étudiants non francophones.

Le savoir lexical, tant en termes d'étendue que de profondeur, est un puissant prédicteur de la réussite académique. Les tests en français d'admission à l'université ne lui accordent généralement que peu d'attention. Il serait peut-être utile de revoir ces tests afin qu'ils prennent davantage en compte les particularités des étudiants dont la langue maternelle n'est pas le français et pour lesquels les compétences lexicales ne sont pas vraiment comparables à celles des LN.

D'autre part, il nous semble que des efforts pourraient être faits pour sensibiliser les étudiants allophones à l'importance de posséder des connaissances solides en vocabulaire de la L2 s'ils souhaitent poursuivre des études supérieures. Notamment, ces étudiants devraient connaître le rôle important de la morphologie dans le vocabulaire français et des liens entre les connaissances en morphologie et la taille du vocabulaire:

Il n'existe pas beaucoup d'études sur les connaissances lexicales et en morphologie dérivationnelle chez des apprenants adultes du français L2 et cet état de fait a suscité notre intérêt pour cette recherche. Nous avons pu mesurer que même des LNN avancés en français et à l'aise à l'oral, peuvent avoir des compétences lexicales relativement limitées, ce qui constitue un obstacle à la poursuite avec succès d'études supérieures dans la langue seconde. Nous pensons qu'il y a des avenues de recherche pour pousser un peu loin la compréhension des besoins des LNN en regard du lexique de la L2. Outre une analyse plus fine de la taille du vocabulaire en contrôlant les niveaux de fréquence, il serait intéressant, par exemple, d'évaluer les connaissances lexicales et en morphologie dérivationnelle de tous les LNN lors de la passation du test d'admission en français à l'université et de comparer les résultats entre ceux qui le réussissent et ceux qui l'échouent. Les résultats d'une telle étude pourraient éclairer les institutions d'enseignement dans la recherche de moyens pour aider les étudiants allophones à poursuivre avec succès leur cheminement académique.

APPENDICE A

Formulaire de consentement éclairé.....p. 144

Formulaire de consentement éclairé

TITRE DU PROJET

Le rôle de la morphologie dans le développement des connaissances lexicales en français.

CHERCHEUSE

Doris Allard

Département de linguistique et
de didactique des langues
Université du Québec à Montréal

DIRECTRICE

Lori Morris, professeure

Département de linguistique et
de didactique des langues
Université du Québec à Montréal

OBJECTIF DE LA RECHERCHE

La recherche vise à évaluer la connaissance lexicale des locuteurs natifs et non natifs.

CUEILLETTE DE DONNÉES

La cueillette de données se fera à partir des 2 tests écrits, à 12 semaines d'intervalle.

CONFIDENTIALITÉ DES INFORMATIONS

Il est entendu que toutes les informations recueillies demeureront strictement confidentielles.

Ces informations seront conservées sous clé dans un classeur. Le nom du participant sera codifié. Aucune information nominative ne sera transmise dans le cadre de communications orales ou écrites.

BIENFAITS ET AVANTAGES

Chaque participant recevra, sur demande, à la fin du cours LIN-1002, les résultats des tests.

RISQUES

Le risque encouru par l'étudiant qui participe à l'étude est sous le seuil du risque minimal, c'est-à-dire que la probabilité et l'importance des éventuels inconvénients associés à cette recherche sont comparables à ceux auxquels les sujets s'exposent dans les aspects de leur vie quotidienne.

INFORMATIONS CONCERNANT LE PROJET

La chercheuse s'engage à répondre à toute question posée par les participants relativement à cette recherche. Pour ce faire, ces derniers pourront communiquer directement avec la personne identifiée au début du présent document.

RETRAIT DU PROJET

La participation de l'étudiant au projet de recherche décrit ci-dessus est tout à fait libre. Il est également entendu que l'étudiant pourra, à tout moment, mettre un terme à sa participation sans aucune pénalité et sans qu'aucune pression ne soit exercée.

Toute critique ou plainte peut être adressée à la personne identifiée au début du présent document. Si votre plainte ne peut être réglée directement avec la chercheuse principale, vous pouvez faire valoir vos droits auprès du Comité institutionnel d'éthique de la recherche avec des être humains de l'Université du Québec à Montréal.

APPENDICE B

Épreuve, version 1.....	p. 147
Épreuve, version 2.....	P. 152

1- Cet exercice porte sur la **signification** des mots.

Lisez attentivement les mots de la liste.

Si vous connaissez la signification du mot, inscrivez **O** (pour oui) dans les parenthèses.

Si vous ne connaissez pas la signification du mot ou si vous n'êtes pas certain de sa signification, inscrivez **N** (pour non) dans les parenthèses.

Vous devez faire cet exercice rapidement (Maximum 5 minutes)

Attention! Pour s'assurer que vous ne répondez pas au hasard, il y a aussi des mots imaginaires, qui n'existent pas en français, dans la liste.

- | | | |
|-------------------|------------------|--------------------|
| 1 [] choint | 2 [] délicat | 3 [] urgence |
| 4 [] avenir | 5 [] climat | 6 [] plassant |
| 7 [] chiffre | 8 [] disponible | 9 [] soutenir |
| 10 [] augmenter | 11 [] agratable | 12 [] rôle |
| 13 [] grâce | 14 [] chemin | 15 [] bletter |
| 16 [] matrer | 17 [] mauvais | 18 [] structure |
| 19 [] ratoire | 20 [] morceau | 21 [] prétenser |
| 22 [] limite | 23 [] certain | 24 [] contraire |
| 25 [] interdire | 26 [] morler | 27 [] égale |
| 28 [] évirelle | 29 [] qualité | 30 [] conseiller |
| 31 [] final | 32 [] docher | 33 [] critiquer |
| 34 [] muteau | 35 [] région | 36 [] espérer |
| 37 [] veindre | 38 [] huile | 39 [] riquer |
| 40 [] matte | 41 [] face | 42 [] poursuivre |
| 43 [] tellène | 44 [] principal | 45 [] engager |
| 46 [] dépense | 47 [] scène | 48 [] conséquence |
| 49 [] taitelle | 50 [] jadis | 51 [] domaine |
| 52 [] vide | 53 [] sombre | 54 [] mallenier |
| 55 [] faveur | 56 [] type | 57 [] permanent |
| 58 [] s'empâtrer | 59 [] pentée | 60 [] précis |

2- Lisez la phrase et complétez le mot.

- 1) La plai _____ autour du fleuve St-Laurent est une région très fertile.
- 2) L'Empereur de Chine était le chef su _____ de son pays.
- 3) On appelle cols bleus les travailleurs ma _____, alors que les cols blancs font un travail administratif.
- 4) L'Égypte est réputé pour ses excellents tis _____ de coton.
- 5) Je crois connaître cette personne, son visage m'est fa _____ .
- 6) Selon les médias, les fortes pluies pourraient in _____ une bonne partie de la capitale et causer beaucoup de dégâts.
- 7) Le tarif pour l'envoi d'un col _____ postal sera augmenté en janvier prochain.
- 8) Selon les sondages, le soccer est le sport fa _____ chez les adolescents québécois.
- 9) Paul n'arrête pas de penser au voyage. L'idée de partir le rend anx _____.
- 10) Chaque matin, je fais le même tra _____ pour me rendre au travail. Je prends l'autobus 51, puis le 31 jusqu'à la rue Rosemont.

3) Lisez la phrase et complétez le mot.

- 1) Le président a remis à Pierre Dubois la médaille des Arts, la plus haute distinction _____ dans ce domaine.
- 2) Conform _____ aux souhaits du directeur, la réunion s'est tenue dans le plus grand secret.
- 3) Le niveau record du prix du pétrole ne va pas arrang _____ les choses.
- 4) Hong-Kong a longtemps fait partie de l'empire coloni _____ britannique.
- 5) La compagnie ACME se spécialise dans la vente par correspond _____ de livres et de disques allemands.
- 6) J'offre une récompens _____ de 50\$ à la personne qui retrouvera mon chat Minou.
- 7) La danseuse avait des mouvements légers et grac _____ qui démontraient tout son talent
- 8) C'est un joueur très rus _____ pour qui tous les moyens sont bons.
- 9) Elle a du s'absen _____ pour des raisons de santé
- 10) Heureusement, le passag _____ du véhicule n'a pas été blessé.

4- Paires de mots

Cette activité porte sur la signification des mots. Pour chaque paire de mots, vous devez indiquer s'il y a un lien, une ressemblance dans la signification.

Il y a 2 réponses possibles :

0) Il n'y a pas de ressemblance de signification entre les deux mots.

1) Il y a une ressemblance de signification entre les deux mots.

Cochez ✓

	0	1
attention-attentif		
siffler-sifflet		
dur-durée		
fréquent-fréquemment		
drame-dramatique		
examen-examiner		
perdre-perdrix		
clair-clarté		
voir-vision		
chéri-chérir		
défi-déficit		
soi-soixante		
printemps-printanier		
bond-bondir		
banc-bancaire		

5- Complétez le tableau avec les mots de la même famille

Exemple :

NOM	VERBE	ADJECTIF	ADVERBE
amour	aimer	aimable	aimablement

NOM	VERBE	ADJECTIF	ADVERBE
			regrettablement
		avantageux	
	industrialiser		
priorité			

1- Cet exercice porte sur la **signification** des mots.

Lisez attentivement les mots de la liste.

Si vous connaissez la signification du mot, inscrivez **O** (pour oui) dans les parenthèses.

Si vous ne connaissez pas la signification du mot ou si vous n'êtes pas certain de sa signification, inscrivez **N** (pour non) dans les parenthèses.

Vous devez faire cet exercice rapidement (pas plus de 5 minutes).

Attention! Pour s'assurer que vous ne répondez pas au hasard, il y a aussi des mots imaginaires, qui n'existent pas en français, dans la liste.

- | | | |
|--------------------|-------------------|-------------------|
| 1 () rayelle | 2 () température | 3 () jupon |
| 4 () rôtir | 5 () condamner | 6 () attocier |
| 7 () politique | 8 () aptitude | 9 () beau-fils |
| 10 () gorge | 11 () songer | 12 () minorité |
| 13 () rapidement | 14 () cas | 15 () université |
| 16 () intente | 17 () défaut | 18 () malgré |
| 19 () hausse | 20 () crayelle | 21 () citeau |
| 22 () temance | 23 () juttemant | 24 () vendange |
| 25 () section | 26 () foit | 27 () essentiel |
| 28 () cheville | 29 () masculin | 30 () taler |
| 31 () tache | 32 () patet | 33 () revaître |
| 34 () domaine | 35 () ainsi que | 36 () réduire |
| 37 () écrivain | 38 () parti | 39 () digérer |
| 40 () limaret | 41 () empâtrer | 42 () niveau |
| 43 () parfait | 44 () prétint | 45 () s'enfoncer |
| 46 () soubouler | 47 () expression | 48 () moisseaux |
| 49 () fuite | 50 () matte | 51 () bettiaux |
| 52 () craindre | 53 () bouger | 54 () colle |
| 55 () instituteur | 56 () pillet | 57 () pourrir |
| 58 () adjectif | 59 () renoncer | 60 () vollénique |

2- Lisez la phrase et complétez le mot.

- 1) Cet arbre est vig_____. Il résiste au vent et aux grands froids de l'hiver.
- 2) Ce n'est pas une ville, mais plutôt une petite lo_____ qui compte environ 800 habitants.
- 3) Tout était tranquille lorsque su_____, sans qu'on s'y attende, le chien s'est mis à gronder.
- 4) L'église, très ancienne, constitue le principal at _____ de la région.
- 5) La foule, enth_____, a chaudement applaudi les vainqueurs.
- 6) L'unique bou_____ qu'on avait allumé éclairait faiblement la pièce.
- 7) Prends un peu d'aspirine, ça devrait te sou_____.
- 8) Ta question l'a mis dans l'em_____, il ne sait plus quoi répondre.
- 9) Ce n'est plus une enfant, mais une femme ma_____ capable de prendre une décision.
- 10) La peau de la pêche est douce comme du ve_____, celle de l'orange est rude.

3- Lisez la phrase et complétez le mot.

- 1) Il n'y a aucune indi _____ claire que le gouvernement ira dans cette direction.
- 2) Les deux pays se sont mis d'accord sur une éventuelle attaque terr _____
- 3) Il succède à ce poste à M. Bertrand, élu en 1996, qui ne souhaitait pas un renouv _____ de son mandat.
- 4) Le chef du parti s'est engagé à atteindre plus de justice sans sacrif _____ la liberté.
- 5) Un expert en construction nous a donné de préc _____ conseils pour notre maison.
- 6) Elle a demandé, sur un ton plaint _____, quand allait passer le médecin.
- 7) Après la guéri _____ d'un cancer, John est devenu un grand sportif.
- 8) Il faut maintenir un taux d'humid _____ d'environ 40% dans la maison durant l'hiver.
- 9) La crainte de blessures lui fait vivre un stress continu _____ au travail.
- 10) Une seule goutte de ce produit suffit pour parfum _____ toute une pièce .

4- Cette activité porte sur la signification des mots. Pour chaque paire de mots, vous devez indiquer s'il y a un lien, une ressemblance dans la signification

Il y a 2 réponses possibles :

0) Il n'y a pas de ressemblance de signification entre les deux mots.

2) Il y a une ressemblance de signification entre les deux mots.

Cochez ✓

	0	1
secours-secouriste		
ouvrir-ouvrier		
réagir-réaction		
sembler-semblable		
absent-absence		
pousse-poussière		
bond-bondir		
défi-définir		
étroit-étroitesse		
aise-aisance		
heure-heureux		
élan-élancer		
coup-couper		
parti-partisan		
curieux-curiosité		

5- Complétez le tableau avec les mots de la même famille

Exemple :

NOM	VERBE	ADJECTIF	ADVERBE
Amour	Aimer	Aimable	Aimablement

NOM	VERBE	ADJECTIF	ADVERBE
Facilité			
	Doubler		
		libéral	
			formellement

RÉFÉRENCES

- Anderson, R. C. et Nagy, W. E. 1991. Word meanings. In *The Handbook of Reading Research*, vol. II, R. Barr, M. L. Kamil, R. Rosenthal et P. D. Pearson (éd.), p. 690-724. New York : Longman.
- Beard, R. 1998. Derivation. In *Handbook of Morphology*, A. S. Spencer et A.M. Zwicky (éd.), p. 44-65. Oxford, Malden, Mass.: Blackwell Publishers.
- Beauvillain, C. 1996. The Integration of Morphology and Whole Word Form Information during Eye Fixations on Prefixed and Suffixed Words. *Journal of Memory and Language*, vol. 35, p. 801-820.
- Beck, I. L. , Perfetti, C. et McKeown, M. 1982. The effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, vol. 74, p. 506-521.
- Beck, I. L. et McKeown, M. G. 1991. Conditions of vocabulary acquisition. In *The Handbook of Reading Research*, vol. 2 , R. Bart, M. Kamill, P. Mosenthal et P. D. Pearson (éd.) p. 789-814. New York : Longman.
- Bogaards, P. 1994. *Le vocabulaire dans l'apprentissage des langues étrangères*. Paris: Hatier.
- Bransford, J., Brown, A. L. et Cocking, R. R. 1999. *How people learn: brain, mind, experience and school*. Washington : National Academic Press.
- Butterworth, B. 1983. Lexical access in speech production. In *Lexical Representation and Process*, W. D. Marslen-Wilson (éd.), p. 108-135. Cambridge: MIT Press.
- Bybee, J. L. 1988. Morphology as Lexical Organisation. In *Theoretical Morphology and Approaches in Modern Linguistics*, M. Hammond et M. Noona (éd.), p. 119-141. San Diego: Academic Press.
- Caramazza, A, Laudanna, A. et Romani, C. 1988. Lexical access and inflectional morphology. *Cognition*, vol. 28, p. 297-332.
- Caramazza, A. 1997. How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology*, vol. 14, p. 177-208.

- Carlisle, J. F. 1995. Morphological awareness and early reading achievement. In *Morphological aspects of language processing*, L. B. Feldman (éd.), p. 189-209. Hillsdale, NJ: Lawrence Erlbaum.
- Carlisle, J. F. 2000. Awareness of the structure and meaning of morphologically complex words : Impact on reading. *Reading and Writing: An Interdisciplinary Journal*, vol. 12, no 3-4, p. 169-190.
- Casalis, S. et Louis-Alexandre, M. F. 2000. Morphological analysis, phonological analysis and learning to read French: A longitudinal study. *Reading and Writing: An Interdisciplinary Journal*, vol.12, no 3-4, p. 303-335.
- Chall, J. S. 1987. Two Vocabularies for Reading: Recognition and Meaning. In *The Nature of Vocabulary Acquisition*, M. G. McKeown et M. E. Curtis (éd.), p. 7-17. Hillsdale, N. J. : Lawrence Erlbaum.
- Chapnik-Smith, M. 1997. How do bilinguals access lexical information? In *Tutorials in Bilingualism*, A. M. B. de Groot et J. F. Kroll (éd.), p. 145-168. Mahwah, NJ: Lawrence Erlbaum.
- Chartrand, S. G., Aubin, D., Blain, R. et Simard, C. 1999. *Grammaire pédagogique du français d'aujourd'hui*. Boucherville, Québec: Graficor.
- Chialant, D. et Caramazza, A. 1995. Where is Morphology and How is it Processed? The Case of Written Word Recognition. In *Morphological Aspects of Language Processing*, L. B. Feldmans (éd.), p. 55-76. Hillsdale, N J: Lawrence Erlbaum.
- Clark, E. 1993. *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.
- Clark. E. 1998. Morphology in language acquisition. In *Handbook of Morphology*, A. S. Spencer et A.M. Zwicky (éd.), p. 374-389. Oxford: Blackwell Publishers.
- Coady, J. et Huckin T. N. 1997. *Second Language Vocabulary Acquisition : A Rationale for Pedagogy*, Cambridge: Cambridge University Press.
- Coady, J. 1997. L2 vocabulary acquisition through intensive reading . In *Second Language Vocabulary Acquisition : A Rationale for Pedagogy*, J. Coady et T. N. Huckin, (éd.), p. 225-237. Cambridge : Cambridge University Press.
- Cobb, T. 1999. Breadth and depth of lexical acquisition with hands-on concordancing. *Computer Assisted Language Learning*, vol. 12, no 4, p. 345-360.
- Cobb, T. 2000. One Size Fits All? Francophone Learners and English Vocabulary Tests. *The Canadian Modern Language Review*, vol. 57, no 2, p. 1-21.

- Cobb, T. et Horst, M. 2000. Vocabulary size of some City University students. *City University (HK) Journal of Language Studies*, vol. 1, p. 59-68.
- Cobb, T. et Horst, M. 2004. Is there room for an Academic Word List in French? In *Vocabulary in a Second Language : Selection, Aquisition, and Testing*, B. Laufer, et P. Bogaards, (éd.), p. 15-38. Amsterdam : John Benjamin.
- Cobb, T. *The Compleat Lexical Tutor*. Site internet : <http://www.lexutor.ca>, consulté le 20 avril 2006.
- Colé, P., Marec-Breton, N., Royer, C. et Gombert, J-E. 2003. Morphologie des mots et apprentissage de la lecture. *Rééducation orthophonique*, vol. 213, p. 57-76.
- Corbin, D. 1987. *Morphologie dérivationnelle et structuration du lexique*. Lille : Presses universitaires de Lille.
- Corson, D. 1985. *The lexical bar*. New York : Pergamon Press.
- Corson, D. 1997. The Learning and Use of Academic English Words. *Language Learning*, vol. 47, no 4, p.671-718.
- de Groot, A. M. B. 1993. Word-type effects in bilingual processing tasks: Support for a mixed representational system. In *The Bilingual Lexicon*, R. Schreuder et B. Weltens (éd.), p. 27-51. Amsterdam-Philadelphia: John Benjamins.
- Département de linguistique et de didactique des langues, Université du Québec à Montréal. 2005. *Rapport du groupe de réflexion sur l'enseignement de la grammaire*, 70 p.
- Derwing, B. L. 1976. Morpheme recognition and the learning of rules for derivational morphology. *Canadian Journal of Linguistics*, vol. 21, no 1, p.38-66.
- Dijkstra, T., Morcoso del Prado Martin, F. M., Schulpen, B., Schreuder, R. et Baayen, R. 2005. A roommate in cream: Morphological family size effects on interlingual homograph recognition. *Language and Cognitive Processes*, vol. 20, p. 7-41.
- Dumay, N., Gaskell, G. et Xianja, F. 2004. *A Day in the Life of a Spoken Word. Proceedings of the 26th Annual Meeting of the Cognitive Science Society*, Chicago, Illinois.
Document tire du site :<http://eis.bris.ac.uk/~psnmrvd/Cogsci2004.pdf>, consulté le 5 mai 2006.
- Durgunoglu, A. 1997. Bilingual Reading: It's Components, Development and Other Issues. In *Tutorial in Bilingualism*, de Groot, A. M. B. et J. Kroll (éd.), p. 255-276. Mahwah, NJ: Lawrence Erlbaum.

- Durgunoglu, A. et Öney, B. 1999. A cross linguistic comparison of phonological awareness and word recognition. *Reading and Writing*, vol. 11, no 4, p. 282-299.
- Ellis, N. 1994. Consciousness in second language learning: Psychological perspectives on the role of conscious process on vocabulary acquisition. *AILA Review*, vol. 11, p. 37-56.
- Ellis, N.C. 2002. Frequency effects in language processing. *Studies in second language*, vol. 24, p. 143-188.
- Franceschini, R., Zappatore, D. et Nitsch, C. 2003. Lexicon in the brain : what neurobiology has to say about languages. In *The Multilingual Lexicon*, J. Cenoz, B. Hufeisen et U. Jessner (éd.), p. 153-166. Dordrecht, Boston: Kluwer Academic Publishers.
- French, R. H. et Jacquet, M. 2004. Understanding bilingual memory: models and data. *Trends in Cognitive Sciences*, vol. 8, no 2, p. 87-93.
- Freyd, P. et Baron, J. 1982. Individual Differences in Acquisition of Derivational Morphology. *Journal of Verbal Learning and Verbal Behavior*, vol. 21, p. 282-295.
- Frost, R. et Grainger, J. 2000. Cross-linguistic perspectives on morphological processing : An introduction. *Language and Cognitive Processes*, vol. 15, no 4-5, p. 321-328.
- Frost, R., Grainger, J. et Rastle, K. 2005. Current issues in morphological processing : An introduction. *Language and Cognitive Processes*, vol. 20, no 1-2, p. 1-5.
- Ghadirian, S. 2002. Providing controlled exposure to target vocabulary through the screening and arranging of texts. *Language Learning and Technology*, vol. 6, no 1, p. 147-164.
- Goulden, R., Nation, P. et Read, J. 1990. How large can a receptive vocabulary be ? *Applied Linguistics*, vol. 11, no 4, p. 341-363.
- Grainger, J. 1993. Visual Word Recognition in Bilinguals. In *The Bilingual Lexicon*, R. Schreuder et B. Weltens (éd.), p. 11-25. Amsterdam-Philadelphia: John Benjamins.
- Greidanus, T., Bogaards, P., Vanderlinden, E., Nienhuis, L. et de Wolf, T. 2004. The construction and validation of a deep word knowledge test for advanced learners of French. In *Vocabulary in a Second Language*, P. Bogaards et B. Laufer (éd.), p. 191-208. Amsterdam : John Benjamins.

- Greidanus, T., Beks, B. et Wakely, R. 2005. Testing the Development of French Word Knowledge by Advanced Dutch and English-Speaking Learners and Native Speakers. *The Modern Language Journal*, vol. 89, no ii, p. 221-233.
- Grevisse, M. et Goosse, A. 1989. *Nouvelle grammaire française*, 2^e édition. Paris: Duculot.
- Grévisse, M. et Goosse, A. 1993. *Le bon usage*. 13^e édition revue et refondue par A. Goosse. Paris : Duculot.
- Grosjean, F. 1982. *Life with Two Languages. An Introduction to Bilingualism*. Cambridge: Harvard University Press.
- Grosjean, F. 1998. Studying bilinguals : Methodological and conceptual issues. *Bilingualism : Language and Cognition*, vol. 1, p. 131-149.
- Gu, Y. 2003. Vocabulary Learning in a Second Language: Person, Task, Context and Strategies. *TESL-EJ*, vol. 7, no 2.
Article tiré du site : <http://tesl-ej.org/ej26/a4.html>, consulté le 2 août 2006.
- Haastrup, K. 1991. *Lexical inferencing procedures or talking about words*. Tübingen, Germany: Gunter Nar Verlag, 249 p..
- Hancin-Bhatt, B. et Nagy, W. 1994. Lexical transfer and second language morphological development. *Applied Psycholinguistics*, vol. 15, p. 289-310.
- Hankamer, J. 1989. Morphological Parsing and the Lexicon. In *Lexical Representation and Process*, W. D. Marslen-Wilson (éd.), p. 392-408. Cambridge: MIT Press.
- Hay, J. 2001. Lexical frequency in morphology : Is everything relative ? *Linguistics*, vol. 39, no 6, p. 1041-1070.
- Hazenbergh, S. et Hulstijn, J. 1996. Defining a Minimal Receptive Second-Language Vocabulary for Non-Native University Students: An Empirical Investigation. *Applied Linguistics*, vol. 17, no 2, p. 145-163.
- Henderson, L. 1989. On Mental Representation of Morphology and Its Diagnosis by Measures of Visual Access Speed. In *Lexical representation and process*, W. D. Marslen-Wilson (éd.), p. 355-391. Cambridge: MIT Press.
- Henriksen, B. 1999. Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, vol. 21, no 2, p. 303-317.

- Hillis, A. E. 2000. The Organization of the Lexical System. In *The Handbook of Cognitive Neuropsychology*, B. Rapp (éd.), p. 185-210. Philadelphia: Taylor and Francis.
- Hirsch, D. et Nation, P. 1992. What vocabulary size is needed to read unsimplified texts for pleasure ? *Reading in a Foreign Language*, vol. 8, no 2, p. 689-696.
- Horst, M. et Meara, P. 1999. Test of a Model for Predicting Second Language Lexical Growth Through Reading. *Canadian Modern Language Review*, vol. 56, no 2, p. 308-328.
- Horst, M., Cobb, T. et Meara, P. 1998. Beyond a Clockwork Orange: Acquiring Vocabulary through Reading. *Reading in a Foreign Language*, vol. 11, no 2, p. 207-223.
- Hu, M. et Nation, I. S. P. 2000. Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, vol. 13, no 1, p. 403-430.
- Huckin, T. et Coady, J. 1999. Incidental Vocabulary Acquisition in a Second Language. *Studies in second language acquisition*, vol. 21, no 2, p. 181-193.
- Hulstijn, J. H., Hollander, M. et Greidanus, T. 1996. Incidental Vocabulary Learning by Advanced Foreign Language Students : The Influence of Marginal Glosses, Dictionary Use, and Reoccurrence of Unknown Words. *Modern Language Journal*, vol. 80, no 3, p. 327-339.
- Hulstijn, J. H. 2002. What does the impact of frequency tell us about the acquisition device? *Studies in second language acquisition*, vol. 24, p. 269-273.
- Hunt, D. et Beglar, D. 2005. A framework for developing EFL reading vocabulary. *Reading in a Foreign Language*, vol. 17, no 1, p. 23-59.
- Huot, H. 2001. *Morphologie : forme et sens des mots du français*. Paris : Armand Colin.
- Ijaz, I. H. 1986. Linguistic and cognitive determinants of lexical acquisition in a second language. *Language Learning*, vol. 36, p. 401-451.
- Jiang, N. 2000. Lexical Representation and Development in a Second Language. *Applied Linguistics*, vol. 21, no 1, p. 47-77.
- Jiang, N. et Forster, K. I. 2001. Cross-language Priming Asymmetries in Lexical Decision and Episodic Recognition. *Journal of memory and language*, vol. 44, p. 32-51.

- Jiang, N. 2002. Form-meaning mapping in vocabulary acquisition. *Studies and Second Language Acquisition*, p. 617-637.
- Jiang, N. 2004. Semantic Transfer and Its Implication for Vocabulary Teaching in a Second Language. *The Modern Language Journal*, vol. 88, no iii, p. 416-432.
- Jiang, N. 2004b. Semantic Transfer and Development in Adult L2 Vocabulary Acquisition. In *Vocabulary in a second language*, P. Bogaards et B. Laufer (éd.), p. 101-126. Amsterdam, Philadelphia: John Benjamins.
- Koda, K. 1999. Development of L2 Intraword Orthographic Sensitivity and Decoding Skills. *The Modern Language Journal*, vol. 83, no I, p. 51-64.
- Koda, K. 2000. Cross-linguistic variations in L2 morphological awareness. *Applied Psycholinguistics*, vol. 21, p. 297-330.
- Kormi-Nouri, R., Moniri, S et Nilsson, L. 2003. Episodic and semantic memory in bilingual and monolingual children. *Scandinavian Journal of Psychology*, vol. 44, no 1, p. 47-54.
- Krashen, S. 1989. We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. *The Modern Language Journal*, vol. 73, no iv, p. 440-464.
- Kroll, J. et Curley, J. 1988. Lexical memory in novice bilinguals. The role of concepts in retrieving second language words. In *Practical aspects of memory*, vol. 2, M. Grunberg, P. Morris et R. Sykes (éd.), p.389-395. London: John Wiley and Sons.
- Kroll, J. F. et Stewart, E. 1994. Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections between Bilingual Memory Representations. *Journal of Memory and Language*, vol. 33, p. 149-174.
- Kroll, J. F. et de Groot, A. M. B. 1997. Lexical and Conceptual Memory in the Bilingual: Mapping Form to Meaning in Two Languages. In *Tutorials in Bilingualism*, A. M. B. de Groot et J. F. Kroll (éd.), p. 169-199. Mahwah, New-Jersey: Lawrence Erlbaum.
- Kroll, J. F. et Tokowicz, N. 2001. The Development of Conceptual Representation for Words in a Second Language. In *One mind two languages: Bilingual language processing*, J. Nicol (éd.), p. 49-71. Malden, MA: Blackwell Publishers.
- Laudanna, A. et Cermele, A. 1997. Morpho-lexical Representations in Naming. *Language and Cognitive Processes*, vol. 12, no 1, p.49-66.

- Laufer, B. et Sim, D. D. 1985. Measuring and Explaining the Reading Threshold Needed for Academic Purposes Texts. *Foreign Language Annals*, vol. 18, no 5, p. 405-411.
- Laufer, B. 1990. 'Sequence' and 'Order' in the Development of L2 Lexis: Some Evidence from Lexical Confusion. *Applied Linguistics*, vol. 11, no 3, p. 281-296.
- Laufer, B. 1991. The development of L2 lexis in the expression of the advanced language learner. *Modern Language Journal*, vol. 75, no 4, p. 440-448.
- Laufer, B. 1992. Reading in a foreign language: how does L2 lexical knowledge interact with the reader's general academic ability. *Journal of research in Reading*, vol. 15, no 2, p. 95-103.
- Laufer, B. et P. Nation. 1995. Vocabulary Size and Use : Lexical Richness in Written Production. *Applied Linguistics*, vol. 16, no 3, p. 307-322.
- Laufer, B. 1997. What's in a word makes it hard or easy: some intralexical factors that affect learning of words. In *Vocabulary: Description, Acquisition and Pedagogy*, N. Schmitt et M. McCarthy (éd.), p. 140-155. Cambridge: Cambridge University Press.
- Laufer, B. et Paribakht, T. S. 1998. The Relationship Between Passive and Active Vocabulary: Effects of Language Learning Context. *Language Learning*, vol. 48, no 3, p. 365-391.
- Laufer, B. et Nation, P. 1999. A vocabulary-size test of controlled productive ability. *Language Testing*, vol. 16, no 1, p. 33-51
- Laufer, B. et Hulstijn, J. H. 2001. Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-Induced Involvement. *Applied Linguistics*, vol. 22, no 1, p. 1-26.
- Laufer, B. et Yano, Y. 2001. Understanding Unfamiliar Words in a Test. Do L2 Learners understand how much they don't understand? *Reading in foreign language*, vol. 13, p. 549-566.
- Laufer, B. 2003. Vocabulary Acquisition in a Second Language: Do Learners Really Acquire Most Vocabulary by Reading? Some Empirical Evidences. *The Canadian Modern Language Review*, vol. 59, no 4, p. 565-587.
- Laufer, B. et Goldstein, Z. 2004. Testing Vocabulary Knowledge: Size, Strength, and Computer Adaptiveness. *Language Learning*, vol. 54, no 3, p. 399-436.
- Leclerc, J. 1989. *Qu'est-ce que la langue?* 2^e édition. Laval, Québec : Mondia.

- Lecocq, P., Casalis, S., Leuwers, C. et Watteau, N. 1996. *Apprentissage de la lecture et compréhension d'énoncés*. Villeneuve d'Ascq: Presses Universitaires du Septentrion.
- Lehman, A. et Martin-Berthet, F. 1998. *Introduction à la lexicologie*. Paris: Dunod.
- Lexiquum, base de données textuelles. Site : <http://iro.umontreal.ca/cgi-bin/lexiquum>, consulté le 15 août 2005 et le 20 mars 2008.
- Libben, G. et Jarema, G. 2002. Mental Lexicon Research in the New Millenium. *Brain and Language*, vol. 81, p. 2-11.
- Liu, N. et Nation, I. S. P. 1985. Factors affecting guessing vocabulary in context. *RELC Journal*, vol. 16, no 1, p. 33-42.
- Longtin, C-M., Segui, J. et Hallé, P. A. 2003. Morphological priming without morphological relationship. *Language and Cognitive Processes*, vol. 18, no 3, p. 313-334.
- Longtin, C-M. et Meunier, F. 2005. Morphological decomposition in early visual word recognition. *Journal of Memory and Language*, vol. 53, no 1, p. 26-41.
- Mahony, D., Singson, M. et Mann, V. 2000. Reading ability and sensitivity to morphological relations. *Reading and Writing*, vol. 12, no 3-4, p. 191-218.
- Marslen-Wilson, W. D. 2001. Access to lexical representations: Cross-linguistic issues. *Language and Cognitive Processes*, vol. 16, no 5-6, p. 699-708.
- McQueen, J. M. et Cutler, A. 1998. Morphology in Word Recognition. In *Handbook of Morphology*, A. S. Spencer et A. M. Zwicky (éd.), p. 406-427. Oxford: Blackwell Publishers.
- Meara, P. 1984. The Study of Lexis in Interlanguage. In *Interlanguage*, C. Davis, C. Cripier et A. P. R. Howatt (éd.), p. 225-235. Edinburgh: Edinburgh University Press.
- Meara, P. et Buxton, B. 1987. An alternative to multiple choice vocabulary tests. *Language Testing*, vol. 4, no 2, p. 142-154.
- Meara, P. 1993. The bilingual lexicon and the teaching of vocabulary. In *The Bilingual Lexicon*, R. Schreuder et Weltens, B. (éd.), p.279-297. Amsterdam : John Benjamins.
- Meara, P. 1996. The Classical Research in L2 Vocabulary Acquisition. In *Words, Words, Words: The Translator and the Language Learner*, G. Anderman et M. Rogers (éd.), p. 27-39. Clevedon: Multilingual Matters.

- Meara, P. 1996b. The Dimensions of Lexical Competence. In *Competence and Performance in Language Learning*, Grown, G., Malmkjaer, K. et Williams, J. (éd.). Cambridge: Cambridge University Press, p. 33-51.
- Meara, P. 1997. Towards a new approach to modelling vocabulary acquisition. In *Vocabulary: Acquisition and Pedagogy*, N. Schmitt et M. McCarthy (éd.), p. 109-121. Cambridge: Cambridge University Press.
- Meara, P. 1999. Lexis: Acquisition. In *Concise Encyclopedia of Educational Linguistics*, B. Spolsky et R. E. Asher (éd.), p. 565-567. Amsterdam: Elsevier.
- Meara, P. 2005. Lexical frequency profile: A Monte Carlo analysis. *Applied Linguistics*, vol. 26, p. 32-47.
- Mirhassani, A. et Toosi, A. 2000. The impact of word-formation knowledge on reading comprehension. *International Review of Applied Linguistics*, vol. 38, p. 301-311.
- Mochizuki, M. et Aizawa, K. 2000. An affix acquisition order for EFL learners: An exploratory study. *System*, vol. 28, p. 291-304.
- Morcoso del Prado Martin, M. F., Kostic, A. et Baayen, R. H. 2004. Putting the bits together: An information theoretical perspective on morphological processing. *Cognition*, vol. 94, p. 1-18.
- Morin, R. 2003. Derivational Morphology Analysis as a Strategy for Vocabulary Acquisition in Spanish. *The Modern Language Journal*, vol. 87, no 2, p. 200-221.
- Morris, L. et Cobb, T. 2004. Vocabulary profiles as predictors of the academic performance of teaching English as a second language trainees. *System*, vol. 32, p. 75-87.
- Mulligan, N. W. et Lorito, J. P. 2006. Asymmetry Between Memory Encoding and Retrieval. *Psychological Science*, vol. 17, no 1, p. 7-11.
- Nagy, W. E., Herman, P. A. et Anderson, R. C. 1985. Learning words from context. *Reading Research Quarterly*, vol. 20, no 2, p. 233-253.
- Nagy, W. E. et Herman, P. E. 1987. Breadth and Depth of Vocabulary Knowledge: Implications for Acquisition and Instruction, p. 19-35, In *The Nature of Vocabulary Acquisition*, M. G. McKeown et M. E. Curtis (éd.), p. 19-35. Hillsdale, N. J.: Lawrence Erlbaum.

- Nagy, W. E., Anderson, R. C., Schommer, M., Scott, J. A. et Stallman, A. 1989. Morphological families in the internal lexicon. *Reading Research Quarterly*, vol. 24, p. 263-282.
- Nagy, W. E., I. N. Diakidoy et R. C. Anderson 1993. The acquisition of morphology: Learning the contribution of suffixes to the meanings of derivatives. *Journal of Reading Behavior*, no 15, p. 155-169.
- Nagy, W. E. 1997. On the role of context in first and second language vocabulary learning. In *Vocabulary: Description, Acquisition and Pedagogy*, N. Schmitt et M. McCarthy (éd.), p. 64-83. Cambridge: Cambridge University Press.
- Nagy, W. E. et Scott, J. A. 2004. Vocabulary Processes. In *Theoretical Models and Processes of Reading*, 5th edition, R. B. Ruddel et N. J. Unrau (éd.), p. 574-593. Newark, Delaware: International Reading Association.
- Nassaji, H. 2004. The Relationship Between Depth of Vocabulary Knowledge and L2 Learners' Lexical Inferencing Strategy Use and Success. *The Canadian Modern Language Review*, vol. 61, no 1, p. 107-134.
- Nation, I.S.P. 1990. *Teaching and Learning Vocabulary*. New York: Newbury House.
- Nation, I. S. P. 1993. Vocabulary Size, Growth and Use. In *The Bilingual Lexicon*, R. Schreuder et B. Weltens (éd.), p. 115-134. Amsterdam, Philadelphia: John Benjamins.
- Nation, I. S. P. et Newton, J. 1997. Teaching vocabulary. In *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*, J. Coady, et T. Huckin (éd.), p. 238-254. Cambridge: Cambridge University Press.
- Nation, I. S. P. 2001. *Learning Vocabulary in Another Language*, Cambridge: Cambridge University Press.
- Nikolova, O. 2004. Les effets des hyperliens visibles ou invisibles sur l'acquisition lexicale et sur la compréhension en lecture chez des apprenants intermédiaires et avancés en langue étrangère. *ALSIC*, vol. 7, p. 3-28.
- New, B., Pallier, C., Ferrand L. et Matos, R. 2001. Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique*, vol. 101, p. 447-462. Site internet: <http://www.lexique.org>, consulté le 12 octobre 2007.
- Nurweni, A. et Read, J. 1999. The English Vocabulary Knowledge of Indonesian University Students. *English for specific purposes*, vol. 18, no 2, p. 161-175.
- Office québécois de la langue française, *Le grand dictionnaire terminologique*.

Document consulté sur le site : <http://www.granddictionnaire.com>, le 25 septembre 2007.

- O'Gorman, E. 1996. An Investigation of the Mental Lexicon of Second Language Learners. *The Irish Yearbook of Applied Linguistics*, vol. 16, p. 15-31.
- Osborne, A. G. et Mulling, S. S. 2001. Use of morphological analysis by Spanish L2 ESOL learners. *International Review of Applied Linguistics*, vol. 39, p. 153-159.
- Pallier, C. et Argenti, A-M. 2003. *Imagerie cérébrale du langage*. In Etard, O. et Tzourio-Mazoyer, N. (éd.). Paris : Hermès Science. Texte tiré du site de C. Pallier à l'adresse: http://www.pallier.org/papers/Pallier_reviewbil.pdf, site consulté le 22 septembre 2006.
- Paradis, M. 2004. *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.
- Parel, R. 2004. The impact of lexical inferencing strategies on 2nd language reading proficiency. *Reading and Writing: An Interdisciplinary Journal*, vol. 17, p. 847-873.
- Pinker, S. 1994. Rules of language. In *Language Acquisition*, P. Bloom (éd.), p. 472-484. Cambridge: MIT Press.
- Pinker, S. 1999. *Words and rules: the ingredients of language*. New York: Basic Books.
- Plaut, D. C. et Gonnerman, L. M. 2000. Are non-semantic morphological effects incompatible with a distributed connectionist approach? *Language and cognitive processes*, vol. 15, no 4-5, p. 445-485.
- Potter, M. C., So, K. F., Von Eckhardt, B. et Feldman, L. B. 1984. Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, vol. 23, p. 23-38.
- Pressley, M., J. R. Levin et Mc Daniel, M. A. 1987. Remembering Versus Inferring What a Word Means: Mnemonic and Contextual Approaches. In *The Nature of Vocabulary Acquisition*, M. G. McKeown et M. E. Curtis (éd.), p. 107-127. Hillsdale, N. J.: Lawrence Erlbaum.
- Qian, D. 1999. Assessing the Role of Depth and Breadth of Vocabulary Knowledge in Reading Comprehension. *Canadian Modern Language Review*, vol. 56, no 2, p. 282-308.

- Qian, D. 2002. Investigating the Relationship Between Vocabulary Knowledge and Academic Reading Performance: An Assessment Perspective. *Language Learning*, vol. 52, no 3, p. 513-536.
- Rastle, K. Davis, M. H., Marslen-Wilson, W. D. et Tyler, L. K. 2000. Morphological and semantic effects in visual word recognition: A time-course study. *Language and cognitive processes*, vol. 15, no 4-5, p. 507-537.
- Raveh, M. 2002. The Contribution of Frequency and Semantic Similarity to Morphological Processing. *Brain and Language*, vol. 81, p. 312-325.
- Read, J. 1993. The development of a new measure of L2 vocabulary knowledge. *Language Testing*, vol. 10, no 3, p. 355-371.
- Read, J. 2004. Plumbing the depths: How should the construct of vocabulary knowledge be defined? In *Vocabulary in a Second Language*, P. Bogaards et B. Laufer (éd.), p. 209-227. Philadelphia: John Benjamins.
- Rey-Debove, J. 1984. Le domaine de la morphologie lexicale. *Cahiers de lexicologie*, no 45, p. 3-19.
- Riegel, M., Pellat, J. C. et Rioul, R. 1994. *Grammaire méthodique du français*. Paris: Presses universitaires de France.
- Rott, S. et Williams, J. 2003. Making form-meaning connections while reading: A qualitative analysis of word processing. *Reading in a foreign language*, vol. 15, no 1, p. 45-75.
- Saville-Troike, M. 1984. What really matters in 2nd language learning for academic achievement? *TESOL Quarterly*, vol. 18, p. 199-219.
- Schmitt, N. et Meara, P. 1997. Researching Vocabulary Through a Word Knowledge Framework, Word Associations and Verbal Suffixes. *Studies in Second language Acquisition*, vol. 20, p. 17-36.
- Schmitt, N. 1998. Tracking the Incremental Acquisition of Second Language Vocabulary: A Longitudinal Study. *Language Learning*, vol. 48, no 2, p. 281-317.
- Schreuder, R. et Flores D'Arcais, G. 1989. Psycholinguistic Issues in the Lexical Representation of Meaning. In *Lexical Representation and Process*, W. D. Marslen-Wilson (éd.), p. 409-436. Cambridge: MIT Press.
- Segalowitz, N., Poulsen, C. et Komoda, M. 1991. Lower level components of reading skill in higher level bilinguals: Implications for reading instruction. *AILA Review*, vol. 8, p. 15-30.

- Seidenberg, M. S. et Gonnerman, L. M. 2000. Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, vol. 4, no 9, p. 353-361.
- Singleton, D. 1997. Learning and processing L2 vocabulary. *Language Teaching*, vol. 30, p. 213-225.
- Singson, M., Mahony, D. et Mann, V. A. 2000. The relation between reading ability and morphological skills: Evidence from derivational suffixes. *Reading and Writing: An Interdisciplinary Journal*, vol. 12, no 3-4, p. 219-252.
- Slobin, D. I. 1973. Cognitive prerequisite for the acquisition of grammar. In *Studies of child language development*, C. A. Ferguson et D. I. Slobin (éd.), p. 173-208. New York: Holt, Rinehart and Winston.
- Smith, M. C. (1991). On the recruitment of semantic information for word fragment completion: evidence from bilingual priming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, vol. 17, p. 234-244.
- Sökmen, A. 1997. Current trends in teaching second language vocabulary. In *Vocabulary: Description, Acquisition and Pedagogy*, N. Schmitt et M. McCarthy (éd.), p. 237-256. Cambridge: Cambridge University Press.
- Stahl, S. A. 2003. How Words Are Learned Incrementally Over Multiple Exposures. *American Education*, Spring 2003, p. 18-29.
- Swan, M. 1997. The influence of the mother tongue on second language vocabulary acquisition and use. In *Vocabulary: Description, Acquisition and Pedagogy*, N. Schmitt et M. McCarthy (éd.), p. 156-180. Cambridge: Cambridge University Press.
- Taft, M. et Forster, K. I. 1975. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, vol. 14, p. 638-647.
- Taft, M. 1991. *Reading and the Mental Lexicon*. London: Lawrence Erlbaum.
- Ters, F., Reinchenbach, D. et Mayer, G. 1988. *L'échelle Dubois-Buyse de l'orthographe usuelle du français*. Orgeval, France: Éditions MDI.
- Thompson, D. M. et Tulving, E. 1970. Associative encoding and retrieval weak and strong cues. *Journal of Experimental Psychology*, vol. 86, no 2, p. 255-262.
- Tréville, M.-C. 2000. *Vocabulaire et apprentissage d'une langue seconde*. Montréal: Les Éditions Logiques.

- Tschirner, E. 2004. Breadth of Vocabulary and Advanced English Study: An Empirical Investigation. *Electronic Journal of Foreign Language Teaching*, vol. 1, no 1, p. 27-39.
- Tyler, A. et Nagy, W. E. 1989. The Acquisition of English Derivational Morphology. *Journal of Memory and Language*, vol. 28, p.649-667.
- Tyler, A. et Nagy, W. E. 1990. Use of derivational morphology during reading. *Cognition*, vol. 36, p. 17-34.
- Ulmann, M. T. 2001. The Declarative/Procedural Model of Lexicon and Grammar. *Journal of Psycholinguistic Research*, vol. 1, no 1, p. 37-69.
- van Heuven, W. J. B., Dijkstra, T. et Grainger, J. 1998. Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, vol. 39, no 3, p. 458-483.
- Vidal, K. 2003. Academic Listening: A Source of Vocabulary Acquisition ? *Applied Linguistics*, vol. 24, no 1, p. 56-89.
- Wang, M., Koda, K. et Perfetti, C. A. 2003. Alphabetic and non-alphabetic L1 effects in English word identification: a comparison of Korean and Chinese English L2 learners. *Cognition*, vol. 87, p. 129-149.
- Wang, M. et Koda, K. 2005. Commonalities and Differences in Word Identification Skills Among Learners of English as a Second Language. *Language Learning*, vol. 55, no 1, p. 71-98.
- Waring, R. 1998. *Building early vocabulary*. Conférence JALT Hammamatsu, printemps 1998.
Document tiré du site : <http://www1.harenet.ne.jp/~waring/papers/papers.html>, consulté le 2 mai 2006.
- Waring, R. et Nation, P. 2004. Second Language Reading and Incidental Vocabulary Learning. *Angles on the English-Speaking World*, vol. 4, p. 11-23.
- Waring, R. et Tagaki, M. 2003. At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a foreign language*, vol. 5, no 2, p. 130-163.
- Weber, R. M. 1991. Linguistic Diversity and Reading in American Society. In *Handbook of Reading Research*, vol. III, R. Barr (éd.), p. 97-119. New York: Longman.
- Wesche, M. et Paribakht, T. S. 1996. Assessing Second Language Vocabulary Knowledge: Depth Versus Breadth. *Canadian Modern Review*, vol. 53, no 1, p. 13-40.

White, T. G., Power, M. A. et White, S. 1989. Morphological analysis: Implications for teaching and understanding. *Reading Research Quarterly*, vol. 24, no 3, p. 285-304.