

Université de Montréal

L'erreur lexicale au secondaire
analyse d'erreurs lexicales d'élèves de 3^e secondaire et
description du rapport à l'erreur lexicale d'enseignants de français

par

Dominic Anctil

Département de didactique
Faculté des sciences de l'éducation

Thèse présentée à la Faculté des sciences de l'éducation
en vue de l'obtention du grade de doctorat
en didactique

décembre, 2010

© Dominic Anctil, 2010

Université de Montréal
Faculté des études supérieures et postdoctorales

Cette thèse intitulée :

L'erreur lexicale au secondaire : analyse d'erreurs lexicales d'élèves de 3^e secondaire
et description du rapport à l'erreur lexicale d'enseignants de français

Présentée par :
Dominic Anctil

a été évaluée par un jury composé des personnes suivantes :

Marie-Claude Boivin, présidente-rapporteuse
Pascale Lefrançois, directrice de recherche
Françoise Boch, co-directrice
Françoise Armand, membre du jury
Tom Cobb, examinateur externe
François Bowen, représentant du doyen de la FES

RÉSUMÉ

Cette recherche vise à décrire 1) les erreurs lexicales commises en production écrite par des élèves francophones de 3^e secondaire et 2) le rapport à l'erreur lexicale d'enseignants de français (conception de l'erreur lexicale, pratiques d'évaluation du vocabulaire en production écrite, modes de rétroaction aux erreurs lexicales).

Le premier volet de la recherche consiste en une analyse d'erreurs à trois niveaux : 1) une description linguistique des erreurs à l'aide d'une typologie, 2) une évaluation de la gravité des erreurs et 3) une explication de leurs sources possibles. Le corpus analysé est constitué de 300 textes rédigés en classe de français par des élèves de 3^e secondaire.

L'analyse a révélé 1144 erreurs lexicales. Les plus fréquentes sont les problèmes sémantiques (30%), les erreurs liées aux propriétés morphosyntaxiques des unités lexicales (21%) et l'utilisation de termes familiers (17%). Cette répartition démontre que la moitié des erreurs lexicales sont attribuables à une méconnaissance de propriétés des mots autres que le sens et la forme. L'évaluation de la gravité des erreurs repose sur trois critères : leur acceptation linguistique selon les dictionnaires, leur impact sur la compréhension et leur degré d'intégration à l'usage. Les problèmes liés aux registres de langue sont généralement ceux qui sont considérés comme les moins graves et les erreurs sémantiques représentent la quasi-totalité des erreurs graves. Le troisième axe d'analyse concerne la source des erreurs et fait ressortir trois sources principales : l'influence de la langue orale, la proximité sémantique et la parenté formelle entre le mot utilisé et celui visé.

Le second volet de la thèse concerne le rapport des enseignants de français à l'erreur lexicale et repose sur l'analyse de 224 rédactions corrigées ainsi que sur une série de huit entrevues menées avec des enseignants de 3^e secondaire.

Lors de la correction, les enseignants relèvent surtout les erreurs orthographiques ainsi que celles relevant des propriétés morphosyntaxiques des mots (genre, invariabilité, régime), qu'ils classent parmi les erreurs de grammaire. Les erreurs plus purement lexicales, c'est-à-dire les erreurs sémantiques, l'emploi de

termes familiers et les erreurs de collocation, demeurent peu relevées, et les annotations des enseignants concernant ces types d'erreurs sont vagues et peu systématiques, donnant peu de pistes aux élèves pour la correction. L'évaluation du vocabulaire en production écrite est toujours soumise à une appréciation qualitative, qui repose sur l'impression générale des enseignants plutôt que sur des critères précis, le seul indicateur clair étant la répétition. Les explications des enseignants concernant les erreurs lexicales reposent beaucoup sur l'intuition, ce qui témoigne de certaines lacunes dans leur formation en lien avec le vocabulaire.

Les enseignants admettent enseigner très peu le vocabulaire en classe au secondaire et expliquent ce choix par le manque de temps et d'outils adéquats. L'enseignement du vocabulaire est toujours subordonné à des tâches d'écriture ou de lecture et vise davantage l'acquisition de mots précis que le développement d'une réelle compétence lexicale.

Mots-clés : erreur lexicale, vocabulaire, analyse d'erreurs, langue première, production écrite, correction, didactique du français au secondaire, rapport à l'erreur, Théorie Sens-Texte.

ABSTRACT

This research aims to describe 1) francophone students' lexical errors in writing and 2) teachers' relation to lexical errors (conception of error, vocabulary evaluation practices, feedback provided).

The first part of the research consists in a three-level error analysis: 1) a linguistic description based on an error typology, 2) an evaluation of error gravity and 3) an explanation of the possible sources of error. The corpus analyzed is composed of 300 texts written in French class by 3rd year high school students.

The analysis revealed 1144 lexical errors. The most common are semantic problems (30%), errors related to morphosyntactic properties of words (21%) and the use of colloquial words (17%). This distribution shows that half of the lexical errors are due to a lack of knowledge of word properties other than meaning and form. The evaluation of error gravity is based on three criteria: their acceptability according to dictionaries, their impact on comprehension and their degree of integration to language. Problems related to register are usually those perceived as less serious and semantic problems represent the vast majority of serious errors. The third level of analysis concerns the possible causes of the errors and identifies three main sources: influence of oral language, semantic proximity and formal similarity between the word used and the target word.

The second part of the thesis concerns French teachers' relation to lexical errors and is based on the analysis of 224 corrected essays and eight interviews.

When correcting, teachers focus their attention on errors involving morphosyntactic properties of words (gender, invariability, government pattern), which they consider as grammatical errors. The more genuine lexical errors (semantic errors, use of colloquial words and collocation errors) are rarely pointed out, and comments provided regarding these types of errors are vague and inconsistent, giving students very few hints for correction. The evaluation of vocabulary in written production is always subject to a qualitative assessment, based on the teacher's general impression rather than specific criteria, the only clear indicator being

repetition. The explanations teachers provide about lexical problems rely heavily on intuition, which shows some deficiencies in their training in regards to vocabulary.

Keywords : lexical error, vocabulary, error analysis, French teaching, written production, correction, high school, error treatment in L1 writing, error gravity, Meaning-Text Theory.

TABLE DES MATIÈRES

Résumé	i
Abstract	iii
Table des matières	v
Liste des tableaux	xiii
Liste des figures	xiv
Liste des conventions d'écriture.....	xv
Liste des abréviations.....	xv
Remerciements.....	1
Introduction.....	1
I. PROBLÉMATIQUE.....	5
1. Contexte de la recherche.....	7
2. Problème général	10
3. Analyse du problème.....	11
3.1 La difficulté d'évaluer la maîtrise lexicale	11
3.2 L'erreur lexicale : un concept flou... mais définissable!	13
3.3 Les erreurs lexicales des élèves du secondaire.....	16
3.4 Les enseignants du secondaire et l'erreur lexicale.....	18
4. Problème spécifique	20
5. Objectifs de recherche	21
6. Pertinence et retombées de la recherche	21
CADRE CONCEPTUEL.....	23
II. LE LEXIQUE.....	25
1. Notre cadre théorique : la Théorie Sens-Texte	25
1.1 La lexicologie explicative et combinatoire	27
1.2 Justification de notre positionnement théorique.....	28
2. Lexique et vocabulaire.....	30
3. Qu'est-ce qu'un mot?	31
3.1 Le signe linguistique.....	32
3.2 Le lexème.....	33
3.3 La locution	33
3.4 L'unité lexicale (UL).....	34
3.5 Le sens.....	35

3.6 La combinatoire restreinte	37
3.6.1 La combinatoire grammaticale.....	37
3.6.2 La combinatoire lexicale.....	38
3.6.2.1 La modélisation des collocations à l'aide des fonctions lexicales	40
4. La structure du lexique.....	41
4.1 La polysémie et l'homonymie.....	42
4.2 Les classes de mots	44
4.3 Hiérarchisation sémantique.....	45
4.4 La dérivation lexicale	46
4.4.1 Synonymie	47
4.4.2 Antonymie	48
4.4.3 Conversivité	48
4.4.4 Dérivation sémantique par un changement de classe de mots	49
4.4.5 Dérivation sémantique liée à la structure prédicative	49
Conclusion.....	50
III. LA NORME.....	51
1. Norme et variation	51
2. La pluralité des normes.....	52
2.1 Les différents types de normes.....	53
3. La norme lexicale.....	59
4. Norme et français québécois	62
Conclusion.....	64
IV. L'ERREUR LINGUISTIQUE.....	65
1. Définition du concept d'erreur linguistique	65
1.1 L'erreur : un écart par rapport à quoi?	66
1.2 Notre locuteur de référence : un locuteur expert.....	68
1.3 Éléments à retenir dans notre définition de l'erreur linguistique	69
1.4 Notre définition de l'erreur linguistique.....	70
1.5 Erreur, faute, dysfonctionnement? Remarques sur la terminologie relative à l'erreur	71
2. Les travaux sur l'analyse d'erreurs	75
2.1 Les visées de l'analyse d'erreurs.....	75
2.2 Les composantes de l'analyse d'erreurs	76
2.3 Survol historique des approches de l'analyse d'erreurs.....	77
2.3.1 L'analyse contrastive.....	77
2.3.2 L'analyse d'erreurs (AÉ)	79
2.3.3 L'interlangue.....	81
Conclusion.....	83
3. Apports de ces travaux pour l'analyse d'erreurs lexicales.....	84
3.1 Les étapes de l'analyse d'erreurs	85
3.1.1 Le repérage des erreurs.....	86
3.1.2 La description/classification des erreurs	87

3.1.3 L'explication des erreurs.....	90
3.1.4 Le traitement des erreurs.....	96
3.1.5 L'évaluation des erreurs.....	99
3.2 La gravité des erreurs.....	100
3.2.1 Les critères de gravité selon Johansson (1978).....	100
3.2.2 Les critères de gravité selon James (1998).....	102
3.2.3 Survol des travaux sur l'évaluation de la gravité des erreurs.....	104
Conclusion.....	108
4. Les différences entre les erreurs en L1 et en L2	109
Conclusion.....	112
V. L'ERREUR LEXICALE	113
1. Définition de l'erreur lexicale	113
1.1 L'erreur lexicale selon Agustín Llach (2005a).....	113
1.2 Des définitions plus formelles	115
1.3 Notre définition de l'erreur lexicale	116
2. Survol de quelques travaux sur l'erreur lexicale	117
2.1 Patris et Vansnick (1992).....	117
2.2 Anctil (2005).....	118
2.3 Granger et Monfort (1994).....	120
2.4 Milićević et Hamel (2007)	122
2.5 Zimmermann (1986).....	123
2.6 Agustín Llach (2004).....	123
2.7 Engber (1995).....	125
2.8 Arnaud (1984 ; 1992).....	126
3. Éléments pour une typologie descriptive des erreurs lexicales	127
3.1 La structure de la typologie	127
3.1.1 Distinction des niveaux d'analyse	128
3.1.2 Les grandes divisions d'une typologie des erreurs lexicales.....	129
3.1.3 Le degré de raffinement des classes d'erreurs proposées.....	131
3.2 L'analyse des problèmes de sens	133
3.2.1 Les problèmes de sens chez Zimmermann (1986).....	134
3.2.2 Les problèmes de sens chez Anctil (2005).....	137
3.2.3 Les problèmes de sens chez Milićević et Hamel (2007).....	141
3.3 L'analyse des problèmes de nature formelle	143
3.4 L'analyse des problèmes liés à la combinatoire restreinte.....	146
3.5 L'analyse des problèmes lexicopragmatiques.....	147
3.6 Un mot sur les problèmes de style.....	148
Conclusion.....	149
4. L'analyse des sources des erreurs : vers une typologie explicative des erreurs lexicales.....	150
4.1 Warren (1982)	151
4.2 Zughoul (1991).....	153
4.3 Martin (1984)	154
4.4 Laufer (1992).....	155
4.5 Polguère (2007).....	156

4.6 Patris et Vansnick (1992)	157
4.7 Quelle structure pour une typologie explicative des erreurs?	157
Conclusion.....	158
VI. LE RAPPORT À L'ERREUR LEXICALE DES ENSEIGNANTS	159
1. Représentations, conceptions et attitudes	159
1.1 Les représentations sociales	160
1.2 Les représentations en sciences de l'éducation	161
1.3 Représentations et conceptions : du pareil au même?	163
1.4 Les attitudes	166
1.5 Quelques définitions	167
2. Un concept unificateur : le « rapport à ».....	168
2.1 Le rapport au savoir	169
2.2 Un exemple d'opérationnalisation de la notion de « rapport à » : le rapport à l'écrit.....	170
3. Le rapport à l'erreur lexicale des enseignants	173
Conclusion.....	176
PREMIER OBJECTIF - MÉTHODOLOGIE.....	177
VII. MÉTHODOLOGIE – PREMIER OBJECTIF.....	179
1. Constitution du corpus d'analyse.....	179
1.1 Considérations liées au corpus.....	179
1.2 Considérations liées à la population visée.....	181
1.3 La collecte des données et les considérations éthiques.....	182
2. Présentation de la démarche d'analyse.....	183
2.1 Opérationnalisation de la notion de locuteur expert et choix des ouvrages de référence.....	187
3. Présentation des outils d'analyse	192
3.1 La typologie descriptive des erreurs lexicales	192
3.1.1 Survol des classes d'erreurs de la typologie descriptive.....	194
3.1.1.1 Les erreurs formelles	195
3.1.1.2 Les erreurs sémantiques	196
3.1.1.3 Les erreurs relevant de la combinatoire restreinte.....	197
3.1.2 Le double codage	201
3.2 Proposition d'une échelle de gravité des erreurs	202
3.3 La typologie explicative des erreurs lexicales.....	205
4. Le contre-codage.....	206
5. Analyse des données	208
6. Les limites associées à notre premier objectif	209

PREMIER OBJECTIF -RÉSULTATS.....213

VIII. LES ERREURS LEXICALES EN PRODUCTION ÉCRITE EN 3^E SECONDAIRE ..215

1. Le corpus de textes.....215

**2. La description des erreurs à l'aide de la typologie des erreurs lexicales
.....217**

2.1 Les erreurs lexicales au secondaire : portrait global..... 217

2.2 Tour d'horizon des différentes classes d'erreurs du corpus 221

2.2.1 Les erreurs sémantiques 221

2.2.1.1 Les cas d'utilisation d'une UL ayant un sens proche ou synonyme du sens visé.....222

2.2.1.2 Les impropriétés 224

2.2.1.3 Le non-respect des contraintes pesant sur les actants sémantiques 225

2.2.1.4 Les pléonasmes..... 227

2.2.2 Les erreurs formelles 229

2.2.2.1 Les barbarismes 229

2.2.2.2 Les cas d'utilisation d'une UL fictive..... 232

2.2.2.3 Les cas d'utilisation d'une forme analytique..... 233

2.2.3 Les erreurs pragmatiques..... 235

2.2.3.1 Les cas d'utilisation d'un registre de langue inapproprié 235

2.2.3.2 Les autres erreurs liées à la combinatoire pragmatique 238

2.2.4 Les erreurs liées à la combinatoire morphosyntaxique 239

2.2.4.1 Les erreurs lexicomorphologiques..... 240

a. Les cas de non-respect de l'invariabilité 240

b. Les cas de non-respect du genre nominal 243

c. Les autres erreurs lexicomorphologiques 244

2.2.4.2 Les erreurs lexicosyntaxiques..... 247

a. L'utilisation d'une mauvaise préposition régie 247

b. Erreurs dans la construction d'un complément 250

c. Absence d'un complément obligatoire..... 252

d. Les erreurs lexicosyntaxiques spécifiques aux verbes 254

2.2.5 Les erreurs de combinatoire lexicale..... 260

2.2.5.1 Les collocations boiteuses ou inexistantes 260

2.2.5.2 Les cas de non-respect des contraintes d'emploi du collocatif..... 271

2.2.5.3 L'emploi d'un collocatif valable, mais qui ne correspond pas au sens visé 272

2.2.5.4 Emploi d'un mauvais dérivé sémantique 273

2.3 Les types d'unités lexicales en jeu dans les erreurs lexicales du corpus ... 276

2.4 Les classes de mots des UL en jeu dans les erreurs du corpus..... 278

2.5 Portrait des erreurs orthographiques du corpus..... 283

2.6 Réflexions sur la description des erreurs 287

2.6.1 L'identification des erreurs 287

2.6.2 Le double codage des erreurs 288

2.6.3 La typologie des erreurs lexicales 290

3. La gravité des erreurs 292

3.1 Le contre-codage et l'évaluation de la gravité des erreurs 292

3.2 Répartition des erreurs selon la gravité 295

3.3 Répartition des erreurs selon chacun des critères..... 298

3.3.1 L'acceptabilité linguistique 298

3.3.1.1 Un mot sur les dictionnaires comme outils de diagnostic de l'erreur lexicale.... 299

3.3.2 L'impact de l'erreur sur la compréhension du message..... 303

3.3.3 La perception de l'intégration de l'erreur à l'usage.....	304
3.4 Réflexion méthodologique sur notre échelle de gravité des erreurs	307
4. L'explicitation des erreurs lexicales	311
4.1 Les hypothèses des contre-codeuses à propos des sources des erreurs....	311
4.2 Les principales sources d'erreurs lexicales	314
4.2.1 L'intégration de l'erreur à l'usage.....	315
4.2.2 La proximité sémantique entre l'UL erronée et l'UL visée	317
4.2.3 La parenté formelle entre l'UL erronée et l'UL visée	319
4.2.4 L'analogie de construction entre deux UL	323
4.2.5 Une lacune dans le vocabulaire de l'élève.....	326
4.2.6 Une connaissance partielle du sens ou de la forme d'une UL.....	329
4.2.7 Le contexte syntaxique.....	332
4.2.8 La dérivation morphologique.....	336
4.2.9 Quelques hypothèses plus marginales	339
4.3 La multiplicité des sources d'erreurs possibles.....	341
Conclusion.....	344
DEUXIÈME OBJECTIF - MÉTHODOLOGIE.....	345
IX. MÉTHODOLOGIE – DEUXIÈME OBJECTIF.....	347
1. Les méthodologies liées au recueil des représentations	347
1.1 Les méthodes interrogatives	348
1.1.1 L'entretien.....	348
1.1.1.1 L'entretien d'explicitation.....	351
1.1.2 Le questionnaire	352
1.1.3 Les méthodes associatives.....	354
1.1.4 Les méthodes d'observation.....	355
1.1.5 Le recours aux données invoquées.....	357
1.2 Importance de la triangulation.....	358
Conclusion.....	358
2. Constitution d'un corpus pour la description du rapport à l'erreur lexicale d'enseignants	359
2.1 La triangulation de nos données d'analyse	360
2.2 Le recrutement des sujets	362
2.2 Échantillon.....	363
2.3 Les entrevues.....	363
2.3.1 Le premier volet de l'entretien	364
2.3.2 Le second volet de l'entretien	366
3. Le traitement et l'analyse des données.....	366
4. Les limites associées au deuxième objectif	367
DEUXIÈME OBJECTIF - RÉSULTATS.....	369
X. LES ENSEIGNANTS DE FRANÇAIS DU SECONDAIRE ET L'ERREUR LEXICALE..	371
1. Les erreurs lexicales et la correction des enseignants	372

1.1 Les erreurs pénalisées, relevées et non relevées et le classement des erreurs	372
1.2 Le traitement des erreurs et leur gravité.....	377
2. Diagnostic de sept erreurs lexicales par des enseignants de français ..	379
2.1 Une erreur de sens proche	380
2.2 Une erreur liée au typage des actants sémantiques.....	381
2.3 Une erreur de registre de langue	383
2.4 Une erreur liée à l'invariabilité	384
2.5 Une erreur liée au genre nominal.....	386
2.6 Une erreur de préposition régie	388
2.7 Une erreur de collocation.....	389
Conclusion.....	391
3. Le rapport à l'erreur lexicale d'enseignants de français du secondaire	392
3.1 La dimension conceptuelle du rapport à l'erreur lexicale	393
3.1.1 Une conception orthographique de la maîtrise de la langue	393
3.1.2 Une conception de l'erreur lexicale centrée sur le sens et la répétition.....	394
3.1.3 Une perception plutôt positive du vocabulaire des élèves.....	395
3.2 La dimension praxéologique du rapport à l'erreur lexicale.....	397
3.2.1 Les pratiques d'enseignement du vocabulaire	397
3.2.1.1 Les pratiques d'enseignement du vocabulaire en lien avec l'écriture	399
3.2.1.2 Les pratiques d'enseignement du vocabulaire en lien avec la lecture.....	400
3.2.1.2 Les activités « décrochées » en lien avec le vocabulaire.....	401
3.2.2 L'erreur lexicale et la correction de productions écrites	403
3.2.2.1 Des grilles de correction qui laissent peu de place au vocabulaire.....	403
3.2.2.2 L'évaluation du vocabulaire : une tâche floue allégée par l'expérience	405
3.2.2.3 L'évaluation du vocabulaire : un jugement global ayant peu d'impact sur le résultat final	407
3.2.2.4 Des annotations vagues et peu systématiques.....	410
3.2.2.5 Peu de suivi pour l'erreur lexicale	413
3.2.3 Le dictionnaire en classe de français.....	414
3.2.3.1 Pour les enseignants : un fidèle compagnon... pourtant méconnu	414
3.2.3.2 Pour les élèves : un outil sous-exploité, difficile d'accès... et dépassé!.....	416
Conclusion.....	418
CONCLUSION	421
1. Rappel des principaux résultats de la thèse.....	423
1.1 Les erreurs lexicales en production écrite en 3 ^e secondaire.....	423
1.2 Le rapport à l'erreur lexicale d'enseignants de français du secondaire.....	426
2. Mise en parallèle des deux objectifs de recherche.....	429
2.1 L'analyse des erreurs	429
2.2 Les conceptions des enseignants à propos du vocabulaire.....	432
2.3 Les pratiques d'enseignement du vocabulaire.....	433
2.4 Les enseignants et l'erreur de vocabulaire.....	434
3. Limites de la thèse et recherches futures.....	435
3.1 Analyse d'un corpus d'erreurs d'élèves allophones	435
3.2 Analyse d'un corpus plus vaste par plusieurs codeurs.....	436

3.3	Prise en compte de la richesse et de la variété du vocabulaire	436
3.4	Nouvelle approche de la gravité des erreurs	437
3.5	Implication des enseignants et des élèves pour l'explication des erreurs	438
3.6	Observation des pratiques réelles des enseignants liées au vocabulaire..	439
4.	Apports de la thèse, pistes et recommandations	440
4.1	Recommandations à l'intention des lexicographes	440
4.2	Recommandations à l'intention des enseignants	442
4.2.1	En lien avec l'enseignement du lexique	442
4.2.2	En lien avec le traitement de l'erreur lexicale.....	443
4.3	Recommandations à l'intention des formateurs d'enseignants	446
4.4	Recommandations à l'intention des didacticiens du lexique	447
	En guise de conclusion	447
	BIBLIOGRAPHIE.....	449
	ANNEXE 1 Différentes typologies d'erreurs lexicales.....	xix
	ANNEXE 2 Formulaire de consentement	xxvii
	ANNEXE 3 Extrait de codage (objectif 1)	xxxvii
	ANNEXE 4 Guide de contre-codage.....	xlvi
	ANNEXE 5 Questionnaire d'entrevue	liii
	ANNEXE 6 Liste des codes utilisés dans <i>QDA Miner</i>	lvii

LISTE DES TABLEAUX

Tableau 1.	Les différents types de normes selon Brent (1999 : 121).....	58
Tableau 2.	Catégorisation des erreurs lexicales selon Granger et Monfort (1994).....	121
Tableau 3.	Types d'erreurs lexicales selon Arnaud (1984).....	126
Tableau 4.	Grandes divisions des principales typologies.....	131
Tableau 5.	Les différents types d'erreurs sémantiques selon Zimmermann (1986).....	135
Tableau 6.	Les différents types d'erreurs sémantiques selon Anctil (2005).....	138
Tableau 7.	Sous-classes des impropriétés proposées par Anctil (2005).....	140
Tableau 8.	Les erreurs sémantiques chez Milićević et Hamel (2007).....	142
Tableau 9.	Oppositions involved in semantic dissonances (Martin, 1984).....	155
Tableau 10.	Informations fournies lors de l'étiquetage des erreurs.....	187
Tableau 11.	Utilisation des ouvrages pour statuer sur le caractère erroné d'un emploi lexical.....	192
Tableau 12.	Typologie descriptive des erreurs lexicales.....	194
Tableau 13.	Échelle de gravité des erreurs.....	204
Tableau 14.	Typologie explicative des erreurs lexicales.....	206
Tableau 15.	Répartition des erreurs selon les classes de la typologie.....	218
Tableau 16.	Comparaison des proportions de classes de mots avec et sans orthographe et flexion.....	279
Tableau 17.	Répartition des classes de mots dans chaque grande classe d'erreurs (pourcentages).....	280
Tableau 18.	Accord inter-juges pour les critères de compréhensibilité et de perception de l'intégration de l'erreur à l'usage.....	293
Tableau 19.	Écart entre l'évaluation de la gravité des contre-codeuses et celle du chercheur principal.....	295
Tableau 20.	Gravité des 1144 erreurs lexicales du corpus.....	296
Tableau 13.	Échelle de gravité des erreurs.....	298
Tableau 21.	Répartition des 1144 erreurs dans chacun des trois critères de gravité.....	298
Tableau 22.	Questions posées aux enseignants lors du premier volet de l'entrevue.....	365
Tableau 23.	Traitement de l'erreur en fonction des types d'erreurs.....	373
Tableau 24.	Traitement des erreurs en fonction de la gravité.....	378
Tableau 25.	Typologie des maladroites et des fautes lexicales (Patris et Vansnick, 1992).....	xxi
Tableau 26.	Typologie des problèmes lexicaux (Anctil, 2005).....	xxii
Tableau 27.	Typologie des erreurs lexicales (Milićević et Hamel, 2007).....	xxiv
Tableau 28.	Linguistic typology of lexical errors (Zimmermann, 1986).....	xxv
Tableau 29.	Classification of lexical errors (Engber, 1995: 146).....	xxvi

LISTE DES FIGURES

Figure 1.	Niveaux de représentation et formalismes correspondants dans un modèle Sens-Texte (Polguère, 1998: 14)	26
Figure 2.	Exemple de hiérarchisation sémantique.....	45
Figure 3.	Error analysis and error evaluation (Johansson, 1978: 3)	76
Figure 4.	Les différents courants de l'analyse d'erreurs (James, 1998).....	83
Figure 5.	The evaluation of errors (Johansson, 1978: 5).....	101
Figure 6.	Levels of structural similarity (Zimmermann, 1987: 56).....	155
Figure 7.	Répartition des 1144 erreurs lexicales (sans orthographe et flexion)	220
Figure 8.	Répartition des 342 erreurs sémantiques	221
Figure 9.	Répartition des 241 erreurs formelles (sans orthographe et flexion).....	229
Figure 10.	Répartition des 195 erreurs de combinatoire pragmatique.....	235
Figure 11.	Répartition des 111 erreurs de combinatoire morphologique	240
Figure 12.	Répartition des 128 erreurs de combinatoire syntaxique	247
Figure 13.	Répartition des 127 erreurs de combinatoire lexicale	260
Figure 14.	Répartition des erreurs en fonction des classes de mots.....	278
Figure 15.	Proportion des classes de mots pour chaque grande classe d'erreurs	280
Figure 16.	Répartition des 1528 erreurs orthographiques.....	284
Figure 17.	Gravité des 1144 erreurs lexicales du corpus	296
Figure 18.	Représentation des différents types hypothèses sur les 896 posées	315
Figure 19.	Prise en compte des erreurs par les enseignants dans leur correction	374
Figure 20.	Les niveaux de compétence pour le critère vocabulaire dans la grille de correction du MELS (2003: 15)	404

LISTE DES CONVENTIONS D'ÉCRITURE

Signifié (sens linguistique)	: 'décision', 'coup de main'
Signifiant (mot-forme)	: <i>décision, coup de main</i>
Unité lexicale (lexie)	: DÉCISION, COUP DE MAIN
Acceptions d'un vocable	: FEU1, FEU2
Vocables homonymes	: AVOCAT1, AVOCAT2
Agrammaticalité	: <i>*une instant magique</i>
Asémantisme	: <i>#La table sourit.</i>
Extrait de texte d'étudiant	: (112) <i>Cet homme était très apprécié par [du] le village. [4]</i>
Code d'une rédaction	: [4]
Extrait d'entrevue	: <i>c'est pas adapté au conte. [E4]</i>
Code d'un sujet d'entrevue	: [E4]
Numéro de chapitre	: I, II, III...
Numéro de section	: II-3.1. (= chapitre II, section 3.1)

LISTE DES ABRÉVIATIONS

UL	=	unité lexicale
TST	=	Théorie Sens-Texte
LEC	=	Lexicologie Explicative et Combinatoire
DEC	=	Dictionnaire Explicatif et Combinatoire

*À mon éternelle amie Julie, avec qui
j'aurais tant aimé partager cet instant.*

REMERCIEMENTS

La rédaction d'une thèse est une activité résolument solitaire, et il s'agit là, selon moi, de la plus grande difficulté à surmonter lorsqu'on s'engage dans un parcours doctoral. Les longues journées passées seul devant l'écran d'ordinateur, les questionnements si pointus qu'on ne peut vraiment les partager, les matins gris lors desquels la tâche semble insurmontable, les centaines d'heures à parcourir les rayons des bibliothèques et des milliers de pages à la recherche du chaînon manquant de notre argumentation, les tracasseries informatiques au goût de fin du monde à la veille des dépôts importants, les limbes de la procrastination, le purgatoire de l'examen de synthèse... La solitude. Autant d'instant où l'on se révèle à soi-même, à travers nos coups de génie et nos côtés moins glorieux, alors que s'échafaude notre recherche.

Heureusement, tout n'est pas noir dans cette quête en solitaire. Gouter au plaisir de la satisfaction d'une argumentation solide, d'une phrase bien tournée, d'une suite d'idées qui s'enchaînent de façon élégante. Connaître la joie de trouver LA référence tant espérée, d'échanger avec des gens qui partagent les mêmes intérêts, de trouver sa place dans le monde de la recherche en développant son petit domaine d'expertise. Savourer le bonheur de tenir entre ses mains le fruit de son travail... Bref, devenir chercheur apporte aussi son lot de petites victoires et de joies intenses, auxquelles on s'accroche pour mener à terme notre projet.

Bien entendu, malgré le caractère résolument solitaire du parcours doctoral, on n'atteint pas seul le bout de cette longue aventure. Beaucoup de gens ont contribué à mon travail, d'une façon ou d'une autre, et je voudrais ici leur rendre hommage.

Mes premiers remerciements vont bien sûr à Pascale Lefrançois, ma directrice de recherche depuis 7 ans maintenant. Par sa grande disponibilité, sa rigueur, son humour et son amour des mots, elle a su, malgré mes échecs sans cesse repoussés, croire en moi et en mon projet et trouver toujours les paroles pour me pousser vers l'avant.

Un immense merci aussi à Françoise Boch, de l'Université Stendhal-Grenoble 3, qui, après deux brèves rencontres seulement, a accepté de codiriger ma thèse. Elle m'a permis, en m'accueillant à Grenoble pour quelques semaines en début de thèse, d'entreprendre mon travail sur des bases nouvelles. Malgré mes longs silences, elle a su me procurer aux moments charnières une rétroaction des plus stimulantes en posant un regard extérieur sur mon travail.

Bien sûr, ma recherche doit énormément aux quelque 400 élèves qui m'ont donné accès à leur production écrite et aux enseignants, conseillers pédagogiques et directions d'école qui m'ont aidé dans mon recueil de données. Un merci tout spécial aux huit enseignants qui ont accepté de m'accorder une heure de leur temps pour un entretien, mais surtout, qui ont bien voulu lever le voile sur leurs techniques de correction et se sont prêtés au jeu de l'analyse de problèmes

linguistiques souvent difficiles à expliquer. Leur collaboration contribue fortement à la validité de cette thèse en la rapprochant du terrain.

Merci à Ophélie Tremblay, fidèle collègue et amie depuis mon arrivée aux cycles supérieurs, et à Isabelle Gauvin pour avoir assuré le contre-codage d'une partie de mes données. J'exprime aussi ma gratitude à Simon Collin, qui m'a initié au logiciel d'analyse qualitative QDA miner, me sauvant ainsi plusieurs heures précieuses durant la période de sprint de ma fin de thèse.

J'aimerais aussi témoigner ma reconnaissance aux membres de mon jury de thèse. Merci à Françoise Armand, qui m'a initié à la recherche, m'a montré à bâtir une problématique et m'a appris à faire dialoguer les auteurs. Merci aussi à Marie-Claude Boivin qui, par son esprit critique, sa rigueur linguistique et ses remarques pertinentes, a grandement contribué à la qualité de cette thèse. Mes remerciements vont aussi à Tom Cobb, professeur au département de didactique des langues de l'UQAM, qui a accepté d'agir en tant qu'examineur externe sans pourtant connaître mes recherches; j'espère que mon travail saura vous intéresser.

Merci à Alain Polguère et Igor Mel'cuk pour m'avoir donné la pique de la linguistique et m'avoir initié à la Théorie Sens-Texte, ainsi qu'à Jasmina Milicevic pour m'avoir accompagné, l'instant d'une collaboration, dans ma réflexion linguistique sur l'erreur lexicale.

Au-delà de tous ces gens qui gravitent dans le monde universitaire, j'aimerais aussi exprimer ma profonde gratitude à ceux qui, sans avoir contribué directement à ma thèse, ont partagé ma vie tout au long de cette longue aventure. Ces piliers qui m'ont entouré et soutenu, avec qui j'ai partagé mes états d'âme, mes questionnements, mes angoisses. Ceux qui ont dû, parfois malgré eux, vivre avec celle que j'ai affectueusement surnommée THÉRÈSE et qui ont aussi eu la sagesse de m'en détourner quelques fois. Ceux qui, par leurs encouragements, leurs coups de pouce et leur rire, m'ont permis de mener à terme ce projet. Moana, Hélène, Sara-Anne, Alexandre, Josée, Philippe et Audrey, je vous dois en grande partie cette thèse

Je ne peux conclure sans un mot pour ceux que je ne pourrai jamais assez remercier : mes parents, Janine et Réjean, pour leur appui inconditionnel, leur présence rassurante, leur fierté. Je vous aime.

Je désire finalement souligner l'appui précieux du CRSH, du FQRSC, du Département de didactique et de la Faculté des sciences de l'éducation ainsi que de la Faculté des études supérieures et postdoctorales de l'Université de Montréal, qui m'ont permis, par leur soutien financier, de mener à terme cette thèse et de faire ma place en tant que chercheur en participant à divers colloques et congrès internationaux.

INTRODUCTION

Notre projet de maîtrise sur l'erreur lexicale chez les étudiants universitaires nous avait laissé avec une foule de questions qu'il nous semblait important d'explorer. En effet, les informations obtenues alors nous paraissaient trop incomplètes pour pouvoir proposer de réelles interventions didactiques en lien avec l'enseignement du vocabulaire ou des outils efficaces pour le traitement des erreurs. Il nous fallait pour ce faire disposer de données plus complètes sur les erreurs des élèves et entrer en contact avec les enseignants, qui doivent réagir quotidiennement aux erreurs lexicales de leurs élèves.

C'est donc dans un esprit de continuité que nous avons, à l'automne 2005, amorcé notre thèse de doctorat sur l'erreur lexicale au secondaire. Cette longue entreprise, qui s'est étalée sur un peu plus de cinq années, nous a permis de pousser beaucoup plus loin nos réflexions, de nous nourrir de davantage de travaux issus de différents domaines de recherche, de raffiner les outils d'analyse proposés dans la maîtrise et d'en proposer de nouveaux, de peaufiner nos analyses linguistiques. Ce travail nous a aussi permis un contact plus étroit avec le monde scolaire, tout d'abord par l'analyse de centaines de rédactions d'élèves de troisième secondaire, mais aussi par de stimulantes rencontres avec des enseignants de français. C'est le fruit de ces cinq années de travail que contient cette thèse, et nous espérons qu'elle saura vous captiver autant qu'elle a su nous passionner, malgré les embuches, les périodes d'errance et les remises en question.

Cette thèse est organisée selon le modèle classique des travaux en didactique. Le premier chapitre expose notre problématique de recherche et présente les deux objectifs que nous nous sommes fixés. Les chapitres II à VI constituent notre cadre conceptuel et définissent les concepts nécessaires à notre réflexion sur l'erreur lexicale. Nous y présentons d'abord les notions linguistiques qui nous permettront d'analyser avec finesse les centaines d'erreurs de notre corpus de

textes. Les chapitres III et IV s'intéressent respectivement aux questions de *norme* et d'*erreur*, questions centrales lorsque l'on veut s'intéresser à tout type d'erreurs. Le chapitre IV contient aussi un survol des grands courants d'analyse d'erreurs et pose les bases de la démarche d'analyse privilégiée dans cette thèse. Le cinquième chapitre définit pour sa part le concept d'*erreur lexicale* et s'intéresse à diverses recherches ayant abordé cette question sous différents angles. Le chapitre VI, dernier du cadre conceptuel, présente une construction du concept de *rapport à l'erreur lexicale des enseignants*, concept qui nous a permis de décrire en détail de quelle façon les enseignants traitent les erreurs lexicales de leurs élèves.

Le chapitre VII détaille la méthodologie mise en place pour l'analyse des erreurs lexicales relevées dans notre corpus et présente les outils d'analyse que nous avons développés. Comme notre démarche d'analyse d'erreurs est assez complexe, nous avons jugé préférable de présenter les résultats liés à notre premier objectif de recherche immédiatement après la méthodologie relative à ce volet de la thèse. Le chapitre VIII présente donc une description détaillée des problèmes lexicaux d'élèves du secondaire. Les résultats et la discussion sont présentés en parallèle tout au long du chapitre, ce qui facilite la compréhension, compte tenu du volume important de données linguistiques présentées. À la suite de chacune des sections du chapitre, nous avons intégré un encadré « implications didactiques », qui fait le pont entre les résultats exposés et les pistes didactiques qu'ils nous inspirent.

Les chapitres IX et X concernent quant à eux notre deuxième objectif de recherche, soit l'élucidation du rapport à l'erreur lexicale d'enseignants de français. Le chapitre IX explore différentes méthodes d'accès aux conceptions/ représentations et détermine celles que nous retenons dans le cadre de notre étude. Le chapitre X présente les résultats relatifs à la façon dont les enseignants traitent l'erreur lexicale lors de la correction de productions écrites et met en lumière certaines de leurs conceptions relatives à l'enseignement du vocabulaire et à l'erreur lexicale.

La conclusion, qui constitue le dernier chapitre de la thèse, rappelle les principaux résultats de notre travail et met en relation certaines conclusions issues des deux volets de notre recherche. Un survol des principales limites de notre thèse nous permet ensuite de proposer des pistes pour des recherches futures, avant de formuler diverses recommandations visant un meilleur traitement de l'erreur lexicale en classe de français et un enseignement du vocabulaire plus efficace.

Lançons-nous donc dans ce long périple au cœur d'une dimension de l'écriture rarement abordée : l'utilisation du vocabulaire.

I. PROBLÉMATIQUE

1. CONTEXTE DE LA RECHERCHE

L'enseignement du lexique est depuis toujours le parent pauvre de la didactique du français langue première (Bastuji, 1978 ; Simard, 1994). Par contre, depuis quelques années, les nombreuses recherches témoignent d'un engouement nouveau pour cette dimension de l'enseignement de la langue¹. La place qu'accorde le dernier programme de français au secondaire du ministère de l'Éducation (2006) dans sa description des contenus d'enseignement (section « éléments d'apprentissage : notions et concepts ») témoigne aussi de cet intérêt pour le lexique.

Par contre, il suffit de s'intéresser un peu à l'enseignement du lexique pour constater quelques faits contradictoires. Par exemple, malgré la multiplication des travaux s'intéressant au lexique dans l'enseignement et les modifications qu'a subies le programme du Ministère, il semble que les enseignants enseignent somme toute assez peu le lexique en classe, ou du moins pas de façon systématique. Si nous ne disposons pas de preuves pour appuyer cette affirmation pour ce qui est de l'enseignement du lexique au Québec – aucune recherche, à notre connaissance, ne s'est intéressée à cette question –, nous savons que c'est le cas en France. Par exemple, Reboul-Touré (2002: 123) affirme que « les enseignants travaillent essentiellement la grammaire, la lecture, l'écriture et délaissent le lexique qui se retrouve souvent au service d'autres apprentissages. » Vancomelbeke (2004: 7) renchérit lorsqu'il affirme :

Force est de constater que l'apprentissage du vocabulaire dans les écoles est souvent considéré, quel que soit le niveau d'enseignement, comme « la cinquième roue du carrosse ». Contrairement à la grammaire, à la conjugaison ou à l'orthographe, qui font l'objet d'études régulières, voire systématiques, le vocabulaire est souvent abordé, cahin-caha, sans qu'un objectif précis soit véritablement formulé.

Dans des entrevues menées par Dreyfus (2004), les enseignants du primaire affirment accorder peu de temps à l'enseignement du lexique, même si le

¹ Voir notamment les ouvrages de Calaque et David (2004) et de Grossmann, Paveau et Petit (2005), ouvrages collectifs visant à jeter les bases d'une « didactique du lexique » qui en est encore à ses balbutiements.

programme français, à ce niveau, est plus étoffé que le programme québécois en ce qui concerne les contenus lexicaux (Tremblay, 2009). Même du côté anglophone, les chercheurs sont arrivés depuis longtemps à des observations similaires et dénoncent un enseignement du lexique « anémique » dans les programmes de formation en lecture (Durkin, 1979 ; Jenkins et Dixon, 1983 ; Zimmerman, 1993). Nous pouvons donc supposer que la situation est semblable au Québec; c'était du moins le cas il y a une dizaine d'années, alors que Préfontaine affirmait en réaction à la recherche du groupe DIEPE (1995) que les contenus liés au vocabulaire (synonymie, antonymie, champ lexical) et le registre de langue « devraient certainement être enseignés avec plus d'intensité qu'ils ne le sont » (Préfontaine, 1998: 96).

De plus, l'approche qualitative de l'enseignement du lexique basée sur une démarche de découverte prônée par la plupart des chercheurs (Simard, 1994 ; Polguère, 2004b ; Vancomelbeke, 2004) semble céder sa place en classe à une approche plutôt quantitative qui s'appuie largement sur les listes de vocabulaire décontextualisé (Fortier, 1994), dont les limites ont été plusieurs fois soulevées (Quivy et Tardieu, 1995 ; Huet, 1997). Plutôt que d'essayer d'inculquer aux élèves les grands principes qui structurent le lexique, l'enseignement du lexique reste résolument un « enseignement du vocabulaire » ciblant des mots spécifiques choisis par l'enseignant, et généralement abordés du point de vue de l'orthographe lexicale (Dreyfus, 2004) qui, nous le verrons, n'est qu'une des dimensions nécessaires à une réelle maîtrise² d'une unité lexicale³.

Certaines raisons expliquent la place peu enviable du lexique dans l'enseignement du français; Vancomelbeke (2004) pose trois hypothèses. Tout d'abord, la taille imposante du lexique – voire l'impossibilité d'en définir les limites – fait qu'on ne peut aspirer à l'enseigner dans son entièreté (Rocquet, 1998). Se pose alors la difficile question de la sélection des mots à enseigner,

² Ce travail est rédigé en nouvelle orthographe.

³ Pour des raisons qui seront exposées dans notre cadre conceptuel, nous éviterons d'employer le terme *mot*, qui dénote plusieurs types d'entités linguistiques. Pour une définition d'*unité lexicale*, ou *lexie*, voir II-3.4.

question devant laquelle les enseignants peuvent se trouver dépourvus, d'autant plus que rares sont les outils destinés à les guider en cette matière. Ensuite, alors que la grammaire est régie par des règles claires et plutôt consensuelles, le lexique est le lieu des irrégularités : la norme lexicale est fuyante et les mots, pour reprendre l'image de Polguère (2008), sont capricieux. Normal, alors, que les enseignants se tournent vers des contenus plus « rassurants », comme les catégories grammaticales, les fonctions des mots et les règles d'accord (Préfontaine, 1998: 96), qui par ailleurs posent énormément de problèmes aux élèves. Finalement, l'acquisition du vocabulaire est perçue comme naturelle par nombre d'enseignants⁴ (Dreyfus, 2004), ce qui légitime leur choix de reléguer l'enseignement du lexique à l'arrière-plan de leurs préoccupations et d'attraper « au vol » l'occasion de parler de vocabulaire lorsque celle-ci se présente. Aux trois hypothèses avancées par Vancomelbeke (2004), ajoutons que les contraintes de temps, les programmes de formation étoffés et les évaluations ministérielles, qui forcent à établir des priorités didactiques, ont sans doute un rôle important à jouer dans la place périphérique qu'occupe l'enseignement du lexique en classe de français.

Pourtant, même s'ils accordent toujours une place limitée au vocabulaire, plusieurs enseignants de français continuent de se plaindre de la pauvreté lexicale de leurs élèves (Rocquet, 1998 ; Dreyfus, 2004 ; Vancomelbeke, 2004). Mais ce « manque de vocabulaire » existe-t-il vraiment? Sur quoi se fonde le jugement des enseignants?

En effet, les résultats obtenus par les élèves québécois aux diverses épreuves de français semblent contredire cette critique. Par exemple, 87,9 % des élèves de 5^e secondaire ont obtenu la cote A ou B pour le critère « utilisation des mots » à l'épreuve unique de juin 2004⁵. Ces résultats en apparence encourageants en

⁴ Ce qui n'est pas faux, l'accroissement du vocabulaire se faisant principalement de façon accidentelle après l'âge de 7 ou 8 ans.

⁵ Nous aurions aimé fournir des statistiques plus à jour concernant l'épreuve unique. Malheureusement, tous les documents récents faisant référence aux résultats obtenus par les jeunes Québécois à cette épreuve ne présentent pas les résultats associés à chacun des critères

disent probablement plus sur le caractère flou des critères (très souvent qualitatifs et subjectifs) permettant de juger du vocabulaire employé en production que sur la réelle compétence lexicale des élèves. En effet, des recherches (Groupe DIEPE, 1995 ; Lefrançois, Laurier, Lazure et Claing, 2005) ont démontré que, si le lexique constitue la dimension la moins problématique en production de textes, il constitue généralement l'aspect le moins bien réussi dans des tests ciblés, par exemple des questions à choix multiples. Ceci nous porte à croire que l'apparente maîtrise lexicale des élèves en production témoigne plutôt de la difficulté à évaluer le vocabulaire d'un individu à partir d'une courte production écrite. Par ailleurs, l'évaluation du vocabulaire en classe reposant presque exclusivement sur des productions écrites dans lesquelles le vocabulaire est apprécié de façon qualitative, il est difficile de savoir quelles sont les lacunes lexicales réelles des élèves.

2. PROBLÈME GÉNÉRAL

Devant les faits exposés plus haut, nous constatons que nous savons peu de choses sur la façon dont se déroule l'enseignement du lexique en classe, puisque les nombreux contenus apparaissant dans les programmes ne semblent pas refléter la place accordée au lexique par les enseignants. Ce manque de connaissances nous semble d'autant plus criant pour le niveau secondaire : le programme du ministère de l'Éducation y est assez étoffé en ce qui concerne le lexique, mais nous ignorons de quelle façon cela se concrétise dans les pratiques d'enseignement. De plus, nous avons des appréciations plutôt contradictoires de la maîtrise lexicale des élèves du secondaire : d'un côté, des enseignants qui se plaignent de la pauvreté du vocabulaire de leurs élèves et d'un autre, d'excellents résultats à différentes épreuves de français. Que sait-on alors véritablement de la maîtrise du vocabulaire des jeunes? En effet, le vocabulaire donnant rarement lieu à des tests ciblés, il est presque exclusivement évalué à travers les productions écrites, dans lesquelles il est généralement apprécié à

de l'épreuve, mais simplement les résultats globaux. Tous nos efforts auprès du Ministère pour obtenir des données plus récentes sont restés vains.

l'aide de critères subjectifs et un peu vagues. Le portrait que nous avons des problèmes lexicaux des élèves du secondaire est, en conséquence, très imprécis.

3. ANALYSE DU PROBLÈME

En partant du constat que nous disposons pour l'instant d'un portrait plutôt lacunaire des problèmes lexicaux des élèves du secondaire, nous avons choisi de nous servir du concept d'erreur lexicale comme pierre angulaire pour éclairer une dimension de l'enseignement du lexique. Bien entendu, l'analyse d'erreurs lexicales ne constitue qu'une facette de l'évaluation du vocabulaire, qui n'est elle-même qu'une composante de l'enseignement du lexique. Pourtant, la notion d'erreur lexicale nous semble fondamentale, puisqu'elle est une source précieuse d'informations sur la façon dont les élèves maîtrisent (et acquièrent) ce réseau imposant et hétéroclite qu'est le lexique. Il nous paraît donc important d'en connaître davantage sur les erreurs lexicales des élèves et le traitement qu'en font les enseignants en classe. La présente section propose donc un survol de ce que nous connaissons à ce jour sur ces questions.

3.1 LA DIFFICULTÉ D'ÉVALUER LA MAÎTRISE LEXICALE

La difficulté pour les enseignants d'évaluer la maîtrise lexicale des élèves constitue sans doute un des problèmes majeurs liés à l'enseignement du lexique. Interroger un élève sur un nombre limité d'unités lexicales particulières ne permet pas de tirer des généralisations sur sa maîtrise lexicale. Les tests ciblés, en effet, nous informent sur la connaissance qu'ont les élèves d'unités lexicales particulières et ne permettent pas une évaluation très précise de leur maîtrise lexicale générale⁶.

⁶ Nous faisons cette affirmation dans un contexte d'enseignement d'une langue première. En effet, certains tests visant à évaluer le vocabulaire d'apprenants d'une langue seconde – nécessairement beaucoup plus restreint que celui d'un locuteur natif –, souvent utilisés à des fins de classement, semblent fournir des portraits assez justes de la connaissance lexicale d'un individu à un stade donné de son apprentissage (voir par exemple le *Vocabulary Levels Test* (Nation, 1983, 1990), le *Eurocenters Vocabulary Size Test* (Meara et Buxton, 1987) ou le *Lexical Frequency Profile* (Laufer et Nation, 1995)).

L'analyse de rédactions donne aussi une perception partielle de la maîtrise lexicale d'un élève, puisque ces dernières ne contiennent qu'un échantillon limité d'unités lexicales choisies par l'élève, probablement celles qui lui sont familières. Lépine (1995: 24) avait déjà émis certaines réserves par rapport à cet outil d'évaluation : « on s'est plutôt rendu compte qu'un texte de 500 mots ne constituait pas un échantillon suffisant pour vérifier la maîtrise du lexique chez un individu – sans parler de la difficulté de cerner et de noter la faute de vocabulaire. » Pourtant, la production écrite demeure le lieu privilégié d'évaluation du vocabulaire dans le cours de français langue première, puisqu'il fait rarement l'objet d'évaluations spécifiques. En effet, la littérature concernant l'enseignement du français langue première ne mentionne que très peu de tests de vocabulaire ciblés, outre peut-être ceux visant l'évaluation de la maîtrise de l'orthographe lexicale, la dictée par exemple. Par contre, le critère *vocabulaire* est présent dans presque toutes les grilles de correction de textes, mais la façon dont il est traité témoigne souvent de la difficulté de l'évaluation de cette dimension du code.

Par exemple, le critère 4 de la grille d'évaluation de l'épreuve unique de français en cinquième secondaire (2002-2003), *utilisation des mots*, demande à ce que le vocabulaire utilisé soit précis, correct et appartienne au registre de langue standard; le tout doit être jugé d'un point de vue normatif et représente 5 % de la note (en comparaison à 45 % pour les aspects grammaticaux). On mentionne que le vocabulaire est évalué à travers la cohérence de l'argumentation, puisque le correcteur doit tenir compte du fait que l'élève utilise un vocabulaire varié suscitant l'intérêt du lecteur. Mais comment apprécier, d'un point de vue normatif, des critères aussi flous que la précision et la variété du vocabulaire, ou encore le fait que le vocabulaire employé suscite l'intérêt? En conséquence, les élèves québécois obtiennent, année après année, des résultats qui laissent croire qu'ils possèdent une excellente maîtrise lexicale.

Certaines études se sont appliquées à décrire en détail les problèmes de français écrit des élèves du secondaire. Nous en citerons deux : *Savoir écrire au*

secondaire : étude comparative auprès de quatre populations francophones d'Europe et d'Amérique (Groupe DIEPE, 1995) et *Évaluation de l'efficacité des mesures visant l'amélioration du français écrit du primaire à l'université* (Lefrançois *et coll.*, 2005). Les deux études ont évalué les élèves à travers deux épreuves : un questionnaire à choix multiples évaluant les connaissances linguistiques et une rédaction. Dans la recherche de Lefrançois *et coll.* (2005), les questions relatives au lexique sont les moins bien réussies dans le questionnaire à choix multiples (58,1 %⁷). Étrangement, les erreurs de lexique sont les moins fréquentes en rédaction (1,5 erreur sur 250 mots). Dans le questionnaire utilisé par le Groupe DIEPE (1995), le taux de réussite pour les items relatifs au lexique est de 56,3 %⁸, ce qui se rapproche beaucoup de l'étude de Lefrançois *et coll.* (2005). Pour la rédaction, le vocabulaire constitue encore une fois la dimension du code écrit présentant le meilleur taux de réussite. À la lumière de ces résultats, il semble donc évident que ni un test ciblant des unités lexicales particulières ni la rédaction ne permettent de donner un portrait exact et véritable de la maîtrise lexicale des élèves, qui se situe probablement quelque part entre les deux extrêmes observés. Par contre, les productions écrites ont cela d'intéressant qu'elles permettent d'observer non pas le vocabulaire que les élèves ne connaissent pas, mais bien les problèmes liés à la maîtrise du vocabulaire auquel ils ont spontanément recours dans un contexte de rédaction. Même si elle ne suffit pas à la tâche, cette prise en compte des problèmes lexicaux en production de textes constitue une facette essentielle de l'évaluation du vocabulaire, puisqu'elle peut révéler des pistes intéressantes sur la maîtrise lexicale en mettant en lumière certaines lacunes récurrentes.

3.2 L'ERREUR LEXICALE : UN CONCEPT FLOU... MAIS DÉFINISSABLE!

Nous venons de soulever le problème de l'évaluation du vocabulaire, notamment en rédaction, mais il existe un problème plus profond à notre avis : la définition

⁷ Il s'agit ici de la moyenne entre le pré-test (58,38 %) et le post-test (57,75 %).

⁸ Moyenne des quatre populations étudiées.

même de ce qu'est une erreur lexicale. En effet, il ne semble pas y avoir de consensus, ni en éducation ni en linguistique, sur ce qu'est une erreur lexicale.

Par exemple, il est intéressant de voir que la conception d'erreur lexicale varie dans les différentes études présentées plus haut. Le Groupe DIEPE (1995) a considéré dans les erreurs de vocabulaire les barbarismes, les impropriétés, les approximations lexicales et les pléonasmes; vu l'objectif de comparaison entre différentes populations, les régionalismes et familiarismes n'ont pas été comptabilisés. Lefrançois *et coll.* (2005: 23) considèrent plutôt comme des erreurs lexicales « celles relatives au sens des mots, au choix de l'auxiliaire dans les temps composés, aux collocations fautives et aux prépositions régies par des mots » ainsi que les mots appartenant au registre familier. La grille d'évaluation du Ministère, évoquée plus haut, semble quant à elle porter une attention spéciale aux problèmes de sens et de registres de langue. Cette même grille, ainsi que celle utilisée par le Groupe DIEPE (1995), tient aussi compte de la richesse et de la variété du vocabulaire, alors que Lefrançois *et coll.* (2005) ont décidé de se concentrer sur les dimensions strictement linguistiques et n'ont donc pas considéré ce critère dans l'évaluation des rédactions. Finalement, le Groupe DIEPE a aussi recensé, dans sa correction, certaines maladresses⁹ dans l'utilisation du vocabulaire, alors que seules les erreurs sont couvertes par les grilles d'évaluation utilisées par le ministère et par Lefrançois *et coll.* (2005).

Nous constatons donc que la notion même d'erreur lexicale demeure floue, puisqu'il en existe des conceptions plus ou moins larges, selon qu'elles prennent en compte des dimensions spécifiques de la maîtrise lexicale, et en fonction de leur position relativement aux maladresses lexicales, type de dysfonctionnements assez fréquent dans les rédactions (Anctil, 2005). Par conséquent, le diagnostic de problèmes lexicaux par les enseignants peut s'avérer difficile, puisque les limites mêmes du concept d'erreur lexicale ne sont pas bien définies (Agustín Llach, 2005a). Dans le même ordre d'idées, Denance

⁹ Selon les auteurs, les maladresses sont des problèmes « parfois plus difficiles à identifier de façon unanime, notamment lorsqu'elles sont liées à des questions de longueur ou de répétition » (Groupe DIEPE, 1995: 189)

(1999) a pointé la difficulté de départager les erreurs de syntaxe des erreurs de vocabulaire :

Un des défis que nous avons en tant que correcteurs et superviseurs, c'est de faire le départage entre les fautes de vocabulaire et celles de syntaxe. [...] La frontière entre syntaxe et sémantique est en effet poreuse : il faut souvent retourner au sens pour détecter et corriger une faute de syntaxe... Si on joue sur le sens, ne sommes-nous pas alors devant une faute de vocabulaire? Beaucoup de fautes hâtivement pénalisées en syntaxe se révèlent en fait de « simples » fautes de vocabulaire.

Patris et Vansnick (1992: 47) soulignent aussi la difficulté de cerner l'erreur lexicale et dénoncent « le manque d'outils adéquats, tant pour analyser les fautes que pour les corriger ». De façon plus large, Calaque et Grossmann (2000: 5) notent que

[les] enseignants se plaignent du manque de vocabulaire des enfants, tout en restant bien désarmés pour y faire face, tant sur le plan des contenus (quoi enseigner en matière de lexique ?) que sur le plan des démarches, des méthodes et des outils.

En conséquence, les enseignants se retrouvent avec peu de ressources pour aider les élèves à mieux comprendre leurs erreurs de vocabulaire et à améliorer leur compétence lexicale. D'autant plus que, comme le constate Polguère (2004b: 3),

aussi bien les élèves que les enseignants sont rarement entraînés à décrire les phénomènes lexicaux et à les comprendre [et que, par conséquent], nombre de problèmes rencontrés par les apprenants sont mal identifiés et, donc, mal traités.

Comme nous l'avons constaté plus haut, les conceptions relatives à l'erreur lexicale sont plutôt variées. Par ailleurs, la littérature s'intéressant à cette notion est plutôt limitée et seuls quelques auteurs en ont proposé une caractérisation (voir chapitre V) ; Meara (1984) soulignait d'ailleurs le peu d'intérêt pour le lexique dans les travaux sur l'analyse d'erreur et invitait ses collègues chercheurs à mener des recherches en ce sens. Il est d'autant plus difficile de comprendre cette pauvreté des travaux sur l'erreur lexicale, puisque le lexique constitue « un objet largement décrit et théorisé en linguistique » (David et Grossmann, 2003: 4) et que nous disposons de tous les outils conceptuels permettant une caractérisation claire et complète de cette notion.

Conscient de ce manque, nous avons nous-même élaboré, dans le cadre de notre maîtrise, une typologie des problèmes lexicaux présents dans un corpus de textes d'étudiants universitaires (Anctil, 2005). Le cadre théorique dans lequel nous avons développé notre typologie est celui de la théorie Sens-Texte (voir II-1), approche linguistique centrée principalement sur la lexicologie et l'étude du sens. En nous basant sur un modèle très fin de description lexicographique, nous avons été en mesure de déterminer les entorses possibles à l'ensemble des propriétés des unités lexicales, et donc à en proposer une caractérisation exhaustive reposant sur un cadre théorique précis. Notre définition de l'erreur lexicale est assez large et prend en compte non seulement le signifiant (forme) et le signifié (sens) des unités lexicales, mais aussi la combinatoire restreinte (voir II-3.6). Ainsi, notre conception de l'erreur lexicale regroupe certains dysfonctionnements souvent considérés par les enseignants comme des problèmes de syntaxe et d'accord.

3.3 LES ERREURS LEXICALES DES ÉLÈVES DU SECONDAIRE

Nous savons très peu de choses sur les problèmes lexicaux des élèves du secondaire. Les deux études s'intéressant à la maîtrise du code écrit chez les élèves du secondaire mentionnées plus haut (Groupe DIEPE, 1995 ; Lefrançois *et coll.*, 2005) soulignaient une faible présence de problèmes lexicaux dans les rédactions d'élèves, mais sans préciser les types spécifiques d'erreurs lexicales commises.

Par contre, certains problèmes observés chez les étudiants universitaires nous poussent à penser qu'ils sont aussi présents au secondaire. Nous avons constaté (Anctil, 2005) que les textes de notre corpus d'étudiants universitaires comportaient en moyenne 3,8 problèmes lexicaux par 250 mots, selon notre définition de l'erreur lexicale. Parmi ces problèmes, les plus fréquents étaient ceux relatifs au sens (impropriétés) et aux registres de langue. Près du tiers des problèmes étaient aussi attribuables à une méconnaissance de certaines propriétés de combinatoire alors qu'une erreur sur cinq concernait l'orthographe lexicale.

Nous avons aussi consulté deux recherches ayant analysé les erreurs de français écrit d'élèves québécois du secondaire : *Le français écrit au secondaire : une enquête et ses implications pédagogiques* (Bureau, 1985) et *Le français écrit au secondaire : une réévaluation par la segmentation des textes, l'étude de la cohésion et l'analyse des erreurs lexico-graphiques* (Majeau, 1992). Bien que leurs auteurs aient encore une fois une conception de l'erreur lexicale qui diffère de celle que nous proposons, ces études nous permettent des observations un peu plus fines que les deux autres mentionnées plus haut, puisqu'elles présentent des résultats relatifs à certaines sous-classes d'erreurs lexicales.

Selon Bureau (1985), les erreurs de sens et l'emploi d'anglicismes constituent 7,8 % de tous les problèmes observés. Par contre, sa division des classes d'erreurs ne nous permet pas d'établir clairement le taux d'erreurs d'orthographe lexicale, mais il semble avoisiner les 24 %, ce qui se rapproche d'ailleurs des 21 % que nous avons observés dans les rédactions d'étudiants universitaires (Anctil, 2005). Par ailleurs, les erreurs relatives au régime des unités lexicales sont classées parmi les erreurs de grammaire et diminuent le nombre de problèmes que nous aurions nous-même classés parmi les problèmes lexicaux. Selon ces résultats, nous pouvons donc déduire que les erreurs lexicales sont aussi présentes dans les rédactions d'élèves du secondaire. Par contre, cette étude date un peu, d'autant plus que le corpus étudié était constitué de textes recueillis en 1975; il se peut donc que la situation ait évolué depuis.

Pour son analyse des erreurs lexico-graphiques trouvées dans les textes de 17 élèves de deuxième et quatrième secondaires, Majeau (1992) s'est servie de la grille élaborée par Bureau (1985). Si nous faisons abstraction des erreurs d'orthographe lexicale (qui constituent environ 20,7 % des 1107 erreurs relevées), les erreurs sémantiques et les anglicismes totalisent 8,5 % de l'ensemble des erreurs. Encore une fois, les problèmes relatifs au régime ont été comptabilisés dans la catégorie syntaxique; selon notre conception large de l'erreur lexicale, nous pouvons évaluer que le total d'erreurs lexicales aurait dépassé 10 % du total, sans compter les problèmes d'orthographe lexicale. Par

contre, le nombre limité de sujets dans cette étude ne permet pas de généralisation à l'ensemble des élèves du secondaire.

Ces quelques travaux laissent donc présager la présence de problèmes lexicaux dans les rédactions d'élèves du secondaire dans une proportion non négligeable. Par contre, nous gagnerions à étudier un corpus plus vaste constitué de textes plus récents pour obtenir une description plus précise de ces problèmes.

3.4 LES ENSEIGNANTS DU SECONDAIRE ET L'ERREUR LEXICALE

Nous avons vu plus haut que certains auteurs soulignent le manque d'outils adéquats pour aider les enseignants dans le diagnostic des problèmes lexicaux dans les productions écrites de leurs élèves (Patris et Vansnick, 1992 ; Calaque et Grossmann, 2000 ; Polguère, 2004b). Par contre, on sait très peu de choses de la façon dont les enseignants de langue eux-mêmes conçoivent et traitent l'erreur lexicale en production écrite. Dreyfus (2004: 5) pointe d'ailleurs ce manque lorsqu'elle note « le peu de connaissances à propos des compétences lexicales réelles des élèves et la difficulté des enseignants à évaluer celles-ci. »

La littérature révèle très peu d'informations quant à la conception qu'ont les enseignants de l'erreur lexicale ; pourtant, il est évident que les conceptions des enseignants influencent leurs pratiques de classe (Gilly, 1989 ; Garcia-Debanc et Plane, 2004) et que, par conséquent, la façon donc ils perçoivent l'erreur guidera l'attention qu'ils y accordent et le traitement qu'ils en font (Porquier, 1977 ; Spsychiger, 2006). Les seuls indices dont nous disposons pour l'instant sont les grilles de correction de production écrite proposées par le Ministère, qui sont susceptibles de conditionner l'idée que se font les enseignants de ce qu'est un problème lexical. Nous pouvons par exemple supposer que les critères de richesse, variété et précision du vocabulaire présents dans la plupart des grilles de correction feront partie intégrante de toute appréciation du vocabulaire d'un texte par les enseignants. Mais de quelle façon opérationnalisent-ils de tels critères lors de l'évaluation? Portent-ils une attention particulière à d'autres types de problèmes, par exemple les termes impropres dont parle Jean-Denis

Moffet, responsable de l'élaboration et de la validation de l'épreuve uniforme de français (Maisonneuve, 1997)?

Les articles portant sur l'évaluation de productions écrites que nous avons consultés (Halté, 1984 ; Garcia-Debanc et Mas, 1987 ; Gueunier, 1989) ne font que peu mention du vocabulaire, sinon pour spécifier qu'il fait toujours partie des critères d'évaluation. Gueunier (1989) a tout de même observé que les futurs enseignants français semblaient démontrer d'importantes lacunes dans le diagnostic de problèmes lexicaux. Elle mentionne aussi que les enseignants, autant que les futurs maîtres, démontrent des « compétences métalinguistiques [...] plus grandes dans les domaines relevant de la surface de la langue qu'en sémantique et grammaire du texte » (p.97). Martin (1984) souligne quant à elle la difficulté d'expliquer aux élèves ce qui ne va pas dans leurs choix lexicaux douteux. Dans une étude à petite échelle dans laquelle elle demandait à des enseignants d'anglais langue seconde d'expliquer certaines erreurs lexicales comme ils l'auraient fait à leurs élèves, Zimmerman (1993) a constaté qu'ils avaient de fortes intuitions leur permettant d'expliquer aisément certains types de problèmes lexicaux – notamment ceux liés à la connotation –, mais qu'ils peinaient à fournir une rétroaction pertinente pour d'autres, particulièrement les problèmes de collocation. Malgré ces observations, nous ignorons si les enseignants eux-mêmes se sentent dépourvus devant des problèmes de nature lexicale.

Si nous ignorons à peu près tout de la façon dont les enseignants notent l'erreur lexicale, nous disposons d'encore moins d'informations relativement à la façon qu'ils ont de rétroagir auprès de leurs élèves à propos de telles erreurs. Nous savons en effet bien peu de choses quant à la façon de signaler l'erreur lexicale lors de la correction de productions écrites. Beefun (2001) a tout de même noté que les enseignants corrigent la plupart du temps l'erreur en la soulignant et en écrivant « la bonne réponse »¹⁰, soit une reformulation plus heureuse dans le cas

¹⁰ Il existe plusieurs types d'annotations pour rétroagir aux erreurs en production écrite. Pour un survol des avantages et des limites de chacun, voir Roberge (2006).

d'erreurs de vocabulaire ; ces observations découlent cependant de l'analyse d'un corpus de rédactions produites dans un contexte de français langue étrangère et touchent tant les erreurs lexicales que morphologiques et syntaxiques. Par ailleurs, nos propres observations en milieu scolaire nous ont permis de constater l'omniprésence de la traditionnelle annotation « *voc* » placée dans la marge, mais nous ignorons si cette annotation donne lieu à des interventions didactiques de la part des enseignants.

4. PROBLÈME SPÉCIFIQUE

Nous avons donc vu que nous disposons pour l'instant d'un portrait plutôt lacunaire des problèmes lexicaux des élèves du secondaire, puisque les recherches n'ont pas fourni de descriptions fines de ce type particulier de problèmes. Le fait que la caractérisation même de ce qu'est une erreur lexicale ne fait pas consensus et qu'aucune définition claire et complète n'en est généralement admise dans le monde de l'éducation ne peut que compliquer la tâche des enseignants, qui doivent malgré tout évaluer la maîtrise lexicale de leurs élèves par le diagnostic des problèmes lexicaux rencontrés dans les rédactions de ces derniers, et ce, sans outils adéquats. Par ailleurs, nous ignorons à peu près tout de la façon dont les enseignants eux-mêmes conçoivent et traitent l'erreur lexicale. Quels sont leurs critères pour affirmer que leurs élèves manquent de vocabulaire? Que considèrent-ils comme une erreur lexicale? De quelle façon indiquent-ils de telles erreurs sur les rédactions? Quel impact ont-elles sur le résultat final obtenu?

5. OBJECTIFS DE RECHERCHE

Voici donc, à la lumière des questions soulevées plus haut, les objectifs visés par la présente recherche.

1. *Analyser les erreurs lexicales d'élèves francophones de troisième secondaire¹¹ en production écrite;*
2. *Décrire le rapport d'enseignants de français du secondaire à l'erreur lexicale en ce qui concerne*
 - *leur conception de ce qu'est une erreur de nature lexicale;*
 - *leurs pratiques d'évaluation du vocabulaire en production écrite;*
 - *leurs modes de rétroaction par rapport à des erreurs lexicales.*

6. PERTINENCE ET RETOMBÉES DE LA RECHERCHE

La question de l'erreur lexicale a jusqu'à maintenant été très peu exploitée en didactique du français langue première. Pourtant, de tels problèmes existent dans les rédactions d'élèves et les enseignants y sont sans cesse confrontés. En fournissant une description complète des erreurs lexicales des élèves du secondaire, cette recherche permettra de mieux comprendre les lacunes des élèves dans leur emploi du vocabulaire et offrira certaines pistes quant à la façon dont le lexique devrait être abordé en classe.

La description du rapport des enseignants à l'erreur lexicale, notamment de la façon dont cette dernière est traitée en classe, apportera quant à elle des connaissances dans un champ de recherche jusqu'ici très peu exploré (Agustín Llach, 2004). Nous espérons qu'une meilleure compréhension de ce rapport nous permettra de cibler davantage les écueils qui peuvent surgir lors du diagnostic de problèmes lexicaux. Par le fait même, ces constatations pourront amener des suggestions de modifications qui pourraient être apportées à la formation des futurs maîtres de français au secondaire.

Finalement, nous aimerions que cette recherche descriptive constitue un premier pas vers le développement d'un outil d'aide au diagnostic des

¹¹ Le choix de nous intéresser spécifiquement aux élèves de troisième secondaire sera justifié au chapitre VII, dans lequel nous présenterons la méthodologie de recherche associée à notre premier objectif.

problèmes lexicaux destiné aux professeurs du secondaire, puisque, comme nous l'avons souligné, un tel outil manque toujours. En s'appuyant non seulement sur des concepts issus de la linguistique, mais aussi sur un portrait précis des problèmes lexicaux des élèves, des conceptions des enseignants et de leurs pratiques, cet outil gagnera en pertinence didactique.

CADRE CONCEPTUEL

II. LE LEXIQUE

La présente thèse s'intéressant à l'erreur lexicale, il est important de bien définir les différents concepts linguistiques qui sous-tendront tant notre définition de l'erreur lexicale que l'analyse que nous ferons des problèmes rencontrés dans les productions écrites des élèves. Ce chapitre vise à définir ce qu'est le lexique, en mettant l'accent sur la notion importante d'unité lexicale et sur les différents liens existant entre les « mots ».

1. NOTRE CADRE THÉORIQUE : LA THÉORIE SENS-TEXTE

Pour assurer validité et cohérence à notre conception de l'erreur lexicale, il nous apparaît primordial que notre réflexion s'inscrive dans un cadre théorique précis. Ce cadre théorique sera la *Théorie Sens-Texte* (TST), et plus précisément la branche lexicologique de cette théorie : la *Lexicologie explicative et combinatoire*.

La TST a vu le jour en Russie dans les années soixante¹² et s'est développée depuis en Russie, en Europe et au Canada, notamment à travers les nombreux travaux d'Igor Mel'čuk. Cette théorie linguistique, issue des courants structuraliste et fonctionnaliste, vise la construction de modèles (*modèles Sens-Texte*) permettant de rendre compte de l'association que tout locuteur d'une langue L est capable de faire entre un sens donné de L et l'ensemble des énoncés paraphrastiques de L exprimant ce sens, le tout dans une perspective de traitement automatique des langues.

L'appareillage terminologique et formel de la TST, très riche et sophistiqué, permet une modélisation de cette association Sens-Texte à travers les différents niveaux de fonctionnement des langues : sémantique, syntaxique, morphologique et phonologique. La correspondance Sens ↔ Texte est préférablement décrite dans le sens de la synthèse, c'est-à-dire du sens vers le

¹² L'article de Žolkovskij et Mel'čuk (1967) est souvent cité comme le premier texte publié dans cette théorie.

texte, puisque la production est une activité « plus linguistique » que la compréhension dans la mesure où elle n'implique que des connaissances purement linguistiques et écarte le problème complexe de la désambiguïisation inhérent à toute tâche de compréhension (Polguère, 1998 ; Milićević, 2006). Tous les niveaux de description, à l'exception du niveau sémantique, se divisent en deux : un niveau profond, plus près du sens, et un niveau de surface, plus près du texte. La figure suivante résume les niveaux de description de la TST.

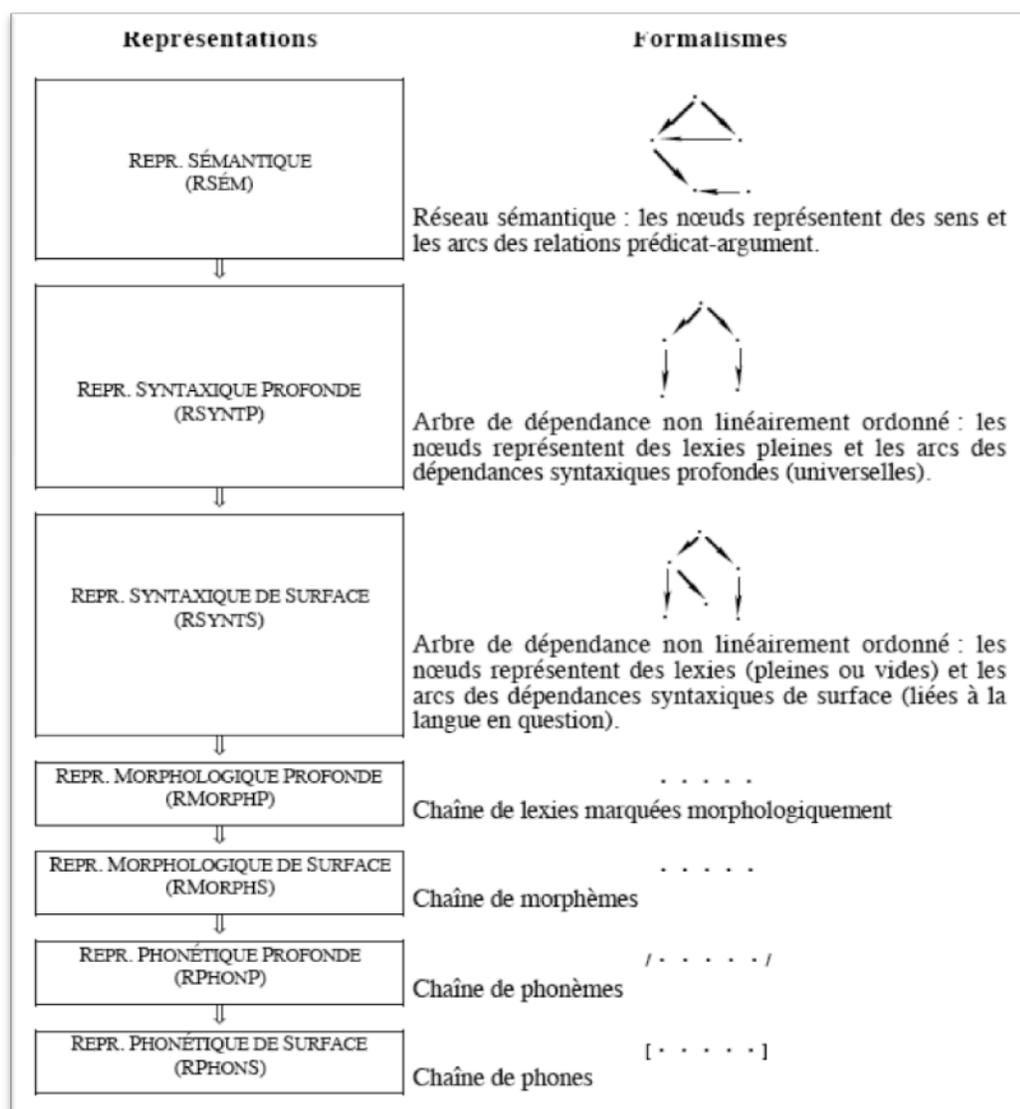


Figure 1. Niveaux de représentation et formalismes correspondants dans un modèle Sens-Texte (Polguère, 1998: 14)

Bien entendu, notre propos n'est pas ici de détailler l'ensemble des niveaux de description utilisés par la TST. Nous pouvons toutefois souligner que, contrairement aux modèles générativistes, qui constituent sans contredit le courant dominant en linguistique, la TST prend le sens (et non la syntaxe) comme point de départ de sa description linguistique ; par ailleurs, la syntaxe mise de l'avant par la TST est une syntaxe de dépendance inspirée des travaux de Lucien Tesnière (voir Tesnière, 1959), un des précurseurs du structuralisme en linguistique. Mentionnons aussi que les travaux menés jusqu'à maintenant dans le cadre de la TST se sont surtout concentrés sur les niveaux sémantique, syntaxique et morphologique du modèle. Pour une introduction plus détaillée à la TST, nous vous renvoyons aux textes de Mel'čuk (1997), de Polguère (1998) et de Milićević (2006).

1.1 LA LEXICOLOGIE EXPLICATIVE ET COMBINATOIRE

Comme nous l'avons mentionné plus haut, c'est surtout la *Lexicologie explicative et combinatoire* (LEC), composante lexicologique de la TST, qui nous servira d'assise théorique dans le cadre de cette thèse. La TST plaçant le sens au cœur de sa description linguistique, il est normal que la description des « mots », principaux éléments porteurs de sens de la langue, y occupe une place de choix : « [le] “linguiste Sens-Texte” est donc nécessairement un lexicographe » (Polguère, 1998: 24).

La LEC est une méthode de description rigoureuse des unités lexicales¹³ qui s'actualise dans un type de dictionnaire bien particulier : un dictionnaire explicatif et combinatoire (DEC). Quatre volumes de ce type de dictionnaire ont déjà été publiés pour le français sous le titre *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques* (Mel'čuk, Arbatchesky-Jumarie, Elnitsky, Iordanskaja et Lessard, 1984 ; Mel'čuk, Arbatchesky-Jumarie, Dagenais, Elnitsky, Iordanskaja, Lefebvre et Mantha,

¹³ L'emploi du terme *unité lexicale* à la place de *mot* sera justifié plus loin. Voir point 3.

1988 ; Mel'čuk, Arbatchewsky-Jumarie, Iordanskaja et Mantha, 1992 ; Mel'čuk, Arbatchewsky-Jumarie, Iordanskaja, Mantha et Polguère, 1999).

Le nom même de la LEC rend compte des spécificités de ce type de lexicologie. La LEC est *explicative*, car elle utilise la définition analytique, qui consiste en une « explication » du sens de l'unité lexicale par une décomposition de ce sens en termes de sens plus simples; cette façon de définir sera détaillée à la section 3.5. La LEC est *combinatoire* dans la mesure où elle

décrit la COMBINATOIRE SYNTAXIQUE ET LEXICALE associée à chaque unité lexicale de la langue. Par combinatoire syntaxique, on entend les phénomènes habituellement désignés en linguistique par les termes de sous-catégorisation ou régime. Par combinatoire lexicale, on entend les collocations contrôlées par l'unité lexicale en question. (Polguère, 1998: 25)

La combinatoire lexicale est décrite dans un DEC à l'aide d'un outil descriptif puissant : les fonctions lexicales (voir 3.6.2.1). Ces fonctions permettent de rendre compte des liens lexicaux qui existent à la fois sur l'axe paradigmatique – c'est-à-dire qui connectent l'unité lexicale décrite à d'autres avec lesquelles elle est liée sémantiquement (synonymes, antonymes, etc.) – et sur l'axe syntagmatique – c'est-à-dire les liens qui connectent l'unité lexicale décrite à celles avec lesquelles elle est susceptible d'entrer en cooccurrence (cf. la collocation).

Une telle approche lexicographique permet donc une description exhaustive des unités lexicales en mettant en lumière l'ensemble des informations qu'implique la maîtrise d'un mot, ce qui sera bien entendu extrêmement précieux pour délimiter la notion d'erreur lexicale. Pour une présentation détaillée de la LEC et du formalisme des fonctions lexicales, voir Mel'čuk, Clas et Polguère (1995).

1.2 JUSTIFICATION DE NOTRE POSITIONNEMENT THÉORIQUE

Malgré son appareillage conceptuel élaboré, sa cohérence et l'intérêt grandissant qu'elle suscite chez les linguistes – et dans une moindre mesure chez les didacticiens – la TST demeure relativement marginale (Milićević, 2006), notamment par rapport à la grammaire générative, qui a d'ailleurs beaucoup

influencé l'enseignement du français au cours des dernières décennies. Voici les raisons qui nous poussent à inscrire notre réflexion sur l'erreur lexicale dans ce courant de recherche.

Tout d'abord – nous l'avons déjà mentionné et le verrons plus en détail dans les prochaines sections –, la TST propose une caractérisation très claire et exhaustive de la notion de « mot », notion cruciale pour toute réflexion sur l'erreur lexicale; la description lexicographique constitue la clé de voute de cette approche linguistique, contrairement à d'autres théories, comme la grammaire générative, dans lesquelles la place centrale est occupée par la syntaxe. Il faut par ailleurs mentionner que les concepts lexicaux auxquels nous ferons appel dans cette thèse ne sont pas exclusifs à la TST; la plupart sont des concepts généralement admis en linguistique, mais dont les définitions sont parfois élastiques, ce qui peut conduire à une certaine confusion, d'où la nécessité de choisir un cadre théorique précis. L'intérêt de la TST pour nous est qu'elle fournit des définitions claires et non ambiguës de l'ensemble des concepts linguistiques essentiels à notre travail et qu'elle les ordonne dans un système cohérent où les liens qui les unissent sont mis en évidence. Elle associe aussi à ce réseau conceptuel une terminologie précise, qui permet d'atteindre un niveau d'abstraction favorisant la distanciation nécessaire à toute réflexion « méta » sur la langue. Par ailleurs, les formalismes puissants et la rigoureuse méthode de description que propose la LEC constituent quant à eux des outils formidables pour la description lexicographique qui, nous en sommes convaincu, se révéleront précieux pour aborder la tâche difficile de description d'erreurs linguistiques.

D'un point de vue plus pragmatique, il faut souligner que notre participation à divers projets de recherche liés à la LEC (travail sur le LAF¹⁴, sur le dictionnaire

¹⁴ Le *Lexique actif du français* (LAF) est un dictionnaire spécialisé et formalisé qui se concentre sur les phénomènes de dérivation sémantique et de collocation dans le but d'aider ses utilisateurs à encoder linguistiquement le message qu'ils veulent véhiculer. Un premier volume de ce dictionnaire, qui constitue en quelque sorte une version plus accessible du DiCo, a été publié en 2007 (Mel'čuk et Polguère, 2007).

DiCo¹⁵, développement d'une ontologie des étiquettes sémantiques) et nos tâches d'assistantat d'enseignement auprès d'Igor Mel'čuk et d'Alain Polguère nous ont permis d'approfondir notre connaissance de la TST et d'acquérir une certaine aisance relativement aux différents formalismes qu'elle propose. Le fait d'évoluer dans un « environnement Sens-Texte » durant quelques années nous a permis de constater la cohérence et la pertinence de cette théorie, de sorte que nous appréhendons naturellement les phénomènes lexicaux de ce point de vue et qu'il nous semble essentiel de tester cette approche dans un contexte plus didactique.

Finalement, la TST est citée par plusieurs auteurs comme une théorie linguistique tout à fait pertinente pour l'enseignement du lexique (Martin, 1984 ; Picoche, 1992 ; Cauchon, 2003 ; Tremblay, 2003 ; Mel'čuk et Polguère, 2007 ; Tremblay, 2009). En en faisant la toile de fond de notre thèse de doctorat, nous ajoutons notre voix à la leur et espérons contribuer à l'affirmation du potentiel didactique de ce cadre théorique.

2. LEXIQUE ET VOCABULAIRE

Une confusion assez répandue entre les termes *lexique* et *vocabulaire* nous pousse à croire qu'ils sont synonymes. Pourtant, il existe entre eux une distinction importante.

Le lexique d'une langue est l'ensemble des mots qui composent cette langue. De façon simpliste, nous pourrions affirmer qu'un dictionnaire monolingue comme le *Petit Robert* contient le lexique d'une langue. Bien entendu, une telle affirmation est fautive dans la mesure où le lexique d'une langue est une entité théorique qui ne peut être décrite avec exhaustivité. En effet, où s'arrêter lorsqu'on dresse la liste des mots d'une langue? Que faire des termes spécialisés propres à certaines disciplines? Doit-on noter les termes vulgaires et familiers, les emprunts à d'autres langues comme les anglicismes, les régionalismes? Les

¹⁵ Pour une présentation du projet DiCo, voir le site Web de l'Observatoire de Linguistique Sens-Texte (OLST) à l'adresse <http://idefix.ling.umontreal.ca/recherche/linguistique/dico/lang-pref/fr/>.

expressions doivent-elles aussi figurer dans notre liste? Ces quelques questions nous permettent de prendre conscience de l'impossibilité de décrire le lexique d'une langue dans son entièreté. Pour les mêmes raisons, il sera impossible pour une seule personne de connaître l'ensemble des mots constituant le lexique d'une langue.

Pour désigner l'ensemble des mots connus par un individu, nous utiliserons plutôt le terme *vocabulaire*, qui nous servira aussi à désigner l'ensemble des mots contenu dans un texte. Le vocabulaire d'une personne ou d'un texte est donc le sous-ensemble du lexique que connaît cette personne ou que contient ce texte. On pourra donc dire qu'une personne possède un vocabulaire riche, mais pas que son lexique est étendu.

3. QU'EST-CE QU'UN MOT?

Nous venons d'affirmer que le lexique d'une langue est l'ensemble des mots qui la composent. Pour que cette définition soit tout à fait claire, il faudrait aussi proposer une définition de ce qu'est un mot, ce qui n'est pas une mince affaire! Cette « notion de sens commun », omniprésente tout au long de notre apprentissage de l'écriture, se laisse difficilement définir et pose certains problèmes dans une perspective scientifique (Béguelin, 2000: 37). En effet, ce terme tout simple que nous utilisons quotidiennement renvoie en fait à différents concepts linguistiques. Pour s'en convaincre, il suffit d'observer les phrases suivantes :

- (1) La phrase *Elle écoute du jazz* contient quatre mots.
- (2) Les mots *vais* et *irez* sont deux formes du mot *aller*.
- (3) Le mot *cheval d'arçons* désigne un appareil de gymnastique et s'écrit en trois mots.

À la lecture de ces phrases, on s'aperçoit que nous désignons par le terme *mot* différentes réalités : tantôt diverses formes graphiques séparées par des espaces (phrase 1), tantôt un concept plus abstrait qui se réalise sous plusieurs formes (phrase 2), tantôt des expressions (phrase 3). Bien entendu, la linguistique, science qui étudie entre autres les mots, ne pouvait tolérer une telle ambiguïté et a proposé des termes plus précis pour distinguer chacun des concepts que nous

désignons communément par *mot*. Explorons maintenant ces concepts en prenant comme porte d'entrée la notion de signe linguistique.

3.1 LE SIGNE LINGUISTIQUE

Saussure (1960) a défini le signe linguistique comme une association indissociable entre un signifiant (une forme) et un signifié (un sens). Chaque signe linguistique véhicule donc un sens, et un seul, auquel nous avons accès par l'intermédiaire de sa forme, qui est une chaîne de phonèmes à l'oral ou une suite de caractères à l'écrit. Comme nous présentons ici divers concepts relatifs au lexique, nous nous intéresserons principalement aux signes linguistiques lexicaux¹⁶, que nous appellerons mots-formes. Le signifiant des mots-formes sera noté en italique alors que leur sens (ou signifié) apparaîtra entre guillemets simples, appelés guillemets sémantiques. Par exemple, on pourra dire que le mot-forme *drapeau* véhicule le sens 'drapeau'. La phrase (1) présentée plus haut aurait donc pu se lire comme suit : La phrase *Elle écoute du jazz*. contient quatre mots-formes.

Cette conception bidimensionnelle du signe linguistique, telle que proposée par Saussure, a été modifiée par certaines approches linguistiques plus contemporaines. La Théorie Sens-Texte (TST), notamment, conçoit plutôt le signe linguistique comme une entité linguistique possédant trois dimensions : un signifiant, un signifié et un syntactique (Mel'čuk, 1997). Le syntactique d'un signe linguistique – que nous désignerons par *combinatoire restreinte* – est un ensemble de propriétés, appelées propriétés de combinatoire, qui régissent en quelque sorte le comportement de ce signe dans la langue. Polguère (2003: 37) définit les propriétés de combinatoire comme les « contraintes propres à [un] signe qui limitent sa capacité à se combiner avec d'autres signes linguistiques et qui ne peuvent être déduites ni de son signifié, ni de son signifiant ». En d'autres mots, les propriétés de combinatoire sont les informations que l'on doit

¹⁶ Il existe en effet d'autres types de signes linguistiques. Par exemple, le morphème « -s » dans *carottes* est un signe linguistique à part entière, puisqu'il est une association entre une forme (-s) et un sens ('pluriel'). Par contre, l'étude de ce type de signes relève plutôt de la morphologie et n'est pas essentielle pour les concepts que nous voulons présenter ici.

connaître à propos d'un signe linguistique pour bien pouvoir l'utiliser. Par exemple, pour utiliser correctement un nom dans une phrase, il faut savoir qu'il fait partie de la classe des noms, mais aussi s'il est masculin ou féminin, s'il est invariable, s'il appartient au registre familier, etc. Nous reviendrons plus tard aux propriétés de combinatoire, car ces dernières sont cruciales dans la maîtrise du vocabulaire et influenceront grandement notre conception de l'erreur lexicale.

3.2 LE LEXÈME

À la lumière de ce qui a été dit plus haut, nous pouvons donc affirmer que les termes *vais* et *irez* de la phrase (2) sont des mots-formes, puisqu'ils présentent une association entre une forme et un sens. Pourtant, nous sentons bien que ces deux mots-formes entretiennent un lien privilégié, puisqu'ils sont deux formes fléchies du même verbe ALLER. Nous appelons lexème cette entité linguistique abstraite qui regroupe tous les mots-formes partageant le même sens et ne se distinguant que par la flexion. Les lexèmes seront notés en petites majuscules. Ainsi, nous dirons que le lexème ALLER regroupe les mots-formes *vais*, *irez*, *va*, *allons*, *iraient*, *allâtes* et toutes les autres formes conjuguées de ce verbe. De la même façon, le lexème DOUILLET regroupe les mots-formes *douillet*, *douillets*, *douillette* et *douillettes*, alors que le lexème ABRI ne correspond qu'à deux mots-formes : *abri* et *abris*.

La phrase (2) présentée plus haut aurait donc pu se lire comme suit : Les mots-formes *vais* et *irez* sont deux formes du lexème ALLER.

3.3 LA LOCUTION

Nous avons vu plus haut qu'un signe linguistique correspond à une triple association entre une forme, un sens et une combinatoire restreinte. En lisant la phrase (3) avec cette définition en tête, nous pourrions rester un peu perplexe. En effet, que faire de *cheval d'arçons*? Il s'agit bien de trois mots-formes, qui pourtant ne correspondent pas à trois sens, mais bien à un seul. En effet, pour connaître le sens de l'expression *cheval d'arçons*, il ne suffit pas de connaître les

sens respectifs des mots-formes qui la composent. Ce phénomène très fréquent s'appelle la non-compositionalité sémantique : le sens d'une expression ne correspond pas à la somme des sens des mots-formes qui la composent ('cheval d'arçons' ≠ 'cheval'+ 'de'+ 'arçons').¹⁷

Ce type d'expressions sémantiquement non compositionnelles se nomment expressions idiomatiques ou locutions. Comme ces locutions font aussi partie du vocabulaire des locuteurs, nous les considérerons aussi comme des éléments du lexique d'une langue. De la même façon que les lexèmes pouvaient correspondre à plusieurs mots-formes, les locutions pourront regrouper diverses constructions linguistiques qui se distinguent par la flexion. Par exemple, la locution POMME D'ADAM regroupe les constructions linguistiques *pomme d'Adam* et *pommes d'Adam* tandis que la locution PRENDRE EN GRIPPE correspond aux constructions *prends en grippe*, *prendrons en grippe*, *ont pris en grippe*, *aurez pris en grippe* ainsi que toutes ses formes conjuguées.

La phrase (3) pourrait donc se récrire comme suit : La locution CHEVAL D'ARÇONS désigne un appareil de gymnastique et s'écrit en trois mots-formes.

3.4 L'UNITÉ LEXICALE (UL)

Nous avons donc vu deux types d'entités linguistiques, qui se manifestent dans la parole sous forme de mots-formes ou de constructions linguistiques : les lexèmes et les locutions. Ces entités linguistiques sont les éléments qui constituent le lexique d'une langue, c'est pourquoi nous les désignerons par l'expression (ou locution!) « unité lexicale », qui a aussi pour synonyme le terme « lexie ». Voici la définition d'unité lexicale que propose Polguère (2008: 58).

¹⁷ On dira qu'un énoncé est sémantiquement compositionnel lorsque le sens qu'il véhicule est directement « calculable » par l'addition des sens des mots-formes qui le composent. Par exemple, la phrase *Elle adore la plongée.* est sémantiquement compositionnelle, mais pas la phrase *Elle grimpe dans les rideaux.* au sens de 'elle s'emporte'. Par contre, au sens littéral, par exemple en parlant d'une araignée, la phrase *Elle grimpe dans les rideaux.* serait sémantiquement compositionnelle.

Une **lexie**, aussi appelée **unité lexicale**, est soit un **lexème** soit une **locution**.

Chaque lexie (lexème ou locution) est associée à un sens donné, que l'on retrouve dans le signifié de chacun des mots-formes ou syntagme figé par lesquels elle s'exprime.

Ce survol des différentes réalités linguistiques désignées par le terme MOT nous permet de comprendre pourquoi nous préfererons utiliser des termes plus précis et moins ambigus. Ceci nous mènera aussi à revoir la définition du lexique que nous avons énoncée plus haut. Nous dirons donc maintenant que le lexique d'une langue est l'ensemble des unités lexicales (lexèmes ou locutions) qui composent cette langue.

3.5 LE SENS

Le sens d'une UL¹⁸ est sa propriété centrale, puisque les « mots » existent précisément pour véhiculer un sens. Pourtant, le sens est quelque chose de très complexe à définir.

La façon traditionnelle de décrire le sens d'une UL est d'en fournir une définition, c'est-à-dire une paraphrase constituée de termes plus simples que celui qu'elle définit et qui véhicule le même sens. Par exemple, si nous prenons la définition de PEIGNE proposée par *Le Petit Robert* (2002), 'instrument à dents fines et serrées qui sert à démêler et à lisser la chevelure', nous constatons que cette paraphrase rend compte du sens de l'UL PEIGNE en le décomposant en sens plus simples. Nous remarquons aussi que cette définition, de type analytique, se divise en deux grandes composantes : le genre prochain et les différences spécifiques.

Le genre prochain est la composante de sens centrale de la définition, qui constitue « en quelque sorte une paraphrase approximative minimale du sens de la lexie définie » (Polguère, 2008: 182) et qui correspond souvent à un hyperonyme de l'UL. Le genre prochain choisi dans la définition présentée plus

¹⁸ Nous utiliserons dorénavant l'abréviation UL pour « unité lexicale ».

haut est 'instrument' et agit comme composante centrale de 'peigne', qui est en effet un type d'instrument.

Les différences spécifiques, quant à elles, sont les informations qui caractérisent le genre prochain de sorte qu'il corresponde au sens de l'UL décrite en le différenciant des autres lexies qui partageraient le même genre prochain. Les différences spécifiques sont souvent nombreuses. Dans l'exemple donné plus haut, les composantes de sens 'à dents fines et serrées' et 'qui sert à démêler et à lisser la chevelure' permettent de restreindre le sens du genre prochain 'instrument' de façon à le faire correspondre au sens véhiculé par PEIGNE et à le distinguer des autres lexies dénotant des instruments.

Pour décrire le sens de certaines UL, appelées « unités lexicales prédicatives », un autre type d'informations sera nécessaire : le typage des actants. Certaines UL dénotant des faits ont en effet un sens liant qui implique les participants du fait en question, appelés actants sémantiques (Polguère, 2008). C'est le cas des verbes, qui impliquent toujours la présence d'au moins un actant correspondant syntaxiquement au sujet, mais aussi des adjectifs, des adverbes et de certains noms. Ainsi, à DORMIR correspond un actant ('X dort'), à REGARDER deux actants ('X regarde Y'), VIEUX implique un sujet ('X est vieux'), LENTEMENT implique un procès en déroulement ('X se déroule lentement'), NAISSANCE un actant ('naissance de X'). Les différents actants, représentés ici par des variables, sont exprimés syntaxiquement dans la phrase par des UL qui apparaissent avec l'UL prédicative qui les contrôle. La raison pour laquelle il est important d'en tenir compte à l'intérieur de la définition est que, parfois, des contraintes sémantiques pèsent sur les différents participants. Par exemple, le verbe ENDOMMAGER accepte comme complément une UL dénotant un objet physique (*La collision a endommagé la voiture.*), mais pas un être vivant (**La collision a endommagé le conducteur.*)¹⁹. Bien qu'elles soient rarement spécifiées explicitement par les dictionnaires, il est

¹⁹ Il est à noter que l'astérisque note ici le caractère asémantique de l'énoncé; bien que la phrase présentée soit syntaxiquement correcte, elle est défailante sur le plan sémantique.

essentiel de bien connaître ces contraintes pour utiliser correctement une UL prédicative.

Finalement, certaines UL véhiculent une connotation, qui « est un contenu informationnel associé à une lexie qui, contrairement au sens, n'est pas nécessairement exprimé quand cette lexie est utilisée. » (Polguère, 2008: 129) Par exemple, l'emploi du verbe *AVOUEUR* dans la phrase *Il a avoué son homosexualité*. permet à la personne qui formule cette phrase de prendre position de façon insidieuse face au phénomène de l'homosexualité, puisque *AVOUEUR* véhicule une connotation négative dans un tel contexte. Encore une fois, il est important de connaître le poids connotatif des UL que l'on emploie afin d'éviter que ce contenu entre en conflit avec le message véhiculé et le contexte d'énonciation.

3.6 LA COMBINATOIRE RESTREINTE

Comme nous l'avons mentionné plus haut, la connaissance de la forme et du sens d'une UL n'est pas suffisante pour pouvoir utiliser correctement cette dernière. En effet, de nombreuses autres informations doivent être connues pour assurer la maîtrise d'une lexie : les propriétés de combinatoire. Nous pouvons distinguer deux grands axes dans la combinatoire restreinte : la combinatoire grammaticale²⁰ et la combinatoire lexicale.

3.6.1 LA COMBINATOIRE GRAMMATICALE

La combinatoire grammaticale regroupe l'ensemble des propriétés qui régissent le comportement morphologique et syntaxique d'une UL. Par exemple, la classe de mots et le genre nominal sont des propriétés de combinatoire grammaticale. Des cas un peu plus marginaux seraient l'invariabilité (ex. *OBSÈQUES*, toujours pluriel), le fait pour un verbe d'être défectif, c'est-à-dire de ne pas se conjuguer à tous les temps (ex. *PAITRE*, *FRIRE*) et le fait pour une UL de conditionner l'emploi du subjonctif (ex. *bien que je sois* et non **bien que je suis*).

²⁰ Si nous choisissons de parler de *combinatoire grammaticale* plutôt que de *combinatoire syntaxique*, comme dans notre présentation de la LEC (section 1.1), c'est que nous incluons sous cet axe de combinatoire autant des propriétés purement syntaxiques que des propriétés morphologiques.

Parmi les propriétés de combinatoire grammaticale se retrouvent aussi celles relatives au comportement syntaxique de la lexie²¹; c'est ce que nous appelons le régime de l'UL. Le régime d'une UL est en fait la façon d'exprimer syntaxiquement les actants sémantiques que nous avons déjà définis. Ces actants sémantiques, qui dans une phrase jouent souvent le rôle de complément d'une lexie prédicative²², sont parfois obligatoires, parfois facultatifs, et peuvent être introduits soit directement, soit par une préposition. Par exemple, le verbe ENTAMER exige la présence d'un complément direct (transitif direct), alors que le verbe DOUTER introduit son complément à l'aide de la préposition DE (transitif indirect). La notion de transitivité est propre au régime verbal, mais des UL prédicatives appartenant à d'autres classes de mots régissent aussi des prépositions; *fier de qqch.*, *acquiescement de qqn à qqch.*, *enclin à qqch.*, etc. Encore une fois, le régime d'une UL est une de ses propriétés inhérentes et il est nécessaire de le maîtriser pour bien utiliser la lexie.

3.6.2 LA COMBINATOIRE LEXICALE

Le deuxième axe de combinatoire est la combinatoire lexicale, aussi appelée cooccurrence lexicale restreinte, qui se manifeste dans la parole à travers le phénomène de la collocation. La cooccurrence lexicale est le simple fait pour deux UL d'apparaître ensemble dans un énoncé. La collocation, quant à elle, est un cas particulier de cooccurrence. En voici une définition proposée par Polguère (2008: 165) :

*Une **collocation** est un syntagme AB (ou BA), formé des lexies L_A et L_B et qui est tel que, pour le construire, le locuteur sélectionne L_A librement d'après son sens (L_A), alors qu'il sélectionne L_B pour exprimer auprès de L_A un sens très général (s) en fonction de la combinatoire restreinte de L_A .*

*Les collocations sont donc des **syntagmes semi-phraséologisés**.*

L'expression *faire un choix*, par exemple, est une collocation. La personne utilisant cette expression a d'abord voulu exprimer le sens 'choix' et, en fonction de ce

²¹ Lépine (1995) parle d'ailleurs de syntaxe lexicale.

²² C'est notamment le cas avec plusieurs noms, tous les verbes transitifs, certains adjectifs et l'ensemble des prépositions. Dans le cas des lexies verbales, le premier actant sémantique joue syntaxiquement le rôle de sujet et non de complément.

choix, elle a sélectionné le verbe FAIRE. Nous dirons que CHOIX est la base de la collocation, alors que FAIRE est le collocatif. Il est important de comprendre que dans une telle expression, seule la base est choisie librement, et le collocatif est en quelque sorte imposé par le système lexical de la langue, sans que l'on puisse nécessairement justifier pourquoi. En effet, si la personne avait plutôt choisi pour base le nom DÉCISION, qui est pourtant un synonyme de CHOIX, elle aurait dû utiliser le collocatif PRENDRE pour produire *prendre une décision* (et non **faire une décision* ou **prendre un choix*). La collocation est un phénomène complexe²³, mais pourtant omniprésent dans toutes les langues. Par conséquent, il est important pour bien connaître une UL de savoir quels sont ses collocatifs les plus fréquents. Par exemple, une personne qui maîtrise bien le verbe DORMIR saura que l'on peut dire, pour intensifier le sens du verbe, *dormir profondément/à poings fermés/comme un bébé/comme une souche/du sommeil du juste*, etc.

On se rend peu compte de la fréquence des collocations dans la langue, et le peu d'intérêt qui leur est généralement accordé dans l'enseignement en témoigne. Pourtant, les problèmes relatifs à la combinatoire lexicale sont assez fréquents, notamment en langue seconde, puisque la richesse collocationnelle – c'est-à-dire le fait de connaître plusieurs collocatifs pour une UL spécifique – demeure souvent l'apanage du locuteur natif. Les dictionnaires grand public, ressource lexicale par excellence, font généralement un traitement incomplet et inégal de la collocation en présentant implicitement, à travers des exemples d'utilisation de l'UL décrite, seulement quelques collocatifs parmi les plus fréquents.

Pour sa part, la LEC tente de fournir une description exhaustive et structurée de la combinatoire des UL qu'elle décrit dans ses dictionnaires. Cette description fine de la dimension collocationnelle des UL est fournie à l'aide d'un formalisme rigoureux : les fonctions lexicales. Il nous est impossible, dans le cadre de cette thèse, de présenter en détail ce puissant formalisme, mais en voici tout de même un bref aperçu.

²³ Pour une présentation plus détaillée de la notion de collocation, voir Hausmann (1979) et Mel'čuk *et al.* (1995)

3.6.2.1 LA MODÉLISATION DES COLLOCATIONS À L'AIDE DES FONCTIONS LEXICALES

Les fonctions lexicales (ci-après FL), sur le modèle des fonctions mathématiques, permettent d'associer une UL à un ensemble de valeurs pour une fonction donnée. Prenons comme premier exemple une FL qui ne permet pas d'encoder un lien syntagmatique collocationnel, mais plutôt un lien paradigmatique : la FL *Syn*, qui permet d'associer une UL à ses synonymes.

Syn(difficile) = ardu, complexe, compliqué, corsé, délicat, problématique

La FL prend donc une UL précise comme mot-clé (ici DIFFICILE) et renvoie les valeurs correspondantes.

Pour que l'on considère pertinent d'encoder une certaine relation lexicale par une FL, il faut que cette relation soit récurrente dans la langue, c'est-à-dire qu'elle existe pour un nombre important d'UL, et qu'elle renvoie pour chaque UL des valeurs différentes, comme c'est bien sûr le cas de la synonymie. Bien évidemment, même les dictionnaires commerciaux rendent compte de certaines relations récurrentes, comme la synonymie et l'antonymie. Là où la LEC se distingue, c'est par l'emploi des FL pour illustrer des relations peu représentées dans les dictionnaires grand public, mais néanmoins récurrentes dans la langue, par exemple le nom typique des actants sémantiques associés à une UL. Cette relation entre une UL et ses actants sémantique est encodée par la FL *S*, pour « substantif ». En voici un exemple avec le nom ASSASSINAT (*assassinat de Y par X*).

$S_1(\textit{assassinat}) = \textit{assassin, meurtrier, auteur [de l'~]}$
 $S_2(\textit{assassinat}) = \textit{victime}$

Les numéros apparaissant en indice de la FL font référence à l'actant sémantique que concernent les valeurs retournées par la fonction ; ainsi, les termes ASSASSIN, MEURTIER et *auteur de l'assassinat* permettent de désigner le premier actant d'ASSASSINAT (X), alors que VICTIME dénote le second (Y).

Illustrons maintenant une FL permettant d'encoder une relation syntagmatique en nous servant de la FL *Magn*²⁴, qui véhicule le sens très vague 'très' et associe à une UL les collocatifs permettant de l'intensifier, comme dans l'exemple suivant.

²⁴ Du latin *magnus*.

Magn(*faute*) = *grave, lourde, grossière, irréparable, sérieuse*

Il est aussi possible de combiner les FL de façon à traduire d'autres relations.

AntiMagn(*faute*) = *petite, légère, vénielle*

Il existe environ une cinquantaine de FL standards dans la TST, servant à encoder tant des relations paradigmatisques que syntagmatiques, allant de l'intensification aux verbes supports. La LEC, en présentant pour chaque UL qu'elle décrit l'ensemble des FL pertinentes pour cette UL, met en lumière l'ensemble des liens privilégiés qu'entretient cette UL avec d'autres lexies de la langue, et permet d'anticiper les collocatifs avec lesquels elle est susceptible d'apparaître lorsqu'utilisée par un locuteur natif.

Nous arrêtons ici notre (trop) brève introduction à ce puissant formalisme de modélisation des collocations, en mentionnant toutefois que les FL utiles à notre analyse seront présentées plus en détail au chapitre VIII, point 2.2.5, alors que nous aborderons les erreurs de collocations des élèves. Pour une introduction complète à la notion de FL, nous renvoyons à Mel'čuk *et al.* (1995).

4. LA STRUCTURE DU LEXIQUE

La définition du lexique comme un ensemble d'UL peut nous entraîner à percevoir celui-ci comme une simple liste de mots, une nomenclature. Cette vision semble en outre confirmée par la façon dont les dictionnaires de langue présentent l'information lexicale. En effet, le dictionnaire – outil de description privilégié du lexique – fonctionne comme une nomenclature, c'est-à-dire que les mots y sont ordonnés alphabétiquement, sans autre forme de logique. Pourtant, le lexique n'est pas un inventaire de « mots » pêle-mêle et il est fort peu probable que notre lexique mental soit ordonné alphabétiquement! En effet, le lexique possède une structure et il est possible d'en rendre compte en effectuant des regroupements d'UL sur diverses bases. Nous présenterons différents liens permettant d'appréhender le lexique comme un réseau structuré d'unités lexicales.

4.1 LA POLYSÉMIE ET L'HOMONYMIE

Comme les sens à exprimer sont infinis, mais que les langues possèdent un nombre fini de phonèmes et ne sauraient s'encombrer de signifiants extrêmement longs et difficiles à mémoriser, un « principe général d'économie » (Huet, 1997: 120) fait qu'il est fréquent d'avoir un même signifiant pour plusieurs signifiés. En fait, la majorité des « mots » courants d'une langue ont plus d'un sens. Comme nous savons que chaque lexie est associée à un sens donné et un seul, nous pouvons donc déduire qu'il y aura plusieurs lexies partageant un signifiant identique. Deux cas de figure sont possibles, comme l'illustre la phrase (4).

(4) *Les mots FEU et AVOCAT ont plus d'un sens.*

Il existe en effet de nombreuses UL FEU en français (le *Petit Robert* en répertorie plus d'une vingtaine). Pour limiter notre explication, nous considérerons ici seulement trois UL :

FEU1 [*Le feu éclairait la pièce.*] : 'dégagement d'énergie calorifique et de lumière accompagnant la combustion vive'

FEU2 [*Laisser mijoter à feu doux.*] : 'Source de chaleur (à l'origine, foyer enflammé) dans la transformation des aliments'

FEU3 [*Le chauffard ne s'est pas arrêté au feu rouge.*] : 'Signal lumineux destiné à guider, à réglementer le déplacement des véhicules'

Bien que les trois UL FEU présentées correspondent à des sens bien distincts, il existe un lien sémantique évident entre les sens 2 et 3 et le sens premier. Ce cas de figure correspond au phénomène bien connu de **polysémie**. Nous dirons donc que ces trois UL font partie d'un même vocable FEU, qui est polysémique. Chacune des UL composant un **vocable polysémique** est appelée **acception** de ce vocable. La plupart des « mots » courants d'une langue sont en fait des vocables polysémiques, c'est-à-dire à plusieurs acceptions. Dans les dictionnaires, les différentes acceptions d'un vocable sont identifiées par des numéros à l'intérieur de l'article du vocable. Bien entendu, si la polysémie est un phénomène très fréquent, il existe aussi des vocables que nous qualifierons de **monosémiques**, c'est-à-dire qu'ils ne correspondent qu'à une seule UL ou, en

d'autres mots, qu'ils n'ont qu'une seule acception. Le vocable LÉVRIER, par exemple, est monosémique.

Le cas d'AVOCAT présent dans la phrase (4) est un peu différent. En effet, il existe aussi en français différentes UL AVOCAT. En voici deux :

AVOCAT¹¹ [*Son avocat s'est retiré de la cause.*] : 'Personne qui, régulièrement inscrite à un barreau, conseille en matière juridique ou contentieuse, assiste et représente ses clients en justice'

AVOCAT²¹ [*Dominic est allergique aux avocats.*] : 'Fruit de l'avocatier, de la grosseur d'une poire, à peau verte ou violette, dont la chair a la consistance du beurre et un goût rappelant celui de l'artichaut'

Dans ce cas-ci, nous ne dirons pas que ces deux UL font partie d'un même vocable, puisqu'il n'existe entre elles aucun lien sémantique évident. Il s'agit ici d'un cas typique d'**homonymie**, c'est-à-dire une relation entre deux vocables ayant un signifiant identique, mais sans partager de lien de sens. Dans le dictionnaire, des vocables homonymes auront des entrées distinctes. Nous disons que l'homonymie est une relation entre des vocables, puisqu'il se peut tout à fait que des vocables homonymes soient aussi polysémiques. Par exemple, le vocable AVOCAT¹ présenté plus haut contient une autre acception métaphorique :

AVOCAT¹² [*Il s'est fait l'avocat de cette idée.*] : 'Toute personne qui défend une cause, une idée, qui intercède pour une autre personne'

En résumé, la polysémie est une propriété d'un vocable possédant plusieurs acceptions liées sémantiquement, alors que l'homonymie est une relation entre différents vocables partageant une même forme, mais ne présentant aucun lien de sens évident. Évidemment, c'est la polysémie que nous présentons ici comme un phénomène structurant du système lexical, l'homonymie étant en quelque sorte un « accident ».

Dernier fait à noter : l'homonymie peut se manifester seulement à l'oral (homophonie) ou à la fois à l'oral et à l'écrit (homographie). Par exemple, ÇA et SA sont des homophones, mais pas des homographes, alors que les cas étudiés jusqu'à maintenant étaient tous des cas d'homographie, qui intéressent

davantage le lexicographe. Le phénomène de la paronymie s'apparente un peu à celui de l'homonymie, dans la mesure où deux paronymes partagent une forme (sonore ou écrite) très similaire, bien qu'elle ne soit pas identique, mais entre lesquelles il y a absence de liens sémantiques (ex. LAGUNE et LACUNE).

4.2 LES CLASSES DE MOTS²⁵

Si l'on nous demandait spontanément d'effectuer des regroupements entre des UL, notre premier réflexe serait sans doute de les regrouper selon les classes de mots auxquelles elles appartiennent : les noms avec les noms, les verbes avec les verbes et ainsi de suite. Il y aurait donc autant de regroupements que de classes de mots : verbe, nom, adjectif, adverbe, pronom, déterminant, conjonction, préposition. Dans un tel classement, on retrouve deux grands types de classes : les classes ouvertes et les classes fermées. Les classes ouvertes sont celles qui contiennent le plus grand nombre d'UL; ce sont les verbes, les noms, les adjectifs et les adverbes. Ces classes sont qualifiées d'« ouvertes » parce qu'elles sont susceptibles de voir apparaître de nouvelles UL et donc de croître en nombre. Pensons par exemple aux créations lexicales qui sont apparues récemment dans la langue : COURRIEL, TÉLÉCOPIER, CLAVARDAGE, GOOGLER, etc. Il serait très étonnant de voir apparaître de la même façon un nouveau pronom ou une nouvelle préposition; c'est pourquoi nous parlerons plutôt de classes fermées dans le cas de ces classes de mots. Les classes fermées sont en quelque sorte « figées » et le fait d'y ajouter de nouvelles UL provoquerait de grands bouleversements, alors que des UL appartenant aux classes ouvertes font régulièrement leur apparition et s'intègrent tout naturellement dans la langue.

Ce classement par classes de mots est intéressant, mais il ne rend pas compte d'une véritable structuration du lexique; en effet, il serait surprenant que les regroupements par classes de mots constituent l'unique façon d'organiser notre lexique mental. La raison en est bien simple : de tels regroupements sont

²⁵ Bien que la littérature Sens-Texte privilégie l'emploi de *parties du discours* pour faire référence aux classes de mots, nous choisissons tout de même cette dernière appellation, qui correspond à l'usage dans les classes de français. Pour une discussion à propos des choix terminologiques associés aux classes de mots, voir Polguère (2008).

effectués non pas sur la base de critères sémantiques, mais à partir de critères d'ordre syntaxique. Voyons maintenant d'autres façons d'envisager la structure du lexique, qui cette fois s'appuient sur des critères sémantiques.

4.3 HIÉRARCHISATION SÉMANTIQUE

Un premier classement d'ordre sémantique pourrait reposer sur l'existence de relations hiérarchiques entre les UL. On peut penser par exemple à la relation entre les noms ANIMAL et MAMMIFÈRE : MAMMIFÈRE est « subordonné » à ANIMAL, il s'agit d'un hyponyme de ce nom. Inversement, ANIMAL est un hyperonyme de MAMMIFÈRE.

Deux termes sont en relation d'hyperonymie/hyponymie lorsque l'un des deux termes (l'hyperonyme) véhicule un sens moins riche qui est inclus dans le sens de l'autre (son hyponyme).²⁶ Par exemple, le sens d'ANIMAL est moins précis que celui de MAMMIFÈRE et ANIMAL constitue le genre prochain de MAMMIFÈRE (= '**animal** vertébré, dont les femelles allaitent leurs petits avec des mamelles...'). À l'inverse, on pourra dire que l'hyponyme est un type de son hyperonyme : le mammifère est un type d'animal. Évidemment, la relation d'hyperonymie ne fonctionne pas que sur deux niveaux et il est tout à fait possible que l'hyponyme d'une UL constitue l'hyperonyme d'une autre, comme l'illustre le schéma ci-dessous.

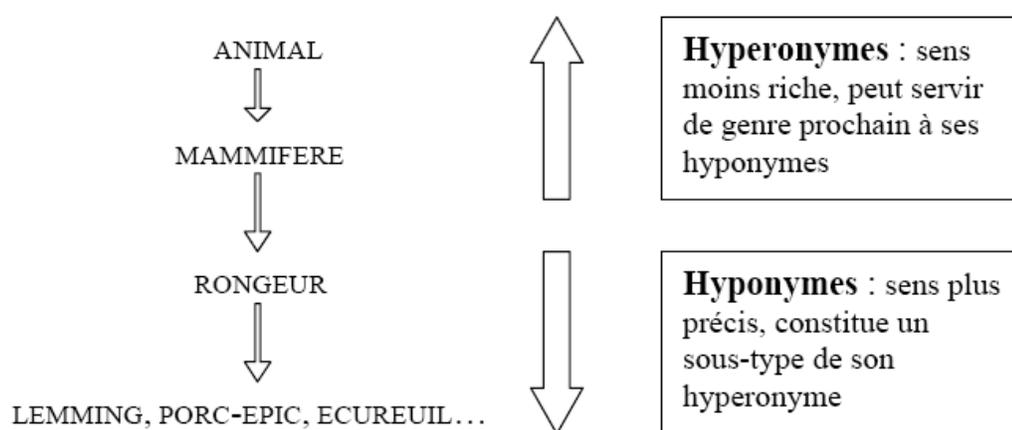


Figure 2. Exemple de hiérarchisation sémantique

²⁶ Dans le *Programme de formation de l'école québécoise*, on utilise plutôt la terminologie *terme générique* et *terme spécifique* pour désigner respectivement l'hyperonyme et l'hyponyme.

Ce regroupement est plus structurant que le premier, mais il comporte encore quelques inconvénients. D'abord, il existe en fonction d'un classement de nature plutôt référentielle; il se fonde souvent sur une connaissance du monde plutôt que sur des considérations de nature purement linguistique. Les UL sont forcément en relation avec les objets du monde qu'elles dénotent, mais le lexique est une entité linguistique et on peut penser qu'il a un fonctionnement interne. Enfin, ce classement, comme le précédent, fait intervenir des UL appartenant toutes à la même classe de mots, ce qui est quelque peu contraignant, bien que cette fois elles soient tout de même réunies sur la base de liens sémantiques.

4.4 LA DÉRIVATION LEXICALE

L'option que nous choisirons est de considérer le lexique comme un réseau où tout se tient. Comme le rappelle Polguère (2008), il suffit de « tirer » sur une UL pour voir apparaître à sa suite toute une série de lexies. La raison en est que le lexique se caractérise par deux propriétés structurantes : la polysémie et la connexion lexicale. Nous avons abordé plus haut la polysémie ainsi que le phénomène de la collocation, qui constitue un type de connexion lexicale; nous nous intéressons ici à un autre type de connexion lexicale : la dérivation.

La dérivation morphologique permet de créer de nouvelles UL par l'ajout d'affixes (préfixe ou suffixe), l'UL créée pouvant appartenir ou non à la même classe de mots que l'UL de base (ou *radical*). Ce phénomène morphologique, somme toute assez peu productif en français d'un point de vue synchronique²⁷, est celui qui sous-tend les « familles de mots » telles qu'on les enseigne à l'école. Ainsi, seuls les mots ayant une parenté morphologique font partie d'une même famille de mots; cette façon de voir les choses peut conduire à certaines incohérences. Par exemple, si COURIR et COURSE font clairement partie d'une même famille lexicale, la logique morphologique nous pousserait à dire qu'il n'en est pas de même de TOMBER et CHUTE, alors que le lien qui unit ces deux UL est tout à

²⁷ En effet, hormis quelques rares préfixes et suffixes, comme *pré-*, *dé-* ou *-able/-ible*, que l'on peut accoler à à peu près n'importe quel radical verbal pour créer de nouvelles UL verbales ou adjectivales, la dérivation morphologique en français n'est pas un procédé qui permet aux locuteurs de produire des créations lexicales de façon spontanée.

fait le même qu'entre les deux précédentes. Il est fort probable (et assez fréquent) qu'un lien de forme traduise l'existence d'un lien de sens, mais ce n'est pas parce qu'il n'existe pas de lien morphologique qu'il y a absence de lien sémantique! Et l'inverse est aussi vrai; une apparente ressemblance de forme n'implique pas nécessairement un lien de sens (par exemple, LUXURE, LUXE et LUXURIANCE n'ont rien à voir ensemble). Sans vouloir nier l'intérêt d'une étude morphologique des mots-formes, notamment dans une visée d'explication de l'évolution de la langue (perspective diachronique) ou de maîtrise de l'orthographe, nous croyons qu'une vision plus large de la dérivation nous permet de mieux saisir l'organisation du lexique.

Plutôt que de parler de dérivation morphologique, la TST parle de *dérivation sémantique*. La dérivation sémantique permet de mettre en lumière les relations de sens récurrentes entre les unités du système lexical, et ce, indépendamment de leur parenté morphologique. Voici les principaux liens de dérivation sémantique qui pourraient nous être utiles dans cette thèse.

4.4.1 SYNONYMIE

La synonymie est une relation qui unit deux UL appartenant à la même classe de mots et dont les sens sont semblables ou très rapprochés, comme CÉDER et VENDRE dans l'exemple suivant :

*Robert a **vendu** sa maison à Raymonde. → Robert a **cédé** sa maison à Raymonde.*

Nous ne définissons pas la synonymie comme une relation entre deux UL ayant des sens identiques, mais bien semblables, puisque la synonymie exacte est un phénomène plutôt rare dans la langue; nous ne retrouvons que quelques paires d'UL véhiculant exactement le même sens (VÉLO ~ BICYCLETTE, AUTOMOBILE ~ VOITURE, UNITÉ LEXICALE ~ LEXIE, etc.). La raison en est simple : il est somme toute assez peu économique d'avoir deux UL ayant exactement le même sens et les langues ne sauraient s'encombrer d'un grand nombre de paires comme celles présentées. Nous considérerons donc comme des synonymes des UL ayant des sens très rapprochés qui peuvent se substituer l'une à l'autre dans une grande variété de contextes; c'est souvent le cas, par exemple, entre une UL et un terme

plus vague pouvant lui servir d'hyperonyme (MARTEAU ~ OUTIL).

Une dernière remarque mérite d'être faite à propos de la synonymie : il s'agit d'une **relation entre deux UL**, et non entre deux vocables. Par conséquent, chacune des acceptions d'un vocable (les différentes UL le composant) aura des synonymes différents.

4.4.2 ANTONYMIE

L'antonymie est la relation qui unit deux UL ayant des sens opposés et appartenant à une même classe de mots (GRAND ~ PETIT, PRÉSENT ~ ABSENT, PROCHE ~ LOIN, MAJORITÉ ~ MINORITÉ, etc.); il s'agit donc du phénomène opposé à la synonymie. Tout comme la synonymie, l'antonymie regroupera des paires de termes qui seront aussi des antonymes exacts (comme ceux présentés plus haut), mais aussi des antonymes approximatifs. D'ailleurs, nous retrouvons dans les dictionnaires autant les antonymes exacts qu'approximatifs. Comme dans le cas de la synonymie, il est important de se rappeler que les différentes acceptions d'un vocable seront associées à des antonymes différents.

4.4.3 CONVERSIVITÉ

Un autre exemple de dérivation sémantique est la conversivité. Polguère (2008: 154) en donne la définition suivante :

*Deux lexies appartenant à la même partie du discours sont **conversives** si :*

- 1) ce sont soit des prédicats sémantiques dénotant une même situation, soit des quasi-prédicats sémantiques dénotant deux entités impliquées dans une même situation;*
- 2) ces (quasi-)prédicats s'expriment dans la phrase avec une inversion de l'ordre d'au moins deux de leurs actants.*

Par exemple, le verbe ACHETER est un conversif du verbe VENDRE. La situation dénotée par les deux verbes est exactement la même, seul l'ordre des participants sémantiques change. Malgré une première impression en ce sens, la conversivité n'est pas de l'antonymie. En effet, *prêter quelque chose à quelqu'un* ne signifie pas 'ne pas emprunter quelque chose à quelqu'un'. Il s'agit plutôt d'une forme de synonymie, dans laquelle seul le point de vue sur la situation dénotée diffère.

4.4.4 DÉRIVATION SÉMANTIQUE PAR UN CHANGEMENT DE CLASSE DE MOTS

Un cas typique de dérivation sémantique est le changement de classe de mots d'une UL donnée, changement qui n'affecte pas le sens véhiculé par les deux UL mises en rapport. Par exemple, DÉPART et PARTIR signifient exactement la même chose, mais ce ne sont pas des synonymes, car ils n'appartiennent pas à la même classe de mots. Évidemment, la dérivation sémantique ne se limite pas à la relation entre un verbe et un nom, mais peut aussi s'appliquer à d'autres classes de mots : FIER et FIERTÉ véhiculent le même sens (*Sophie est très fière de son fils.* ≈ *Sophie éprouve une grande fierté envers son fils.*), tout comme RAPIDE et RAPIDEMENT (*L'écoulement de l'eau est rapide.* ≈ *L'eau s'écoule rapidement.*).

Bien entendu, il ne faudrait pas croire que le changement de classe de mots implique nécessairement une parenté morphologique entre les deux UL. Par exemple, le nom qui est l'équivalent sémantique du verbe TOMBER est CHUTE; bien que rien ne relie ces deux lexies d'un point de vue formel, elles véhiculent pourtant un sens identique (*Dominic est tombé dans l'escalier.* ≈ *Dominic a fait une chute dans l'escalier.*). Voilà tout l'intérêt de la notion de *dérivation sémantique* : les liens qui unissent ces UL sont d'abord des liens sémantiques, qui peuvent parfois être accompagnés de liens morphologiques.

4.4.5 DÉRIVATION SÉMANTIQUE LIÉE À LA STRUCTURE PRÉDICATIVE

Les participants (ou actants) sémantiques d'une UL se réalisent parfois au niveau lexical. Un cas typique est la relation entre un verbe et son premier participant, la personne qui fait l'action dénotée par le verbe, qui s'exprime souvent morphologiquement (COURIR ~ COUREUR, DANSER ~ DANSEUR, ENSEIGNER ~ ENSEIGNANT). Par contre, il peut encore une fois arriver que ce type de lien ne soit pas marqué par une parenté morphologique. Par exemple, le nom typique du premier participant sémantique du verbe VENDRE au sens de 'dénoncer (qqn) en échange d'une récompense, d'un avantage' est TRAITRE (et non VENDEUR!).

CONCLUSION

Les dernières pages nous ont permis de présenter les principaux concepts qui nous serviront à caractériser la notion d'erreur lexicale et à envisager des scénarios d'explication pour les erreurs analysées. Maintenant qu'ont été posées nos assises théoriques, amorçons notre réflexion sur la norme et l'erreur.

III. LA NORME

Quiconque s'intéresse à l'erreur ne peut éviter une discussion à propos de la norme. En effet, l'erreur est souvent définie comme un écart par rapport à la norme, ou du moins à une norme, et un énoncé n'est considéré erroné que par son caractère déviant par rapport à l'usage normal. Mais au fait, qu'est-ce qu'un usage normal? Quelle est la norme qui nous servira de point de référence lorsque nous parlerons d'erreur? Qu'est-ce que la norme lexicale? Existe-t-il une norme propre au français du Québec? Voilà les questions auxquelles nous répondrons dans les sections qui suivent.

1. NORME ET VARIATION

Une langue est un système de signes conventionnel commun aux locuteurs d'une communauté linguistique dont l'efficacité repose en grande partie sur son caractère stable. Cette stabilité du système linguistique est assurée par des *invariants* qui « sous-tendent les *systèmes* de l'oral et de l'écrit » (Honveau, 1995: 53) et en constituent en quelque sorte le *noyau dur* (Corbeil, 1980 ; Chaudenson, Mougeon et Béniak, 1993). Par contre, au-delà de ces invariants, la langue est soumise à divers types de variations qui, sans constituer des obstacles à la communication (Honveau, 1995), exercent sur elle certaines pressions qui conduisent parfois à des changements. Les sources de variations peuvent être diverses (Polguère, 2003) et touchent principalement les dimensions phonologique et lexicale de la langue. Marquilló Larruy (2003: 46) identifie cinq types de variation externes.

1. La variation historique, qui intéresse l'étude diachronique de la langue;
2. La variation géographique ou topolectale (que l'auteure nomme « diatopique »);
3. La variation diastratique, qui est liée aux différents facteurs sociaux;
4. La variation diaphasique, qui concerne la dimension pragmatique de la communication;
5. La variation associée au canal de communication, oral ou écrit.

À ces sources de variation, nous pourrions ajouter celle liée à des domaines techniques spécifiques (Polguère, 2003); par exemple, les termes utilisés pour parler du corps humain au quotidien diffèrent de la terminologie propre à la médecine. Parallèlement aux sources de variations externes, il existe une variation interne, inhérente à la langue, liée à ses potentialités (Berrendonner, 1988 ; Gadet, 1989); cette variation interne fait en sorte que des structures concurrentes cohabitent dans le système linguistique (*un/une météorite, ce qui manque/ce qu'il manque, la montre que Jean a vendue/la montre qu'a vendue Jean, etc.*).

C'est donc dans la mesure où la langue est soumise à certaines variations qu'il est pertinent de s'intéresser à la question de la norme. Quels sont les usages, au-delà du noyau dur commun à tous les locuteurs d'une langue, qui constituent la norme? Quel est le statut des usages se situant dans les zones de variations du système? La section suivante nous fournira certains éléments de réponse.

2. LA PLURALITÉ DES NORMES

Toute discussion sérieuse à propos de la norme devra éviter les écueils liés à la polysémie des vocables NORMAL et USAGE. En fait, le simple fait de mettre en lumière la polysémie de ces termes étroitement liés à la norme permet de mieux comprendre les difficultés inhérentes à la définition de ce concept.

Verreault (1999) s'est intéressé aux différents sens d'USAGE, dont voici les trois principaux :

1. 'le fait d'utiliser un mot dans le discours' [*L'usage de ce mot n'est pas approprié dans ce contexte.*]. SYN. *emploi*.
2. 'manière dont la langue est utilisée dans un milieu, une situation, une époque déterminée' [*Ce terme vieilli est disparu de l'usage en France depuis des siècles.*].
3. 'manière correcte d'utiliser les mots d'une langue' [*L'usage suggère d'éviter la locution « suite à ».*]. SYN. *bon usage*.

La différence sémantique que l'on constate entre les sens 2 et 3 existe aussi entre les diverses acceptions de NORME. Rey (1972, 1983) fait ressortir une opposition semblable avec l'adjectif NORMAL, qui peut être entendu comme « tout ce qui est plus courant, plus habituel que quelque chose d'autre [ou comme] tout

ce qui est conforme à une règle, à un modèle donné » (Verreault, 1999). Toute entreprise de caractérisation du concept de norme demandera donc de tenir compte de cette opposition.

Avant toute tentative pour définir la « norme », la moindre considération lexicologique découvre derrière le terme deux concepts, l'un relevant de l'observation, l'autre de l'élaboration d'un système de valeurs, l'un correspondant à une situation objective et statistique, l'autre à un faisceau d'intentions subjectives. Le même mot, utilisé sans précaution, correspond à la fois à l'idée de moyenne, de fréquence, de tendance généralement et habituellement réalisée, et à celle de conformité à une règle, de jugement de valeur, de finalité assignée. (Rey, 1972: 5)

Il sera donc impossible de fournir une définition unique de la norme, puisque, « [en] réalité, il n'existe pas une norme, mais différents types de normes. » (Verreault, 1999: 33) La section suivante vise à présenter ces différentes normes.

2.1 LES DIFFÉRENTS TYPES DE NORMES

La question de la norme linguistique en est une complexe qui pourrait à elle seule faire l'objet d'une thèse²⁸. Par contre, comme elle ne constitue pas le propos central de notre travail, nous nous contenterons d'en dresser un portrait sommaire. Voici donc un survol des différentes normes auxquelles nous ferons appel dans le cadre de cette thèse.

Il existe tout d'abord une **norme réelle**, aussi appelée *norme de fonctionnement* (Marquilló Larruy, 2003) ou *norme primitive* (Brent, 1999), qui est étroitement liée à la deuxième acception que nous avons donnée d'USAGE. Cette norme réelle²⁹, conséquence de l'existence même de la langue, existe dans l'absolu, peu importe si l'on s'y intéresse ou si l'on tente de la décrire; elle est en fait constituée des régularités qui se dégagent de la somme des productions langagières des locuteurs d'une langue. C'est en respectant les règles implicites de la norme réelle de sa communauté linguistique qu'un locuteur affirme son

²⁸ Certains linguistes y ont notamment consacré plusieurs travaux, dont les plus importants sont sans doute ceux de Hjelmslev (1971) et Coseriu (1973).

²⁹ À noter que notre conception de la norme réelle diffère de celle de de Villers (2005) qui considère que la norme réelle est *l'usage effectif de la langue illustré à l'oral par les locuteurs prestigieux, à l'écrit, par la littérature, les textes officiels et les journaux.* (p.318)

appartenance à cette dernière. Il est important de bien comprendre que la norme réelle n'est pas accessible en tant que telle, puisqu'il est impossible d'analyser l'ensemble des productions langagières des membres d'une communauté linguistique pour en faire ressortir les régularités. C'est par le biais des normes construites, principalement de la norme descriptive, que nous obtenons une image plus ou moins fidèle de la norme réelle.

La **norme descriptive** tente quant à elle de rendre compte de la norme réelle en rendant explicites les règles tacites qui la gouvernent ou les régularités qui lui sont propres. Typique du travail des linguistes, l'approche descriptive se veut objective, c'est-à-dire qu'elle ne porte aucun jugement sur les phénomènes qu'elle observe. Malgré tout, il s'agit d'une norme construite, dont l'objectivité n'est pas totale, puisque la simple sélection des corpus faisant l'objet de la description – par exemple le fait de ne s'en tenir qu'aux « grands auteurs », à certains quotidiens, à des corpus écrits plutôt qu'oraux, etc. – introduit certains biais dans la description. Ainsi, la norme descriptive n'est en quelque sorte qu'un reflet d'une partie de la norme réelle, qu'elle ne peut décrire de façon exhaustive. Cette limite de la norme descriptive est particulièrement frappante si l'on pense au dictionnaire, outil descriptif par excellence qui, même dans le cas d'ouvrages très complets comme *Le Petit Robert*, n'arrive pas à rendre compte de tous les usages lexicaux.

La **norme prescriptive**, typique de la grammaire normative, hiérarchise les usages en en condamnant certains au profit de certains autres. Souvent appelée « bon usage », cette approche prescriptive spécifie quels sont les modèles auxquels doivent se conformer les locuteurs d'une langue s'ils veulent « bien s'exprimer ». Cette norme est subjective en ce qu'elle fonde ses choix normatifs sur des critères souvent flous (esthétisme) ou discutables (racine historique, tradition, justification « logique ») (Leeman-Bouix, 1994 ; Ducrot et Schaeffer, 1995 ; Brent, 1999) et qu'elle prend comme point de référence la langue d'une certaine élite (gens instruits, grands auteurs, lecteurs de nouvelles, etc.). Il arrive par conséquent que les principes dictés par la norme prescriptive ne soient pas

en accord avec l'usage de la majorité; pensons par exemple à l'emploi recommandé de l'indicatif après la locution conjonctive APRÈS QUE (*Après qu'il a été parti...* plutôt que *Après qu'il soit parti...*), qui semble marqué pour nombre de locuteurs. Selon Verreault (1999), le choix de prendre le français des gens instruits comme point de référence pour l'élaboration de la norme est subjectif, mais justifié par des considérations socioculturelles, puisque c'est généralement la langue qui est socialement valorisée par l'ensemble de la société. D'ailleurs, le fait d'adopter un point de vue prescriptif dans l'élaboration de la norme est « lui aussi motivé par des considérations d'ordre socioculturel puisqu'on reconnaît à certaines personnes ou à certains organismes le pouvoir – et dans certains cas le devoir – de se prononcer sur certains usages, d'orienter l'usage » (Verreault, 1999: 35); mentionnons à cet égard l'Académie française et l'Office québécois de la langue française. De tels organismes, en établissant en quelque sorte la norme prescriptive, jouent un rôle important dans la normalisation et la standardisation de la langue française. La norme prescriptive déterminée par ces institutions rejoint les locuteurs principalement par le biais d'ouvrages de référence comme les grammaires, mais surtout à travers les dictionnaires, « qui incarnent le plus commodément la norme linguistique d'une communauté et qui en assure (sic) la diffusion. » (de Villers, 2005: 21) Comme le signalait Rey (1983: 558): « [...] la fonction du dictionnaire est de fournir à ses usagers une référence sur la norme. »

Brent (1999) fait mention d'un autre type de norme davantage idiosyncrasique qui prend sa source dans les autres normes mentionnées : la **norme évaluative**. Lafontaine (1986) caractérise ainsi cette norme, qu'elle nomme pour sa part *norme subjective* :

On désigne par là les jugements de valeur individuels sur la langue, la façon dont l'individu évalue les productions linguistiques d'autrui et les siennes propres, ainsi que les représentations qu'il se fait de différents phénomènes linguistiques. (p.18)

La norme évaluative est implicite et découle d'un besoin d'agir dans l'action³⁰, même en l'absence d'une norme prescriptive. Elle constitue un « compromis pratique entre différents influx » (Brent, 1999: 125): la norme réelle, les ouvrages de référence, les instruments d'évaluation, le matériel pédagogique, sa propre conception du système de la langue, son propre usage de la langue, etc.

Lafontaine (1986) a mis en lumière l'importance de cette norme évaluative chez les enseignants. Dans son étude sur les attitudes et les normes subjectives de 214 enseignants belges du primaire et du secondaire, la chercheuse a démontré que l'idée fortement répandue selon laquelle la variété de français enseigné à l'école correspond à l'usage prescrit dans les ouvrages normatifs n'est pas fondée. En effet, elle a constaté que les enseignants « ignorent un certain nombre de prescriptions normatives [...] particulièrement [...] dans le domaine phonique et dans celui des particularités régionales. » (p.126) Par ailleurs, les enseignants sont souvent plus puristes que les grammairiens normatifs, notamment en condamnant les usages à propos desquels les ouvrages ne se prononcent pas. Force est donc d'admettre que la norme qui les guide n'est pas la norme prescriptive à proprement parler, mais une norme qui s'en inspire, tout en étant propre à chaque enseignant.

Quant aux contenus linguistiques prescrits, il entre dans le pouvoir de chaque enseignant, de déterminer les formes qu'il juge légitime de prôner et celles qu'il juge légitime de condamner. Assurément, l'enseignant n'est pas totalement libre de ses choix; il est plutôt en liberté surveillée, dans la mesure où il subit la pression sociale de plus « autorisé » que lui. Il jouit néanmoins d'une incontestable marge de liberté qui lui permet, soit de contester un point de vue légitime qu'il connaît bien par ailleurs [...], soit, dans l'ignorance du (des) point(s) de vue autorisé(s), de se forger un point de vue propre et de l'imposer aux autres. (Lafontaine, 1986: 138)

Brent (1999) mentionne aussi l'existence d'un cinquième type de norme qui, bien qu'il s'agisse de la norme la plus susceptible de nous intéresser dans le cadre de la présente thèse, est aussi celle à propos de laquelle nous disposons du

³⁰ Brent (1999) signale d'ailleurs que certains auteurs parlent de *norme pragmatique* pour désigner la norme évaluative. Nous éviterons cette étiquette afin de ne pas créer d'ambiguïté avec la norme fonctionnelle, qui prend en compte les dimensions pragmatiques d'utilisation de la langue.

moins d'information : la **norme pédagogique**. Selon l'auteur, cette norme, qui « serait normalement décrite de façon détaillée dans les programmes-cadres du ministère de l'Éducation », demeure largement implicite et se résume « à un renvoi sommaire à la norme prescriptive exogène consignée dans les bons dictionnaires et les bonnes grammaires » (p.124). Ce manque de consensus autour de la norme pédagogique explique pourquoi les enseignants s'en remettent souvent à leur norme évaluative dans l'exercice de leurs fonctions.

Nous reproduisons plus bas un tableau synthèse (Tableau 1), proposé par Brent (1999: 121), qui permet d'opposer les normes évoquées en fonction de différents paramètres ou *dimensions de caractères*, détaillés ci-dessous.

- a. Dimension de caractères linguistique/cognitive – linguistique – sociale : *Une norme langagière est d'autant plus SOCIALE qu'elle contient des éléments d'« évaluation sociale » du comportement langagier [...] Dans la mesure où elle se limite à actualiser le code linguistique ou à décrire son usage effectif, elle est COGNITIVE ou LINGUISTIQUE.* (p.124)
- b. Dimension de caractères implicite – explicite : Les normes explicites correspondent la plupart du temps à des normes construites.
- c. Dimension de caractères objective – subjective : Une norme subjective sera marquée par la présence d'éléments d'évaluation individuelle ou sociale.
- d. Dimension de caractères dynamique [variable] – statique [catégorique] : Une norme est jugée catégorique si elle ne tient *pas compte du caractère dynamique et variable du langage humain, c'est-à-dire [...] le changement du code linguistique à travers le temps ainsi que sa grande variabilité pragmatique, sociale et géographique.* (p.124)
- e. Dimension de caractères endogène – exogène : Une norme endogène prend sa source dans la communauté linguistique où elle se trouve, alors qu'une norme exogène est importée d'ailleurs.

L'observation de ce tableau synthèse nous permet de constater que les normes qui servent de guides à nos pratiques langagières relèvent en grande partie de l'implicite et de la subjectivité. Des pans importants de la langue que nous utilisons n'ont pas fait l'objet de descriptions systématiques et ne sont soumis à aucune prescription normative.

Tableau 1. Les différents types de normes selon Brent (1999 : 121)

Type		Caractères primaires	Caractères secondaires
A	RÉGULARITÉ DE COMPORTEMENT (langagier) ou NORME (langagière) PRIMITIVE ³¹	linguistique ou cognitive; implicite; le plus objective (non tributaire ou peu tributaire de la norme prescriptive ou descriptive)	dynamique [variable]; endogène
B	NORME (langagière) DESCRIPTIVE	linguistique; explicite; assez objective (mais plus ou moins influencée par des modèles théoriques et même par la norme prescriptive)	statique [catégorique] ou dynamique [variable]; endogène
C	NORME (langagière) ÉVALUATIVE ou PRAGMATIQUE	sociale; implicite; subjective (usage linguistique et évaluatif flottant entre la norme primitive, la norme descriptive [si connue], la norme prescriptive et la norme pédagogique [si connue] et se manifestant notamment dans la conduite de l'enseignement et de l'évaluation des apprentissages)	dynamique [variable]; endogène
D	NORME (langagière) PRESCRIPTIVE ou CANONIQUE	sociale; explicite; subjective (faisant appel à l'autorité, aux traditions, aux arguments d'ordre logique, esthétique, pseudo-psychologique)	statique [catégorique]; endogène ou exogène (au Québec elle est, de façon prépondérante, exogène)
E	NORME (langagière) PÉDAGOGIQUE	sociale; explicite ou implicite; subjective (modèle linguistique effectivement appliqué dans un programme d'études linguistiques ainsi que dans le matériel didactique et les instruments docimologiques afférents)	statique [catégorique]; endogène ou exogène (normalement endogène, même si elle s'appuie sur une norme prescriptive exogène)

Nous serions tenté d'ajouter, aux différentes normes présentées plus haut, une dernière qui les recoupe et que l'on pourrait qualifier de **fonctionnelle** (Dallaire, 1997). Cette norme sociale, à la fois implicite et explicite (notamment à travers les marques d'usage des dictionnaires), guide notre comportement langagier en fonction des situations de communication dans lesquelles nous nous trouvons³². Ainsi, une situation plus formelle exigera le recours à un vocabulaire courant ou soutenu et des constructions syntaxiques en accord avec les principes valorisés par la norme prescriptive; une conversation entre amis sera souvent colorée par

³¹ Par *norme primitive*, Brent fait référence à ce que nous avons appelé *norme réelle*.

³² Cette norme rejoint l'idée de variation diaphasique mentionnée plus haut (voir 1).

l'emploi d'un vocabulaire plus relâché et caractérisée par des tournures syntaxiques familières. Nous pourrions souligner en dernier lieu que la norme fonctionnelle est subjective et qu'elle a un caractère plutôt endogène; pour s'en convaincre, il suffit de penser au tutoiement, qui est acceptable au Québec dans un spectre de situations beaucoup plus large qu'en France.

Dans les sections qui suivent, nous nous intéresserons plus précisément à la norme lexicale et à la question du français québécois, qui fait l'objet de nombreux débats.

3. LA NORME LEXICALE

Cette discussion sur la norme nous amène maintenant à nous intéresser à la question de la norme lexicale, qui concerne plus directement nos préoccupations dans le cadre de la présente recherche. Citons d'entrée de jeu Siouffi et Steuckardt, (2001: 9)

De manière générale, on peut dire qu'à la différence de la syntaxe, le lexique est peut-être moins fermement et immédiatement assujéti aux questions de norme. Le domaine du lexique n'impose pas une compréhension très facile de la notion de norme (à la différence du domaine orthographique, par exemple). La variation y est sans doute plus sensible qu'en syntaxe, le changement plus spectaculaire, la créativité plus positivement perçue.

Dans sa réflexion à propos de la possibilité de définir le concept de norme lexicale, Guilbert (1972) arrive aussi à la conclusion que «la norme grammaticale se perçoit aisément, tandis qu'une norme lexicale reste à définir.» (p.29) Paradoxalement, le dictionnaire constitue pour nombre de locuteurs l'incarnation même de la norme linguistique. Pourquoi alors s'interroger à propos de l'existence d'une norme lexicale?

Selon nous, une partie de la réponse à cette question concerne la distinction entre les différents types de normes mentionnés plus haut. En effet, nul ne pourra nier l'existence d'une norme lexicale réelle, puisque cette dernière découle naturellement de l'usage que font les locuteurs du lexique de leur langue. Mais cette norme réelle s'actualise-t-elle dans une norme construite – descriptive ou prescriptive – qui nous permettrait d'y avoir accès?

Un premier réflexe serait sans doute de dire que la norme lexicale de la langue s'actualise dans le dictionnaire. Cette vision un peu naïve est celle d'une majorité de locuteurs, pour qui le dictionnaire constitue la dernière instance de décision en matière de lexique (« Si ce n'est pas dans le dictionnaire, ça n'existe pas! ») : « [...] l'exclusion lexicographique est en général assimilée à une condamnation. » (Simard, 1990: 37) Le dictionnaire de langue, comme l'indique Guilbert (1972: 42), « traduit le besoin social de codifier le lexique de la communauté », mais bien entendu, il ne permet pas de dresser un portrait exhaustif des usages lexicaux. En effet, toute entreprise lexicographique implique des choix :

Il s'agit de retenir du lexique les mots qu'implique un modèle socio-culturel de locuteur d'un niveau élevé, d'écarter les termes trop spécialisés dont l'emploi ne jouit pas d'une certaine diffusion dans la communauté, de n'admettre parmi les mots nouveaux que ceux qui ont acquis droit de cité par une certaine diffusion et jouissent d'un certain label par le texte où ils apparaissent et la personne qui les a produits, de conserver des termes même vieillis en raison de leur statut littéraire, de donner à chaque mot et expression un arbre généalogique qui garantit sa place dans le lexique. Le dictionnaire de langue sert donc à définir un certain usage du lexique, qui constitue une norme lexicale par rapport à l'ensemble de la communauté linguistique. (Guilbert, 1972: 42)

Selon Guilbert (1972), le lexique est régi par deux normes : une norme linguistique et une norme sociale. La norme linguistique se résume à quelques grands principes qui permettent la création d'UL : un nombre limité de règles concernant la construction morphologique (préfixation, suffixation, composition), et le phénomène de mutabilité sémantique, qui, en synchronie, permet à un locuteur certains effets stylistiques en étendant le sens des mots qu'il utilise et, diachroniquement, est à l'origine de la polysémie des vocables. Mais, au-delà de ces grandes zones de création, Guilbert fait remarquer que la norme lexicale est résolument sociale. Cette dimension sociale se manifeste notamment par les règles d'orthographe (souvent aberrantes d'un point de vue purement linguistique), la prononciation enseignée, la hiérarchisation des niveaux de langue de même que par les choix que font les concepteurs de dictionnaires et les recommandations du type « *ne dites pas... dites plutôt...* » émises par différentes instances faisant autorité. Selon Guilbert (1972: 47),

donc, « il existe bien une norme lexicale et [...] elle est essentiellement de caractère social et idéologique, donc évolutive. »

La meilleure illustration de la norme lexicale, selon Guilbert, demeure le dictionnaire. Malgré l'utilité indéniable ce type d'ouvrage comme moyen d'accès à cette norme, certains décalages entre la norme qu'il véhicule et la norme réelle – c'est-à-dire l'usage au sens d'*habitudes linguistiques* – méritent d'être soulignés. Tout d'abord, il faut mentionner que les dictionnaires ont généralement une visée plutôt descriptive, qui se heurte à une première difficulté : l'importance du stock lexical à décrire. En effet, contrairement à la norme grammaticale, qui doit rendre compte d'un nombre somme toute assez limité de phénomènes (systèmes d'accords, règles syntaxiques), la norme lexicale touche un nombre quasiment infini d'items. Ce problème a pour conséquence l'impossibilité de décrire le lexique de façon exhaustive – nous n'avons qu'à penser aux nombreuses acceptions manquantes dans les dictionnaires – et pour corolaire la question, évoquée plus haut, de la sélection des éléments à inclure dans les ouvrages. Deuxièmement, même pour les items retenus dans les ouvrages, la description elle-même n'est pas toujours exhaustive. En effet, chaque UL possède une multitude de caractéristiques propres, dont certaines peuvent échapper aux lexicographes. Par exemple, il est rare que les dictionnaires fournissent systématiquement des informations sur la fréquence des mots décrits (à l'exception de quelques marques d'usage), informations pourtant primordiales en vue d'une meilleure compréhension de la norme réelle. Un dernier facteur expliquant le décalage entre la norme descriptive des dictionnaires et la norme réelle est le tiraillement omniprésent entre l'approche descriptive généralement prônée par les lexicographes et les nombreuses incursions prescriptives qui parsèment ces ouvrages. Ces conflits entre normes, qui se manifestent déjà à l'étape de sélection des UL à décrire, peuvent conduire à des descriptions qui témoignent davantage du bon usage que de l'usage. Nous constatons par exemple l'absence du verbe DÉBUTER à la forme transitive dans nombre de dictionnaires, alors que cette forme transitive est

indéniablement répandue dans la population. Ces « irrégularités » dans l'approche descriptive constituent une preuve du caractère social de la norme lexicale.

En conclusion, la norme lexicale existe bel et bien, mais le dictionnaire ne nous en donne qu'une description souvent incomplète et teintée par des considérations prescriptives. Nous pouvons donc supposer que, devant une norme aussi floue, l'importance de la norme évaluative mentionnée plus haut est grande en matière de lexique, puisque c'est principalement par le vocabulaire que s'exprime l'idiosyncrasie linguistique des locuteurs. Nous verrons lors de la présentation de notre méthodologie de quelle façon nous entendons traiter la question de la norme lexicale (voir VII-2.1).

4. NORME ET FRANÇAIS QUÉBÉCOIS

Après plusieurs décennies de débat, il semble que l'existence d'un français propre au Québec soit aujourd'hui généralement admise. Martel et Cajolet-Laganière (1996: 71) définissent ainsi le français québécois : *l'ensemble des usages linguistiques constituant la variété de français utilisée sur le territoire du Québec*. Cette langue partage bien sûr de nombreux points avec le français de France, mais présente aussi plusieurs particularités, notamment en ce qui concerne les aspects phonologiques de la langue et, dans une moindre mesure, certaines tournures syntaxiques (Ostiguy et Tousignant, 1993). Mais évidemment, c'est surtout par le lexique que s'expriment les spécificités du français québécois (Martel et Cajolet-Laganière, 1996).

Dans une étude sur les particularités lexicales du français québécois dont la méthodologie reposait sur la comparaison de vastes corpus journalistiques québécois (*Le Devoir*) et français (*Le Monde*), de Villers conclut que « le recoupement des vocabulaires est minimalement de 77 % mais, dans les faits, il pourrait s'élever à plus de 85 %.³³ » (de Villers, 2005: 39) Au-delà de ce *tronc*

³³ Cette intersection importante entre le vocabulaire du français québécois et celui de la France serait sans aucun doute moindre dans une recherche qui considérerait aussi des textes oraux

commun important, les chiffres présentés par la chercheuse mettent bien en évidence une proportion non négligeable de vocabulaire propre au discours journalistique québécois, ce qui témoigne de l'existence d'une norme « standard » propre au français d'ici. En effet, les textes des journaux constituent sans doute la meilleure illustration du *modèle à suivre* au sein d'une communauté linguistique (de Villers, 2005: 319), et les particularités qui ressortent de l'étude de de Villers confirment que le modèle à suivre au Québec ne correspond pas nécessairement à la norme telle qu'elle est conçue en France. Mais où en est donc la description de cette norme?

Si la production lexicographique au Québec remonte à plus de 150 ans (de Villers, 2005: 25), c'est surtout au cours des vingt dernières années que l'entreprise de description du français québécois a franchi des pas de géants. En effet, de nombreux dictionnaires fabriqués ici fournissent aujourd'hui des descriptions étoffées de la langue parlée et écrite au Québec. Avec les multiples publications de L'Office québécois de la langue française et des ouvrages comme *Le Multidictionnaire*, le logiciel *Antidote*, *Le grand dictionnaire terminologique* et le dictionnaire *Franqus*, la communauté francophone du Québec dispose aujourd'hui d'outils qui, non seulement contribuent à *l'édification de sa norme linguistique* (de Villers, 2005: 25), mais affirment l'existence de cette dernière. Ainsi, la norme exogène française qui a longtemps dominé le paysage québécois fait lentement place à une norme mieux adaptée à notre langue. Si nous disons « lentement », c'est que les ouvrages que nous venons de mentionner ne jouissent pas encore d'une immense diffusion et que les dictionnaires les plus présents dans les écoles et les foyers québécois sont toujours *Le Petit Robert* et *Le Petit Larousse*, tous deux produits en France (Simard, 1990). L'insécurité linguistique des Québécois est aussi tenace, et nombreux sont ceux qui traînent encore un certain complexe d'infériorité par rapport au français de France, qui constitue à leurs yeux une langue plus légitime que celle de leur quotidien.

ainsi que des textes écrits de registre plus familier que les textes journalistiques sur lesquels repose l'étude de de Villers.

Malgré la progression plus ou moins rapide de la promotion d'un français propre au Québec dans la population en général, notamment dans le monde de l'enseignement où une norme pédagogique endogène reste à définir (Brent, 1999), nous avons aujourd'hui accès à des ouvrages de référence qui décrivent le français québécois avec rigueur et finesse et réussissent à faire émerger la norme d'ici, principalement d'un point de vue lexical. Nous verrons dans notre méthodologie de recherche que certains des ouvrages mentionnés plus haut nous permettront de tenir compte de cette norme du français québécois dans l'analyse de notre corpus de rédactions.

CONCLUSION

Ce chapitre nous aura permis de constater que la norme linguistique est un concept pluriel, au caractère parfois fuyant, notamment en ce qui concerne la norme lexicale. Dans les dictionnaires, incarnation de la norme pour nombre de locuteurs, la norme descriptive tente de rendre compte, partiellement il va sans dire, de la norme réelle, mais cède parfois le pas à une norme plus prescriptive. C'est en gardant en tête cette discussion sur la norme que nous définirons, dans le chapitre qui suit, le concept d'erreur linguistique.

IV. L'ERREUR LINGUISTIQUE

L'erreur constitue la pierre angulaire de notre travail et ce troisième chapitre de notre cadre conceptuel y est consacré. Nous définirons d'abord l'erreur linguistique telle que nous la concevons dans cette thèse et ferons quelques remarques terminologiques quant au vocabulaire utilisé pour parler des erreurs. Nous survolerons ensuite les différents courants de recherche qui se sont intéressés à l'erreur et nous verrons en quoi ils peuvent éclairer la question de l'analyse d'erreurs lexicales. Nous explorerons finalement les différences entre les erreurs linguistiques commises en L1 et en L2 pour justifier notre choix de nous appuyer sur de nombreuses études menées en contexte de langue seconde.

1. DÉFINITION DU CONCEPT D'ERREUR LINGUISTIQUE

Reuter, Cohen-Azria, Daunay, Delcambre et Lahanier-Reuter (2007) mentionnent la difficulté, dans une perspective didactique, de bien caractériser le concept d'*erreur* et le fait qu'il a été peu défini dans les travaux sur l'erreur et l'évaluation. Bien qu'intéressante, la définition qu'ils proposent du *dysfonctionnement à valeur didactique* nous semble un peu large pour vraiment éclairer la question de l'erreur linguistique. Par contre, des définitions de l'erreur (ou de la *faute* dans le cas de Debyser, Houis et Rojas, 1967) apparaissent dans des travaux didactiques portant plus précisément sur la langue. Nous en avons retenu quatre que nous jugeons intéressantes.

... comportement langagier (en compréhension ou en production) qui ne correspond pas à la norme grammaticale, lexico-sémantique, phonologique ou pragmatique, de L2. (Bailly, 1998: 87)

La faute est un écart par rapport à la réalisation attendue de la norme dans un contexte donné. (Debyser, Houis et Rojas, 1967: 10)

... a linguistic form or combination of forms which, in the same context and under similar conditions of production, would, in all likelihood, not be produced by the speaker's native speaker counterparts. (Lennon, 1991a: 182)

Did. lang. *Forme linguistique (mot, terminaison, prononciation, accord, etc.) ne correspondant pas à la forme attendue ou utilisée habituellement par les natifs. (Legendre, 2005: 613)*

Dans les paragraphes qui suivent, nous commenterons certains éléments de ces définitions. Nous en retiendrons ensuite les composantes les plus pertinentes afin de proposer notre propre définition.

1.1 L'ERREUR : UN ÉCART PAR RAPPORT À QUOI?

Le positionnement des auteurs relativement à la caractérisation de l'erreur nous permet d'établir une première distinction parmi les définitions que nous avons retenues. Debyser *et coll.* (1967) et Bailly (1998) définissent l'erreur par rapport à une (ou plusieurs) norme alors que Lennon (1991a) et Legendre (2005) la mettent plutôt en relation avec ce qu'aurait produit un locuteur natif. Bien entendu, ces deux points de vue ne sont pas irréconciliables, puisque la façon typique dont s'exprime un locuteur natif est considérée comme l'usage normal dans sa langue et en constitue donc la norme, du moins la norme réelle. Néanmoins, nous croyons que certaines raisons justifient d'adopter une position plus proche de celle de Lennon et de Legendre.

Tout d'abord, comme le souligne James (1998), le choix de la grammaire de référence – celle faisant office de norme – qu'implique une définition comme celle de Debyser *et coll.* – n'est pas toujours facile à faire. De plus, pour que cette grammaire de référence permette de rendre compte de différents types d'erreurs, il faudrait qu'elle ratisse très large en englobant aussi les dimensions sémantique, phonologique et pragmatique de la langue. La façon dont Bailly définit l'erreur permet de mettre en lumière cette difficulté : « [...] qui ne correspond pas à la norme grammaticale, lexico-sémantique, phonologique ou pragmatique, de L2. » (Bailly, 1998: 87) Si les normes grammaticale et phonologique d'une langue sont généralement assez bien décrites et définies, il en va autrement des dimensions sémantique et pragmatique, qui sont soumises à des normes beaucoup plus floues. Une définition de l'erreur reposant sur la notion de norme est donc difficilement opérationnelle dans la mesure où, comme nous l'avons vu au chapitre III, il existe plusieurs types de normes qui toutes ne sont pas faciles à définir.

James (1998) mentionne qu'il existe différents critères qui nous permettent de statuer sur le caractère erroné ou non d'un énoncé, dont les deux principaux sont la grammaticalité et l'acceptabilité. La grammaticalité ne dépend pas du jugement du correcteur, puisque c'est la grammaire d'une langue qui détermine si l'énoncé produit obéit ou non à ses règles. Ce critère est indépendant du contexte, c'est-à-dire qu'un énoncé agrammatical ne pourra apparaître dans aucun contexte. L'acceptabilité est plus subjective et dépend du jugement du correcteur, qui peut considérer qu'un énoncé est inacceptable en fonction du contexte dans lequel il est produit ou simplement par son caractère étrange. Ainsi, un énoncé tout à fait grammatical peut malgré tout être jugé inacceptable³⁴. James présente différentes raisons qui peuvent expliquer qu'un énoncé soit considéré inacceptable : incompatibilité avec le contexte, nature étrange de l'idée exprimée (situation improbable ou impossible), façon maladroite ou peu idiomatique de faire référence à une réalité envisageable, collocation inusitée, construction grammaticale étrange ou construction phonologique dérangement, construction difficile à décoder (ex. structures imbriquées), phrase mal balancée, non-respect du caractère figé d'une expression (ex. ?*blanc et noir* pour *noir et blanc*), non-respect d'une règle imposée par les puristes.

Cette distinction entre grammaticalité et acceptabilité appuie selon nous le choix de définir l'erreur comme un écart par rapport à ce que produirait un locuteur natif plutôt que par rapport à une norme. En effet, une définition de l'erreur comme celle de Debyser *et coll.* (1967) peut s'avérer utile pour décrire les énoncés quant à leur grammaticalité, mais s'avère moins efficace pour rendre compte de leur acceptabilité. En caractérisant les écarts observés par rapport aux formes qu'aurait utilisées un locuteur natif, nous élargissons le spectre des problèmes couverts. En effet, non seulement un locuteur natif ne produit pas d'énoncés agrammaticaux, mais il sait aussi s'adapter au contexte d'énonciation et maîtrise assez bien sa langue pour s'exprimer de façon idiomatique sans

³⁴ Même si l'on dit souvent que la grammaticalité est une condition *sine qua non* à l'acceptabilité, James (1998) rapporte quelques cas où des énoncés agrammaticaux sont jugés acceptables.

produire des énoncés douteux; par conséquent, ses énoncés sont à la fois grammaticaux et acceptables. Évidemment, nous faisons ici abstraction des erreurs de performance. Marquilló Larruy (2003: 43) rejoint d'ailleurs nos propos lorsqu'elle mentionne que les locuteurs natifs sont susceptibles de produire des énoncés « hors normes » en vertu de certaines positions prescriptives, mais ne seront que très rarement à l'origine d'énoncés « hors système », c'est-à-dire qui iraient à l'encontre du système linguistique.

Positionner notre définition de l'erreur par rapport à la notion de locuteur natif sera d'autant plus pertinent que nous aspirons à caractériser des problèmes lexicaux, dont plusieurs sont soumis à des normes très floues (voir section III-3). En plaçant le locuteur natif au centre de notre définition, nous n'évacuons pas la notion de norme, mais nous la subordonnons à celle de locuteur. Notre position se situe donc résolument du côté de la parole plutôt que de celui de la langue et nous pousse à proposer une définition moins prescriptive de l'erreur, puisqu'elle repose sur l'observation des énoncés produits par des locuteurs.

1.2 NOTRE LOCUTEUR DE RÉFÉRENCE : UN LOCUTEUR EXPERT

Nous jugeons donc qu'il est préférable de définir l'erreur comme un écart par rapport à l'énoncé qu'aurait produit un locuteur natif, comme le suggèrent Lennon (1991a) et Legendre (2005), plutôt que par rapport à une norme. Par contre, si cette référence au locuteur natif est pertinente dans des travaux portant sur l'apprentissage d'une L2 – comme ceux de Debyser *et coll.* (1967), Lennon (1991a) et Bailly (1998) –, elle l'est moins dans une recherche comme la nôtre, qui s'intéresse aux erreurs en L1, donc produites par des locuteurs natifs. Pour pallier ce problème, notre définition de l'erreur devra plutôt faire référence à un locuteur expert, c'est-à-dire un locuteur présentant les caractéristiques suivantes :

- Il s'agit d'un locuteur natif.
- Il a un bon niveau d'éducation (qui a fait des études universitaires, ou qui pourrait en compléter) et une très bonne maîtrise de la langue, c'est-à-dire qu'il maîtrise les règles de grammaire (norme prescriptive) et qu'il possède un vocabulaire précis et étendu.

- Il s'agit d'un locuteur flexible qui sait bien s'adapter à différents contextes d'énonciation (norme fonctionnelle), notamment par l'emploi de différents registres de langue. Il se permet aussi de transgresser parfois certaines règles pour produire des effets de style (humour, jeux de mots, figures de style, etc.).
- Il s'agit d'un locuteur idéalisé³⁵, c'est-à-dire qu'il n'est jamais confronté à des problèmes d'accès au vocabulaire et qu'il ne commet pas d'erreurs de performance.

Nous rappelons que le fait de placer le locuteur expert au centre de notre définition de l'erreur ne nous libère de la notion de norme qu'en apparence, puisque la norme et la façon de s'exprimer d'un locuteur expert sont interreliées et s'influencent mutuellement. De la même façon qu'il aurait fallu établir clairement la norme pour rendre fonctionnelle une définition de l'erreur comme celles de Debyser *et coll.* (1967) ou Bailly (1998), il nous faudra opérationnaliser le concept de locuteur expert. Au-delà de nos propres intuitions, comment pourrions-nous considérer qu'un énoncé est typique ou non de la façon de s'exprimer du locuteur expert? Sur quels ouvrages de référence ou sur quels outils s'appuieront nos jugements? Ces considérations méthodologiques seront abordées à la section VII-2.1.

1.3 ÉLÉMENTS À RETENIR DANS NOTRE DÉFINITION DE L'ERREUR LINGUISTIQUE

La présente section traitera de divers éléments qui nous semblent importants dans une définition de l'erreur.

Hormis celle de Legendre (2005), les définitions présentées plus haut font toutes référence, de façon plus ou moins explicite³⁶, au contexte dans lequel apparaît l'erreur. Cet élément de définition nous paraît essentiel, puisqu'il fait référence au caractère relatif de l'erreur soulevé par Arabski (1979) : une forme

³⁵ Il est à noter que la notion de locuteur idéal n'est pas nouvelle. Elle occupe en effet une place centrale dans la grammaire générative de Chomsky, qui fait référence à un locuteur-auditeur « idéal appartenant à une communauté linguistique complètement homogène, qui connaît parfaitement sa langue et qui, lorsqu'il applique en une performance effective sa connaissance de la langue n'est pas affecté par des conditions grammaticalement non pertinentes, telles que limitation de mémoire, distractions, déplacement d'intérêt ou d'attention, erreurs (fortuites ou caractéristiques) » (Chomsky, 1971: 12).

³⁶ Bailly (1998) y fait allusion par le biais de la « norme pragmatique ».

linguistique acceptable dans un contexte peut être inacceptable dans un autre. Par conséquent, la notion de contexte devra apparaître dans notre définition.

Debyser *et coll.* (1967) et Legendre (2005) mentionnent que l'erreur est une forme qui ne correspond pas à celle attendue. Cette composante de leur définition est très intéressante en ce qu'elle met en lumière une étape inhérente à l'analyse d'erreurs : la restitution de la forme correcte, qui correspond justement la forme attendue. Le fait de considérer la forme erronée comme plus ou moins éloignée de l'item visé permet aussi de rendre compte de la notion d'écart dont parlent Debyser *et coll.* (1967).

Comme James (1998), nous considérons que la meilleure définition de l'erreur est celle proposée par Lennon (1991a). Outre les éléments mentionnés jusqu'à maintenant, un autre retient notre attention dans sa définition : le fait que l'erreur peut concerner à la fois des formes linguistiques et des combinaisons de formes. Cette composante de sa définition est importante pour nous, puisque plusieurs problèmes lexicaux découlent de la combinaison de deux UL, notamment les problèmes de collocations ou d'incompatibilité sémantique ainsi que les pléonasmés.

1.4 NOTRE DÉFINITION DE L'ERREUR LINGUISTIQUE

Voici donc la définition d'*erreur linguistique* que nous proposons dans cette thèse.

Erreur linguistique : forme linguistique ou combinaison de formes linguistiques qui diffère de ce qu'un locuteur expert aurait selon toute probabilité produit dans un même contexte de production.

La définition de l'erreur que nous adoptons est très proche de celle de Lennon (1991a), puisque cette dernière contient presque tous les éléments que nous jugeons essentiels. Seule la référence à un locuteur expert plutôt qu'à un locuteur natif distingue notre définition de celle de l'auteur. Par ce petit changement, nous croyons que nous proposons une définition plus universelle de l'erreur, qui couvre tant les problèmes en L1 qu'en L2.

Il est important de souligner que la définition de l'erreur que nous choisissons ici se limite aux erreurs en production. La définition de Bailly (1998) est la seule que nous ayons rencontrée qui décrit à la fois les erreurs de production et de compréhension. Par contre, la volonté de fournir une définition qui englobe ces deux grandes classes d'erreurs oblige Bailly à utiliser un genre prochain que nous jugeons peu efficace : « comportement langagier ». Ce choix est d'autant plus malheureux, à notre avis, qu'il met l'accent sur le processus ayant mené à l'erreur plutôt que sur l'erreur elle-même, qui n'est qu'une conséquence de ce comportement langagier. Notre définition s'intéresse au produit (l'erreur) et non aux processus qui y ont mené.

Une dernière remarque s'impose par rapport à notre définition. L'erreur, telle que nous la définissons ici, englobe tous les écarts linguistiques par rapport à ce que produirait un locuteur expert, indépendamment de leur « gravité » ou de leur systématicité. Elle permet de désigner tant les problèmes clairement condamnables en vertu d'une norme claire que les emplois simplement marqués par rapport à l'usage normal, qui pourraient être qualifiés plutôt de maladroites. Nous reviendrons à la question de la gravité des erreurs à la section 3.2, mais attardons-nous maintenant à certaines considérations terminologiques relatives à l'erreur.

1.5 ERREUR, FAUTE, DYSFONCTIONNEMENT? REMARQUES SUR LA TERMINOLOGIE RELATIVE À L'ERREUR

Les termes ERREUR et FAUTE ont longtemps été utilisés indifféremment dans l'enseignement, mais la tendance est, depuis quelques décennies, de privilégier l'emploi d'ERREUR³⁷. En effet, FAUTE véhicule, dans certaines de ses acceptions, l'idée de 'manquement à un engagement moral ou religieux, à une loi'; il est donc plus lourdement connoté qu'ERREUR. La didactique n'ayant plus tendance à considérer l'erreur « comme pathologique et comme à éradiquer » (Spychiger,

³⁷ On remarque tout de même quelques « dissidents », par exemple Patris et Vansnick (1992) qui s'intéressent aux fautes lexicales et Lamy, qui propose une pédagogie de la faute (Lamy, 1976) et semble plutôt attaché à ce terme (Lamy, 1984).

2006: 18), mais plutôt comme un phénomène normal du processus d'apprentissage, voire un élément central de ce processus, il est tout à fait compréhensible que l'on préfère un terme avec moins de connotations négatives. Mais au-delà de ces considérations connotatives, ces choix terminologiques cachent-ils des nuances conceptuelles? Cette question est cruciale lorsqu'on s'intéresse à l'analyse d'erreurs, puisque les différents dysfonctionnements dénotés par ces termes n'auront peut-être pas la même pertinence du point de vue de l'analyse.

La première distinction importante entre ERREUR et FAUTE, qui correspond à l'opposition ERROR et MISTAKE de la littérature anglophone, est liée aux notions de *performance* et de *compétence* proposées dans le cadre de la grammaire générative (voir notamment Chomsky, 1965). Un problème de performance, c'est-à-dire un échec d'exécution lors de la mise en œuvre de savoirs linguistiques dans la production ou la compréhension d'énoncés, sera appelé *faute*, alors qu'un dysfonctionnement attribuable à une faille dans la compétence du locuteur, soit sa capacité à produire ou à reconnaître des énoncés appartenant à sa langue, sera désigné par le terme ERREUR (Corder, 1967/1980 ; Besse et Porquier, 1984). Bien que cette opposition semble partagée par de nombreux chercheurs, il existe certaines nuances dans la terminologie employée dans la littérature.

Dans un article portant précisément sur les distinctions entre différents termes liés à l'erreur, Bertrand (1987), propose d'utiliser LAPSUS plutôt que FAUTE pour les dysfonctionnements découlant d'un simple problème dans l'application de règles pourtant connues (et généralement maîtrisées)³⁸. « [...] Parce qu'ils sont "de la faute" de l'élève » et méritent en quelque sorte tout le poids connotatif du terme, Bertrand réserve l'emploi de FAUTE aux problèmes ayant pour source « [le] manque de sérieux dans l'apprentissage, la négligence... » (p.79). L'auteur

³⁸ C'est d'ailleurs la position généralement adoptée par les chercheurs de langue anglaise, qui parlent de *lapses*, mais sans l'opposer à *error* qui en constitue le terme générique ou hyperonyme (voir notamment Reason, 1990: 9).

utilise finalement ECART comme terme générique pour regrouper les trois types de dysfonctionnements : les erreurs, les lapsus et les fautes.

James (1998), pour sa part, croit que la classification des écarts la plus pratique est celle reposant sur la capacité de l'apprenant à rétablir la forme correcte. Les lapsus (*slips* ou *lapses*) peuvent être repérés et corrigés spontanément par l'apprenant sans aide extérieure alors que les fautes (*mistakes*) ne peuvent être corrigées que lorsqu'on lui en fait remarquer la présence³⁹. Les erreurs (*errors*), elles, ne peuvent être corrigées par l'apprenant sans que ne soient faits des apprentissages supplémentaires. Finalement, une dernière catégorie, que James nomme « *solecisms* », regroupe les erreurs découlant de la transgression d'une règle « de correction » imposée par la norme prescriptive, mais qui est en général contre-intuitive même pour les locuteurs natifs; par conséquent, les erreurs de ce type seront souvent présentes dans l'*input* langagier servant de modèle à l'apprenant⁴⁰.

Ce qu'il est important de retenir de cette discussion terminologique est que le terme ERREUR semble le plus répandu (Reuter *et coll.*, 2007) pour dénoter un écart attribuable à une lacune dans les connaissances du locuteur ou découlant de l'application d'une règle fautive faisant partie d'un système linguistique transitoire de l'apprenant⁴¹ : les erreurs de compétence au sens chomskyen. Seules ces « fautes intelligentes » (Berthoud et Py, 1993: 59), par leur récurrence et leur systématisme, seront pertinentes d'un point de vue cognitif et pourront éclairer les processus d'apprentissage de l'apprenant. En bref, ce sont les erreurs qui intéressent l'analyse d'erreurs!

Arabski (1979), qui croit aussi qu'il existe différents niveaux entre l'erreur et la faute, souligne par contre que tous les problèmes sont importants d'un point de

³⁹ James (1998) distingue les « *first-order mistakes* », qui sont corrigées immédiatement lorsqu'on mentionne à l'apprenant qu'il y a un problème, des « *second-order mistakes* », dont la correction exige plus de détails sur la localisation ou la nature du problème.

⁴⁰ Cette idée rappelle la notion d'*erreur induite*, que nous présenterons plus loin (voir IV-3).

⁴¹ À noter cependant que Reuter (2005) préfère le terme *dysfonctionnement*, qu'il considère moins restrictif et qui selon lui permet, par sa parenté formelle avec *fonctionnement*, d'appréhender le dysfonctionnement comme un phénomène « indissociablement lié aux fonctionnements eux-mêmes » (p.213).

vue didactique, puisque l'enseignant vise à les voir tous disparaître des productions de ses élèves. Évidemment, le type d'intervention didactique variera avec le type de problèmes – les fautes nécessitant des exercices pédagogiques intensifs alors que les erreurs demandent des explications – et certains problèmes mériteront davantage l'attention de l'enseignant.

Nous sommes plutôt d'accord avec la position d'Arabski et nous croyons qu'elle est d'autant plus pertinente en matière de lexicque, où les problèmes seront rarement caractérisés par la récurrence et la systématicité. Le caractère ponctuel des problèmes lexicaux et le fait qu'ils soient associés à des UL spécifiques feront qu'il sera souvent impossible d'établir s'il s'agit d'erreurs ou de fautes, particulièrement dans le cas d'une analyse à partir d'un corpus écrit, comme la nôtre, dans laquelle il est impossible de consulter les scripteurs à l'origine des textes. Comment, en effet, savoir si un mauvais emploi lexical témoigne d'une faille dans le vocabulaire de l'élève, s'il résulte d'un problème d'inattention ou simplement de paresse ou de négligence? Nous sommes donc d'avis qu'il serait hasardeux, dans la présente thèse, de tenter de rendre compte du statut des différents problèmes lexicaux relevés par des oppositions terminologiques comme celles proposées par Bertrand (1987).

*Nous emploierons donc, dans cette thèse, le terme **ERREUR** pour désigner tout type de dysfonctionnement lexical, peu importe son origine et indépendamment de sa gravité relative par rapport aux autres problèmes analysés.*

Si nous privilégions ce terme pour désigner tous les problèmes lexicaux, c'est parce qu'il est le plus utilisé dans la littérature et celui auquel se rattachent le mieux les concepts centraux développés dans cette thèse (cf. norme, rapport à l'erreur).

Bien entendu, ceci ne signifie pas que tous les problèmes lexicaux ont pour nous la même importance et doivent recevoir la même attention didactique de la part des enseignants. Bien sûr, il faudra hiérarchiser les erreurs relevées dans notre corpus pour relativiser leur importance et proposer des pistes d'intervention appropriées; par contre, nous ne croyons pas que c'est par l'emploi d'un nombre

limité d'étiquettes (cf. *erreur, faute, lapsus*) qu'il sera possible de rendre compte de cette hiérarchisation. Par exemple, l'opposition *erreur ~ maladresse* utilisée dans notre mémoire (Anctil, 2005) s'était avérée plutôt décevante parce qu'elle ne proposait qu'une gradation binaire qui ne permettait pas de rendre compte du continuum qui existe entre un emploi jugé unanimement comme acceptable et un autre condamné de façon consensuelle; même en multipliant les étiquettes, il serait difficile de positionner de façon satisfaisante les problèmes sur ce continuum. Nous verrons au point 3.2 de quelle façon des auteurs ont tenté de rendre compte de l'importance relative des erreurs dans leurs travaux.

2. LES TRAVAUX SUR L'ANALYSE D'ERREURS

La pertinence didactique de tenir compte des erreurs des apprenants est aujourd'hui solidement établie, notamment par les travaux d'Astolfi (1997). De tout temps, et principalement depuis les années 50, alors que l'erreur était souvent perçue dans le monde de l'enseignement comme la manifestation d'un enseignement non efficace ou d'un manque de connaissances, des linguistes et didacticiens s'y sont intéressés, menant à l'émergence de courants de recherche sur l'analyse d'erreurs. Le présent chapitre fera ressortir l'apport de ces travaux pour la catégorisation, la compréhension et le traitement des erreurs lexicales.

Nous présenterons d'abord les visées des travaux sur l'analyse d'erreurs, puis nous détaillerons les différentes composantes d'une telle démarche avant de faire un survol des trois courants principaux d'analyse d'erreurs. La section 3 s'intéressera quant à elle à décrire les apports des travaux mentionnés pour l'analyse d'erreurs lexicales.

2.1 LES VISÉES DE L'ANALYSE D'ERREURS

Les travaux portant sur l'analyse d'erreurs peuvent avoir deux grandes orientations (Porquier, 1977 ; Corder, 1981 ; Berthoud, 1987) :

1. La première est **didactique**. Elle permet d'abord d'avoir un portrait plus juste des difficultés des apprenants – c'est la dimension *diagnostique* (Porquier, 1977) – pour ensuite adapter les interventions de l'enseignant et, plus globalement, l'enseignement lui-

même (contenu à présenter, ordre de présentation, méthode d'enseignement, matériel didactique, etc.) – il s'agit de la dimension *pronostique*.

2. La seconde est **psycholinguistique**. Du point de vue de l'apprenant, l'erreur constitue un moyen d'apprendre en lui permettant de vérifier l'efficacité du système linguistique qu'il construit (Berthoud et Py, 1993). Pour le linguiste, l'étude des problèmes rencontrés par les apprenants permet en fait de révéler (du moins en partie) leurs processus d'apprentissage, les hypothèses qu'ils formulent à propos de la langue à apprendre et les stratégies qu'ils mettent en œuvre dans leur apprentissage.

Nous pouvons ajouter un troisième point qui, même s'il ne constitue pas véritablement l'une des deux grandes visées de l'analyse d'erreurs, s'avère une retombée intéressante dans le domaine de la linguistique : celle d'éclairer le fonctionnement même de la langue (Porquier, 1977: 38). En effet, une utilisation erronée d'une UL ou d'une structure permet souvent de mettre en lumière certaines contraintes qui autrement seraient restées implicites (Shaughnessy, 1977), ce qui peut s'avérer précieux lors d'activités de description linguistique, notamment la lexicographie.

2.2 LES COMPOSANTES DE L'ANALYSE D'ERREURS

La Figure 3 représente les différentes dimensions à considérer dans l'analyse d'erreurs, particulièrement dans une perspective didactique.

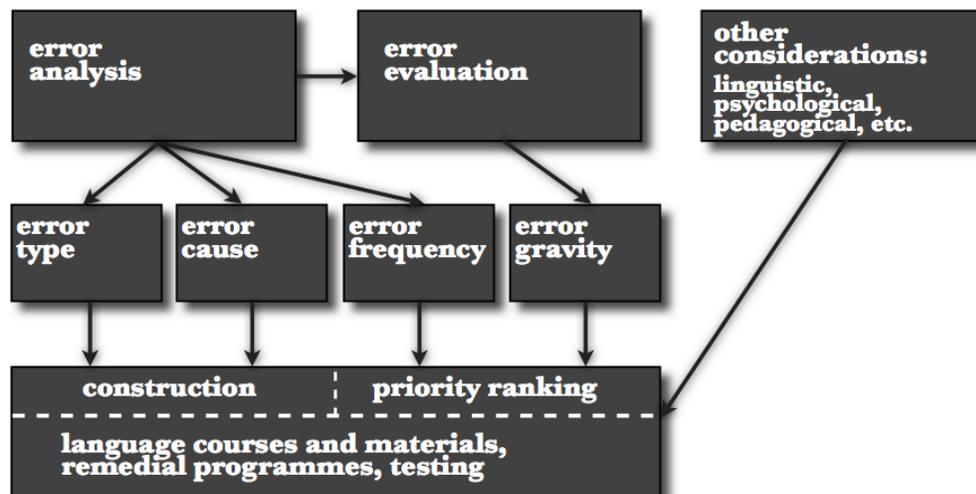


Figure 3. Error analysis and error evaluation (Johansson, 1978: 3)

Cette figure nous paraît particulièrement intéressante, puisqu'elle met en lumière les différentes étapes de l'analyse d'erreurs, telles que nous les verrons plus bas. Elle fait aussi ressortir l'importance d'un aspect souvent négligé dans les travaux sur l'analyse d'erreurs : l'évaluation de la gravité des problèmes décrits (Johansson, 1978). La figure relativise aussi l'importance de l'analyse d'erreurs dans la détermination des contenus d'un cours de langue, et l'ajustement des interventions didactiques ainsi que des modalités d'évaluation; une connaissance des erreurs commises par les apprenants ne peut suffire à elle seule et d'autres considérations entrent en ligne de compte.

2.3 SURVOL HISTORIQUE DES APPROCHES DE L'ANALYSE D'ERREURS

La section qui suit propose un bref survol historique de trois courants de recherche importants s'étant intéressés spécifiquement à l'erreur : l'analyse contrastive, l'analyse d'erreurs et les travaux sur les interlangues.

2.3.1 L'ANALYSE CONTRASTIVE

L'analyse contrastive (AC) apparaît dans les années 50 avec les travaux de C. C. Fries et de R. Lado et donne lieu à de nombreuses recherches jusque dans les années 70. Développée dans le contexte de l'enseignement des langues secondes, l'AC est, comme le mentionne Marquilló Larruy (2003), une analyse à priori : elle tente d'anticiper les erreurs des apprenants et leur analyse à proprement parler ne sert qu'à valider ou infirmer ses hypothèses.

L'AC prend racine dans deux courants théoriques de l'époque. Linguistiquement, elle va puiser dans le structuralisme, qui conçoit la langue comme une structure, c'est-à-dire un système dans lequel les éléments se définissent par les relations (équivalence ou opposition) qu'ils entretiennent avec les autres; par conséquent, toutes les langues sont différentes et ne présentent aucun élément commun (Py, 1984). Du point de vue psychologique, c'est sur le béhaviorisme que repose l'AC : l'apprentissage d'une L2 consiste essentiellement en un transfert à cette dernière des savoirs et habitudes déjà acquis en L1. Comme l'apprentissage repose sur le transfert de savoirs appartenant à un système linguistique vers un

autre qui en est par essence différent, l'AC tente de prévoir, par la comparaison des deux langues en jeu, les points sur lesquels l'apprentissage risque d'achopper; ceux sur lesquels les langues se ressemblent devraient être facilement acquis par les apprenants, alors que de plus grands écarts entre les systèmes linguistiques devraient donner lieu à des interférences (Debyser, 1971). Une fois ces zones sensibles identifiées, des exercices de remédiation sont donnés aux apprenants de sorte à prévenir des transferts négatifs⁴² et à développer chez eux des habitudes propres à la langue cible. L'AC est donc orientée directement vers des applications pratiques et ne vise pas à éclairer les processus d'apprentissage des apprenants; en d'autres termes, elle a une visée résolument didactique, focalisée sur la dimension pronostique (voir 2.1).

Durant sa période la plus féconde, l'AC repose sur ce que l'on a nommé l'*hypothèse forte* : « la confrontation structurelle des langues doit permettre de prévoir **toutes** les difficultés rencontrées par les apprenants. » (Marquilló Larruy, 2003: 64) Par contre, l'observation des erreurs commises pousse les tenants de cette position à nuancer leur hypothèse. En effet, plusieurs des difficultés rencontrées par les apprenants ne sont pas prévues par l'AC (Py, 1984: 34), alors que toutes les erreurs qu'elle anticipe ne se retrouvent pas dans leurs productions et que des erreurs similaires sont produites par des apprenants ayant des L1 différentes (Marquilló Larruy, 2003). Malgré tout, l'AC permet d'assez bonnes prédictions pour ce qui concerne les aspects phonologiques⁴³ (Debyser, 1971).

Selon Py (1984), le succès plutôt mitigé de l'AC s'explique surtout par l'absence de prise en compte de l'apprenant lui-même dans le processus d'apprentissage. En effet, ce type d'analyse repose uniquement sur une comparaison

⁴² Le transfert est le processus psycholinguistique qui consiste à utiliser des ressources de la langue première en langue seconde et se révèle, dans le cas de transfert négatif, par des interférences. Le transfert n'est donc pas observable en soi, mais doit être inféré à la lumière des productions des apprenants, notamment des interférences qu'elles présentent (Porquier et Py, 2004).

⁴³ Il est d'ailleurs intéressant de noter que le structuralisme en linguistique s'est d'abord développé à partir de l'aspect phonologique des langues avant de s'étendre aux autres dimensions.

« linguistique » de deux langues sans tenir compte du rôle de médiation que joue l'apprenant d'une L2, c'est-à-dire sa façon de mettre en relation les deux systèmes en jeu et leur interaction dynamique dans sa façon de les concevoir. Py (1984: 35) explique que chacune des langues se structure sur deux axes : « un axe intrasystémique, qui représente la logique spécifique de la langue concernée; et un axe extrasystémique, qui représente une restructuration inspirée par l'autre langue. »

Porquier (1977) souligne aussi un point important : les interférences n'ont pas lieu entre la L1 et la L2, mais entre un système linguistique déjà acquis et un système provisoire partiel (qui n'est pas celui de la langue cible, voir 3.3), donc entre la connaissance plus ou moins complète ou exacte qu'a l'apprenant des deux systèmes en jeu. Par conséquent, une analyse fondée sur la comparaison de deux systèmes linguistiques « idéalisés » ne saurait prévoir tous les obstacles que peut rencontrer un apprenant. Ceci est d'autant plus vrai que l'AC évacue les erreurs intralinguistiques (Gass et Selinker, 2001), notamment analogiques, dont le rôle et la fréquence ont été démontrés même en L2 (Porquier, 1977 ; Ellis, 1994).

En bref, l'AC, dans sa version « dure », n'a pas su faire ses preuves comme outil prédictif des problèmes rencontrés par les apprenants d'une L2. Par contre, comme le souligne Py (1984), elle peut s'avérer une source extrêmement précieuse pour l'explication à posteriori de problèmes observés chez des apprenants chez qui coexistent, à des niveaux de maîtrise variables, au moins deux systèmes linguistiques.

2.3.2 L'ANALYSE D'ERREURS (AÉ)⁴⁴

Dans les années 70, les chercheurs veulent se dégager de l'orientation purement linguistique de l'AC en étudiant les erreurs avec un regard davantage psycholinguistique, en prenant en compte le rôle de l'apprenant dans

⁴⁴ Nous désignerons l'approche nommée « analyse d'erreurs » par les lettres AÉ afin de marquer la distinction avec l'analyse d'erreurs dans son sens global, c'est-à-dire toute démarche ayant pour objet la description ou l'explication des erreurs.

l'apprentissage de la L2 (Vogel, 1995). Un changement de perception face à l'erreur s'opère pour en donner une vision plus positive; « on voit dans l'erreur la manifestation naturelle de l'apprentissage » (Marquilló Larruy, 2003: 72) et on ne la perçoit plus comme le résultat d'un enseignement imparfait (Berthoud et Py, 1993).

C'est à cette époque qu'il y a un déplacement de l'analyse à priori des erreurs vers une analyse à postériori (Marquilló Larruy, 2003); plutôt que de tenter de prévoir les erreurs des apprenants, les chercheurs en font la cueillette dans les productions de ces derniers afin de les regrouper pour pouvoir les observer. Cette période est marquée par l'apparition de nombreuses typologies visant à catégoriser les erreurs selon différents critères. Par contre, la multiplication des classements, la grande variété des critères utilisés et leur caractère parfois imprécis ainsi que l'hétérogénéité des corpus étudiés (populations et types de productions) ne permettent pas l'extrapolation des observations faites (Porquier, 1977).

L'un des reproches les plus importants adressés à ce courant que l'on nomme l'AE est de demeurer résolument descriptif, sans tenter de remonter aux sources des erreurs, et de concevoir le traitement statistique des erreurs comme une fin en soi. « Il ne suffit pas de savoir quelles erreurs se produisent et sur quoi elles portent, il faut aussi trouver pourquoi elles se produisent. » (Porquier, 1977: 26) Ce reproche met en évidence l'importance de distinguer la description des erreurs (le résultat) et leur explication (les processus qui y ont conduit) et implique l'idée d'une analyse à deux niveaux; Porquier (1977) parle d'approche typologique et d'approche étiologique.

Ellis (1994) résume deux principales critiques relatives au manque de portée de l'AE :

1. On pointe son incapacité à fournir un portrait global de la maîtrise de la langue de l'apprenant, puisqu'elle s'intéresse seulement aux aspects problématiques de sa production langagière (et ne tient pas compte des stratégies d'évitement).
2. L'AE n'ayant donné lieu qu'à très peu d'études longitudinales, elle ne décrit la compétence langagière d'un apprenant qu'à un moment précis,

ce qui ne révèle rien sur sa progression ou les différents niveaux de développement impliqués dans l'apprentissage d'une langue⁴⁵.

Malgré ces faiblesses, l'AE aura eu comme impact principal d'attirer l'attention sur l'erreur en lui attribuant une valeur plus positive et en lançant l'idée qu'elle constitue une fenêtre ouverte sur les processus d'acquisition des apprenants, ce qui correspond à la visée psycholinguistique de l'analyse des erreurs (voir 2.1). C'est ainsi que les linguistes commencent à s'intéresser davantage aux erreurs et que naît la recherche sur les interlangues (Marquilló Larruy, 2003).

2.3.3 L'INTERLANGUE

Le concept d'*interlangue* est proposé par Selinker (1972) et, comme le souligne Vogel (1995), sera maintes fois désigné par d'autres appellations (impliquant parfois des nuances théoriques) : *système intermédiaire* (Py, 1980), *dialecte idiosyncrasique* (Corder, 1981), *langue de l'apprenant* (Vogel, 1995)⁴⁶, etc. L'interlangue est donc un système linguistique propre à l'apprenant⁴⁷, dont les règles partagent des caractéristiques à la fois avec le système de sa L1 et celui de la langue cible (Corder, 1981 ; Gaonac'h, 1982), même si certaines des règles n'ont leur origine dans aucune des deux langues (Gass et Selinker, 2001: 12).

Selon Ellis (1994), la théorie de l'interlangue est particulièrement importante en ce sens qu'elle est la première à avoir tenté de fournir une explication des processus d'acquisition d'une langue seconde. Elle adopte cette approche « diachronique » qui faisait défaut à l'AE; il importe donc de bien distinguer *une interlangue* – qui constitue un « état de langue », un système linguistique construit par un apprenant à un moment précis de son apprentissage – de *l'interlangue* – qui est en fait le continuum sur lequel se positionne les

⁴⁵ Pour reprendre un concept proposé par Saussure (1960), nous pourrions dire que l'AE privilégie une approche synchronique.

⁴⁶ Comme le mentionne J.-P. Confais dans son avant-propos à l'ouvrage de Vogel, qu'il a traduit, le fait d'utiliser *langue de l'apprenant* (« *Lernersprache* ») permet de « bien marquer une évolution entre une approche de départ plutôt linguistique (l'interlangue comme *système* « entre deux langues ») et une approche plus psychologique et psycholinguistique... » (Confais dans Vogel, 1995: 12)

⁴⁷ La notion d'interlangue rejoint en quelque sorte celle d'idiolecte, c'est-à-dire l'ensemble des usages d'une langue propres à un individu donné (Dubois, 1994), voir notamment Labov (1972).

différentes interlangues (Ellis, 1994: 350). C'est par l'observation de ce continuum que des hypothèses peuvent être faites à propos de l'évolution de la langue de l'apprenant. Chaque interlangue est donc un système linguistique transitoire partageant plusieurs règles avec le précédent, mais présentant aussi de nouvelles règles et d'autres modifiées à la suite de l'invalidation de certaines hypothèses; c'est donc un système par essence « instable et fluctuant » (Berthoud et Py, 1993: 60)⁴⁸.

Un des concepts centraux de la théorie de l'interlangue est celui de *fossilisation*, qui correspond en quelque sorte à l'arrêt de l'apprentissage (Gass et Selinker, 2001) et à la stabilisation du système linguistique de l'apprenant, système qui comprendra à la fois des règles correspondant à celles de la langue cible et des items « fossilisés » (générateurs d'erreurs en regard de la langue cible) qui continueront de marquer la production de l'apprenant en dépit de son âge ou de la quantité d'explications ou d'instruction qu'il recevra dans cette langue (Selinker, 1972).

La démarche de l'interlangue ayant pour objet l'élucidation des processus en jeu dans l'apprentissage d'une langue, elle tient compte de l'ensemble des énoncés produits, et non seulement des énoncés erronés, pour faire émerger le système linguistique construit par l'apprenant (Corder, 1981 ; Gass et Selinker, 2001). Selon Corder (1981), il existe trois types d'énoncés: ceux qui sont manifestement erronés (*overtly erroneous*), ceux qui sont bien formés mais qui ne conviennent pas au contexte (*covertly erroneous*) et ceux qui sont bien construits et appropriés au contexte. Il faut donc considérer autant les dimensions formelles que contextuelles avant de poser un jugement, et, dans le cas d'énoncés corrects et appropriés, il ne faut pas exclure la possibilité qu'ils aient été produits par chance, par exemple par l'application d'une règle fautive qui pourrait tout aussi bien générer des énoncés incorrects.

⁴⁸ Il est à noter que le behaviorisme ayant servi de toile de fond à l'analyse contrastive est remplacé par une conception très cognitiviste dans la théorie de l'interlangue.

Les travaux portant sur l'interlangue se distinguent donc de ceux menés dans le cadre de l'AE tant par leur visée, celle d'éclairer les processus en jeu dans l'acquisition d'une langue seconde, que par les données qui sont analysées, soit l'ensemble de la production des apprenants et non seulement les énoncés « erronés ». Dans cette approche, la notion même d'erreur devient problématique dans la mesure où l'apprenant est le seul locuteur de l'interlangue observée et que les énoncés qu'il produit sont conformes aux règles propres à ce système linguistique transitoire et ne constituent des erreurs qu'en référence aux normes de la langue cible. Ces écarts par rapport à la L2 seront malgré tout étudiés pour déterminer s'ils trouvent leur source dans une interaction entre la langue maternelle et l'interlangue (*erreurs interlinguales*) ou à l'intérieur même de ce système (*erreurs intralinguales*). L'analyse des erreurs devient donc, dans les travaux sur l'interlangue, un moyen parmi d'autres d'élucider les processus d'acquisition d'une langue seconde, puisque l'observation d'erreurs récurrentes et systématiques permet de mettre en lumière les règles de l'interlangue de l'apprenant.

CONCLUSION

La figure ci-dessous, adaptée de James (1998), permet de situer les différents courants d'analyse d'erreurs.

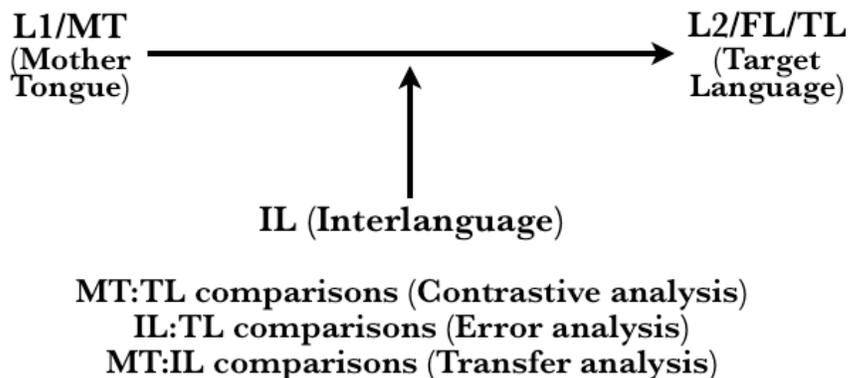


Figure 4. Les différents courants de l'analyse d'erreurs (James, 1998)

Ce bref survol historique des courants s'étant systématiquement intéressés à l'erreur au cours des dernières décennies nous permet certaines observations.

1. Il semble généralement admis que l'erreur mérite une attention particulière, tant didactiquement que linguistiquement, parce qu'elle est révélatrice de l'évolution des savoirs en construction chez l'apprenant et des lacunes à combler.
2. Il ressort que, malgré le potentiel informatif de l'erreur, son analyse ne suffit pas en elle-même à fournir un portrait global de la maîtrise linguistique d'un individu; elle doit être appréhendée dans le cadre plus global de l'ensemble de la production de l'apprenant.
3. Toutes les erreurs relevées n'ont pas la même valeur du point de vue de l'élucidation des processus en jeu dans l'apprentissage (lapsus, faute, etc.).
4. Les chercheurs semblent unanimes sur deux points : la difficulté de catégoriser les erreurs et le caractère hypothétique des sources des erreurs.

Finalement, un dernier constat s'impose : les courants de recherches s'étant intéressés de façon systématique à l'erreur se sont concentrés presque exclusivement à l'acquisition d'une L2, l'accent étant mis sur les erreurs interlinguales (la Figure 4 en témoigne). Par conséquent, les informations dont nous disposons sur les erreurs intralinguales sont limitées et ne proviennent pas toujours de travaux s'étant inscrits clairement dans l'un des courants d'analyse d'erreurs décrits. D'un point de vue méthodologique, nous constatons aussi que les approches présentées tirent essentiellement leurs données de corpus écrits (Gass et Selinker, 2001) et s'intéressent donc principalement aux erreurs de production plutôt qu'à celles de compréhension (Corder, 1974).

3. APPORTS DE CES TRAVAUX POUR L'ANALYSE D'ERREURS LEXICALES

Les différentes approches que nous venons de présenter portaient sur l'analyse d'erreurs de langue en général, mais ont surtout permis des observations intéressantes sur le plan morphosyntaxique. Même si elles ont peu porté sur les aspects lexicaux, ces approches offrent des pistes intéressantes pour la catégorisation et la compréhension d'erreurs lexicales, que nous explorerons dans les sections qui suivent. Pour des raisons d'organisation de notre cadre conceptuel, les pistes léguées par les courants décrits plus haut relativement à la

définition même du concept d'erreur lexicale seront abordées à la section IV-1. Nous nous intéressons ici aux différentes étapes de démarche d'analyse elle-même ainsi qu'à la question de l'évaluation de la gravité des erreurs.

3.1 LES ÉTAPES DE L'ANALYSE D'ERREURS

L'analyse d'erreurs comporte plusieurs étapes que plusieurs auteurs ont tenté de décrire (Corder, 1974 ; Corder, 1981 ; Gass et Selinker, 2001 ; Marquilló Larruy); l'un des apports importants des travaux réalisés jusqu'à maintenant pour l'analyse d'erreurs lexicales concerne donc la démarche à suivre. Une comparaison des démarches proposées par les auteurs nous permet de dégager une approche qui semble faire consensus.

1. **Constitution d'un corpus d'analyse** : Cette étape, mentionnée explicitement par quelques auteurs seulement, est bien sûr sous-entendue par les autres. Nous n'avons pas grand-chose à en dire, sinon que le choix du corpus étudié pourra avoir une influence sur la fréquence et les types d'erreurs relevées (Corder, 1974).
2. **Repérage de l'erreur** : La question de l'interprétation des énoncés contenant des erreurs, cruciale à cette étape, sera discutée plus bas (3.1.1).
3. **Description/classification de l'erreur** : Cette étape implique des critères de classification, souvent opérationnalisés à l'intérieur de typologies (voir 3.1.2).
4. **Explication de l'erreur** : Il s'agit ici d'établir les sources possibles de l'erreur observée. Ce sujet important fera l'objet du point 3.1.3.
5. **Remédiation/prolongements pédagogiques** : Cette étape concerne la dimension pronostique de l'analyse d'erreurs (voir 2.1) et sera développée au point 3.1.4.

Il est étrange de constater qu'un point crucial, du moins dans une perspective didactique, n'apparaît pas dans les étapes proposées par les auteurs : l'évaluation. Cette notion, qui est pourtant très intimement liée à l'erreur, fait partie de la réflexion de tout enseignant confronté à un problème lors de la correction. Nous verrons ce que nous pouvons en dire relativement aux erreurs lexicales au point 3.1.5 ainsi qu'à la section 3.2.

Une autre étape importante de l'analyse d'erreurs n'est mentionnée que par Gass et Selinker (2001) : le traitement quantitatif des données (fréquence et

proportion des différents types d'erreurs). Ellis (1994) souligne d'ailleurs que cette étape fait défaut dans plusieurs recherches menées dans le cadre de l'AÉ; Porquier (1977) dénonce quant à lui l'absence de prise en compte de la dispersion dans l'analyse de certains chercheurs. Pourtant, la prise en considération de l'aspect quantitatif des erreurs aura selon nous un impact tant sur l'évaluation que sur les prolongements pédagogiques, voire sur l'explication des erreurs.

La démarche d'analyse présentée plus haut semble convenir dans le cas d'erreurs lexicales. Dans une recherche portant sur l'analyse d'erreurs lexicales d'étudiants libanais apprenant le français comme langue seconde (FLS), Mroue (2005) propose une démarche qui correspond parfaitement à celle présentée plus haut, sinon que l'évaluation y est explicitement mentionnée. Chez Hamel et Milićević (2007), qui s'intéressent aux erreurs lexicales d'apprenants universitaires du FLS, les étapes sont sensiblement les mêmes, sauf l'étape d'explication de l'erreur qui n'apparaît pas telle quelle (elle est en quelque sorte amalgamée à l'étape de description, ce qui peut s'avérer problématique) et ne semble pas systématiquement donner lieu à des hypothèses.

3.1.1 LE REPÉRAGE DES ERREURS

Cette étape en apparence évidente de l'analyse d'erreurs est pourtant très complexe et repose en grande partie sur l'interprétation du « correcteur »⁴⁹, ce qui implique une certaine subjectivité, particulièrement dans le cas de matériel écrit, puisqu'il est souvent impossible de consulter le scripteur lors de la correction.

Pour paraphraser Corder (1981: 37), repérer une erreur consiste en fait à comparer l'énoncé produit par le scripteur avec celui qu'il aurait dû produire pour exprimer ce qu'il voulait exprimer (de la même façon que l'aurait exprimé un locuteur expert, voir 1.2). Pour générer cette phrase « idéale » servant de

⁴⁹ Nous entendons par *correcteur* toute personne s'intéressant aux erreurs présentes dans un texte (enseignant, linguiste, didacticien, etc.).

point de comparaison, que Corder nomme *reconstruction*⁵⁰, il faut donc inférer le message que voulait exprimer le scripteur à partir des structures produites et du contexte dans lequel elles apparaissent (Corder, 1974). Corder distingue deux types de reconstructions : les reconstructions sûres (*authoritative reconstruction*) – formulées après avoir consulté le scripteur – et les reconstructions plausibles – résultant d’inférences de la part du correcteur. Toute interprétation du sens faite sans vérification auprès du scripteur devra donc être considérée comme une hypothèse et tout diagnostic d’erreur posé sur la base de cette reconstruction sera nuancé en conséquence; la reconstruction choisie influencera bien sûr la description de l’erreur (Ellis, 1994). Cette mise en garde, soulevée par plusieurs auteurs (Corder, 1981 ; Mastin, 2001 ; Marquilló Larruy), sera particulièrement importante dans le cas d’erreurs lexicales sémantiques, puisque la part de subjectivité impliquée dans la formulation de la reconstruction sera souvent plus grande que celle rattachée à d’autres types d’erreurs, celles relevant de l’orthographe lexicale par exemple. Le fait d’avoir recours à plusieurs correcteurs peut diminuer la part de subjectivité en confrontant les différentes interprétations (Read, 2000).

3.1.2 LA DESCRIPTION/CLASSIFICATION DES ERREURS

La description des erreurs relevées lors du repérage consiste à comparer l’énoncé produit par le scripteur avec sa reconstruction pour bien observer ce qui les distingue. Cette étape de description est cruciale dans le processus d’analyse d’erreurs et doit nécessairement avoir lieu en amont de la formulation d’hypothèses sur les sources d’erreurs (Cavalli, Duchêne, Elmiger, Gajo, Marquilló Larruy, Matthey, Py et Serra). Plusieurs critères peuvent être utilisés pour décrire les erreurs et établir des classifications. Nous présenterons ici quelques-unes des oppositions binaires souvent utilisées.

- **Erreur relative/absolue** : Cette distinction, proposée par Debyser *et coll.* (1967), visait à distinguer les erreurs impliquant une forme (ou

⁵⁰ Milićević et Hamel (2007) parlent plutôt de *correction*, alors que Debyser *et al.* (1967) utilisent *restitution*, terme qui nous plait particulièrement, puisqu’il véhicule plus clairement l’idée de rétablissement d’un message déjà existant.

structure) inexistante des erreurs présentant une forme correcte, mais incompatible avec le contexte. La pertinence de cette dichotomie a été par la suite remise en cause, entre autres par Porquier (1977), qui mentionne qu'elle semble sous-entendre que les erreurs absolues sont plus graves, alors qu'elles ont souvent une incidence moins importante sur la compréhension du message. En ce qui concerne l'erreur lexicale, cette opposition serait opérationnelle pour l'orthographe lexicale, mais pourrait être plus difficilement applicable pour des problèmes de nature sémantique.

- **Erreur graphique/orale** : Cette deuxième opposition, aussi présente dans la grille d'analyse développée par Debyser *et coll.* (1967), semble tout à fait pertinente pour la caractérisation d'erreurs liées au signifiant des mots. Cette distinction apparaît d'ailleurs dans la typologie d'erreurs de Catach (Catach, Gruaz et Duprez, 1986) et dans la nôtre (2005).
- **Classable/inclassable** : Mentionné notamment par Debyser *et coll.* (1967), ce critère de classification est crucial pour une analyse qui aspire à éclairer le comportement linguistique d'un apprenant. Il existe deux types d'erreurs inclassables : 1) celles apparaissant dans un énoncé complètement ininterprétable (Corder, 1981) et pour lesquelles il est impossible de fournir une reconstruction et 2) celles qui sont non prévues par la typologie utilisée (Debyser *et coll.*, 1967). Les premières doivent être exclues de l'analyse, alors que les secondes sont mises de côté afin de voir si elles nécessitent l'ajout d'une classe d'erreurs supplémentaire.
- **Erreur interlinguale/intralinguale** : Cette dichotomie est devenue particulièrement importante dans le cadre des recherches sur l'interlangue, alors que les chercheurs tentaient de déterminer si les erreurs observées avaient pour source des interférences entre L1 et L2 ou témoignaient plutôt d'erreurs à l'intérieur même de l'interlangue. Si nous ne remettons pas en cause la pertinence de cette opposition dans la compréhension d'erreurs, nous ne croyons pas qu'elle doive faire partie d'une description à priori des erreurs rencontrées. En effet, ces étiquettes marquent une position par rapport à la source de l'erreur, alors que c'est à l'étape de description que nous nous intéressons présentement; ces étiquettes seront utilisées une fois que des hypothèses auront été formulées à propos des sources des erreurs et pourront servir d'élément classifiant pour regrouper les erreurs après leur analyse. S'il ne fait aucun doute que ces critères s'appliqueront aussi aux erreurs lexicales, il est important de souligner que l'opposition sera surtout pertinente avec des apprenants d'une L2, ou encore pour des locuteurs évoluant dans des environnements où cohabitent plusieurs langues. Malgré le bienfondé d'utiliser cette opposition, Besse et Porquier (1984) font remarquer qu'il n'est pas toujours facile de classer une erreur dans l'une ou l'autre des catégories; pour pallier cette difficulté, Granger et Monfort (1994) établissent une distinction entre les erreurs interlinguales ne s'expliquant

que par le transfert et celles pour lesquelles le transfert n'est qu'un des scénarios d'explication possibles.

- **Isolée/récurrente** : Cette opposition est particulièrement intéressante dans l'optique d'une recherche qui tente de déterminer ce qui fait système chez l'apprenant, de trouver les « règles » qui régissent son interlangue (Marquilló Larruy, 2003). Par contre, à cause du caractère très ponctuel de l'erreur lexicale et de la grande variété de mots utilisés dans le discours (évidemment plus élevé que le nombre de constructions grammaticales), nous ne croyons pas qu'un nombre important d'erreurs lexicales récurrentes seront relevées et que, si c'est le cas, elles témoigneront plutôt de mauvaises connaissances déclaratives (ex. méconnaissance du sens d'un mot) que d'un véritable système transitoire. En effet, comme le mentionnent Granger et Monfort (1994) à propos des erreurs lexicogrammaticales, les erreurs lexicales sont liées à l'emploi d'une unité lexicale précise, contrairement aux « erreurs proprement grammaticales qui sont, quant à elles, de l'ordre du régulier, du généralisable... » (p.63). Ceci rejoint les propos de Patris et Vansnick (1992: 50) qui soulignent aussi que les « fautes » lexicales « ne sont pas généralisables à des classes d'unités; la « règle » transgressée ne peut être formulée en termes de catégories, mais seulement en nommant l'unité concernée. »

Nous venons d'observer certains critères généraux de classification des erreurs en nous interrogeant sur leur pertinence dans l'analyse d'erreurs lexicales. Bien sûr, des classements portant plus spécifiquement sur les erreurs lexicales ont aussi été proposés par les chercheurs. Comme ces classements constitueront les assises théoriques de nos propres typologies des erreurs lexicales, nous les aborderons en détail aux sections 2 et 3 du chapitre V.

Dans tous les classements proposés pour l'analyse d'erreurs, certaines catégories semblent se chevaucher ou se recouper, et un même problème peut, selon le point de vue, être décrit de façon différente. Outre le rôle important d'une certaine subjectivité liée à l'interprétation, cette difficulté découle sans doute du fait que plusieurs niveaux d'analyse sont présents dans les typologies mentionnées, d'où l'intérêt de la proposition de Porquier (1977) d'utiliser, plutôt qu'une seule grille hétérogène, une série de grilles d'analyse complémentaires permettant des recoupements et des regroupements divers. Cette difficulté à bien décrire les erreurs ne doit cependant pas freiner notre volonté de

catégorisation, puisque l'important est de stimuler notre réflexion sur l'erreur (Patris et Vansnick, 1992 ; Mastin, 2001).

3.1.3 L'EXPLICATION DES ERREURS

Cette étape est en quelque sorte le cœur de la démarche d'analyse, puisque c'est celle qui, dans une perspective psycholinguistique, permettra d'éclairer les processus cognitifs en jeu chez l'apprenant et, dans une visée didactique, orientera les interventions de l'enseignant. Nous avons déjà évoqué la part de subjectivité liée à l'interprétation pour la restitution du sens; l'interprétation de la source des erreurs laisse une place encore plus grande à la spéculation. Tout d'abord, une erreur peut avoir plusieurs sources et une même erreur peut avoir des causes différentes selon la personne qui la commet (Stirman-Langlois, 1995). De plus, il est difficile d'avoir accès aux cheminements cognitifs ayant mené l'apprenant à commettre une erreur, d'autant plus que certaines causes possibles ne seront même pas envisagées par le correcteur, qui les « filtre » en quelque sorte par sa vision de la langue; Berthoud (1987: 15) parle d'« adultocentrisme ». Pour valider les sources d'erreurs supposées à la lumière des indices fournis par la production d'un apprenant – les *données textuelles* selon Porquier (1977) – il est recommandé de faire appel à ce dernier afin d'avoir accès à d'autres informations (Cain, 1987 ; Mastin, 2001) que l'on nomme *données intuitives*. Porquier (1977) suggère trois façons de générer des données intuitives : 1) par des commentaires métalinguistiques, 2) par le biais d'une tâche de production contraignante poussant le sujet à utiliser les structures problématiques et 3) par des jugements de grammaticalité. Même s'il n'est pas toujours possible d'établir avec assurance la source d'une erreur, particulièrement à partir de données textuelles seulement, il est important de formuler une hypothèse, puisque celle-ci peut « apparaître comme une indication annexe supplémentaire pour celui qui aurait suivi un autre cheminement. » (Demaizière, 1987: 52) Mais quelles sont donc les hypothèses à avancer pour expliquer les erreurs linguistiques ?

Frei (1929) suggère que les erreurs de langue trouvent leur source dans des « tentative[s] de réparation des aberrations et des déficits de la langue normée » (Bertucci, 1997). Si des formes erronées sont produites, c'est pour répondre à deux grands besoins régissant la production linguistique : l'économie (analogie, invariabilité) et la diversification (différentiation, expressivité).

Ainsi, l'interférence linguistique, liée à ce besoin d'économie, est la source première d'explication des erreurs de langue. Les travaux sur l'analyse d'erreurs se sont surtout intéressés aux erreurs interlinguistiques, c'est-à-dire celles qui surviennent sous l'influence de la L1 (ou d'une autre langue étrangère) alors qu'un locuteur utilise une L2⁵¹. Par contre, les courants suivant l'analyse contrastive ont permis de faire ressortir le rôle des interférences intralinguistique dans le discours d'apprenants, même lorsque ces derniers s'expriment en L2. Voici la définition que propose Polguère (2007: 257) de l'interférence intralinguistique :

Interférence intralinguistique. *Une interférence intralinguistique est un phénomène de parole (au sens saussurien) qui possède les deux caractéristiques suivantes :*

1. *c'est un dysfonctionnement linguistique qui se manifeste par l'usage au sein d'un énoncé en langue **L** d'un élément linguistique (lexical ou grammatical) e_1 qui est soit impropre en contexte soit mal construit ;*
2. *on peut expliquer l'usage que fait le locuteur de e_1 par l'existence dans la même langue **L** d'un élément linguistique e_2 qui présente certaines **analogies** avec e_1 et qui aurait parasité la production du locuteur lors de son accès aux règles linguistiques de **L**.*

Évidemment, il suffirait de spécifier que e_2 , sous l'influence duquel e_1 est produit, appartient à une langue **L₂** pour obtenir une définition de l'interférence interlinguistique (Polguère, 2007). Cette façon de définir l'interférence permet de mettre en évidence le rôle prépondérant de l'analogie, manifestation du principe d'économie de Frei (1929), dans la production d'erreurs linguistiques; l'analogie sera à la base de nombreux scénarios d'explication d'erreurs.

⁵¹ Polguère (2007: 7) souligne d'ailleurs que « le terme *interférence linguistique* est en fait presque toujours utilisé pour signifier 'interférence interlinguistique' ». C'est d'ailleurs ce que l'on constate en consultant (Legendre, 2005), qui définit l'interférence linguistique comme un « Phénomène caractérisé par l'apparition, dans l'utilisation d'une langue, d'éléments ou de structures d'une autre langue. » (p.799)

Bien entendu, l'opposition erreur intralinguale/erreur interlinguale⁵² est omniprésente dans les travaux sur l'analyse d'erreurs, puisque ces derniers s'intéressent à l'explication des erreurs commises par des apprenants d'une L2. Par exemple, Granger et Monfort (1994), dans leur recherche sur la description de la compétence lexicale en langue étrangère, fondent leur classification des sources d'erreurs lexicales sur l'opposition entre les erreurs intralinguales (ou développementales) et les erreurs interlinguales. Ils font par contre une mise en garde intéressante en rappelant « qu'une proportion importante des erreurs permettent les deux interprétations. » (p.68) Ils divisent ainsi, à la suite de Sheen (1988) et Singleton (1990), les erreurs interlinguales en deux catégories : celles qui ne s'expliquent que par le transfert et celles pour lesquelles le transfert ne constitue qu'une explication parmi d'autres. Cette division entre les erreurs interlinguales est d'autant plus intéressante qu'elle fait bien ressortir le fait qu'il n'est pas toujours possible d'établir la source d'une erreur. Selon James (1998), il est même plutôt rare que l'on puisse clairement associer avec assurance une erreur avec une seule source. Pour lui, les erreurs sont généralement soit des erreurs « mixtes » (*compound errors*) – c'est-à-dire qu'elles ont plusieurs sources qui agissent de façon simultanée ou cumulative –, soit des erreurs ambiguës (*ambiguous errors*) – c'est-à-dire que les scénarios d'explication envisagés sont en compétition plutôt qu'en interrelation. Ces deux cas de figure, qui nous poussent à nuancer l'explication d'une erreur leur correspondant, s'ajoutent au fait que l'étape de restitution qui précède le diagnostic des erreurs est souvent faite à partir de données textuelles seulement et que plus d'une correction est parfois envisageable.

L'opposition erreur intralinguale/erreur interlinguale est certes cruciale pour expliquer les erreurs linguistiques, mais James (1998) suggère de considérer

⁵² Nous parlons ici d'*erreur* plutôt que d'*interférence*, puisque l'interférence résulte en une erreur et que ce dernier terme constitue la pierre angulaire de notre travail. Par contre, il aurait été possible d'utiliser le terme *interférence* comme synonyme d'*erreur* dans ce contexte, puisque INTERFÉRENCE, par métonymie, peut dénoter soit le phénomène linguistique ou le résultat de ce phénomène (cf. Legendre, 2005: 799).

aussi deux autres classes d'erreurs : celles relatives à une stratégie communicative et les erreurs induites.

Les erreurs relatives à une stratégie communicative découlent de l'idée que si un item de la langue n'est pas disponible, on peut toujours s'en sortir en en utilisant un autre, ce qui résulte souvent en une approximation. James donne en exemples l'utilisation d'un quasi-synonyme, d'un hyperonyme ou d'un antonyme⁵³, l'invention d'un mot ou l'emprunt direct à une autre langue ainsi que la traduction littérale. Il mentionne aussi le recours aux stratégies analytiques de paraphrase et de périphrase pour pallier, notamment, un manque d'accès au vocabulaire.

Pointer le recours à une stratégie communicative pour expliquer certaines erreurs nous semble extrêmement intéressant et permet d'établir une distinction avec certaines erreurs interlinguales. Par exemple, une même erreur pourrait être expliquée par une confusion entre un item de la L1 et un item de la L2 sur la base d'une parenté formelle (erreur interlinguale) ou par une simple traduction littérale (stratégie communicative). James donne l'exemple de **My watch doesn't **walk** well.*, où *walk* peut avoir été choisi soit sur la base d'une parenté formelle avec *work* (prononcé à la British) au niveau phonologique, soit parce que le scripteur francophone qui a produit cet énoncé a simplement traduit le verbe MARCHER (au sens de 'fonctionner') de sa langue maternelle parce qu'il ignorait le terme de la L2 correspondant à ce sens; selon ce deuxième scénario d'explication, aucune confusion interlinguale n'est en jeu. Il se peut aussi que ces deux sources aient agi en synergie et se soient en quelque sorte validées l'une l'autre. Bien entendu, il sera difficile d'établir avec assurance qu'une stratégie communicative est à la source d'une erreur sans parler directement à la personne l'ayant commise.

Les erreurs induites mentionnées par James ont quant à elle pour origine une source extérieure. Il pourrait par exemple s'agir d'une explication imprécise de l'enseignant qui expliquerait le sens d'un mot en en donnant un quasi-synonyme,

⁵³ Par exemple, un élève qui écrirait *pas aimable* plutôt qu'*antipathique*.

ce qui pourrait conduire l'apprenant à penser que les deux termes présentent un fonctionnement identique et véhiculent exactement le même sens⁵⁴. Les usages linguistiques dans l'environnement d'un locuteur peuvent aussi être à l'origine d'erreurs; par exemple, si le terme AVION est systématiquement utilisé au féminin dans l'entourage linguistique d'un locuteur, il y a fort à parier que celui-ci commettra aussi l'erreur. De la même façon, un enseignant de L2 n'ayant pas lui-même cette langue pour langue maternelle pourrait conduire ses élèves à produire une langue déviante d'un point de vue phonologique, puisque ces derniers risquent de « copier » son accent.

Au-delà de ces grandes « familles » d'erreurs, plusieurs auteurs se sont intéressés à déterminer plus précisément des sources possibles d'erreurs. Nous présenterons ici celles qui nous semblent les plus pertinentes en ce qui a trait à l'erreur lexicale.

Astolfi (1997) établit une typologie des erreurs en fonction de leur source; des huit catégories proposées, cinq nous semblent s'appliquer à l'erreur lexicale en production :

1. *erreurs témoignant des **conceptions alternatives** des élèves* : celles qui découlent de mauvaises connaissances déclaratives à propos d'UL particulières. Ces conceptions alternatives peuvent elles-mêmes découler d'autres sources, par exemple de l'analogie.
2. *erreurs dues à une **surcharge cognitive***, qui seraient en d'autres circonstances évitées⁵⁵ (cf. opposition *erreur* ~ *faute*, voir 1.5).
3. *erreurs ayant leur origine dans une **autre discipline***, ou dans une autre langue. Il s'agit des erreurs de transfert.
4. *erreurs causées par la **complexité propre du contenu***, par exemple pour les mots à l'orthographe compliquée, pour les verbes irréguliers, etc.
5. *erreurs liées aux **opérations intellectuelles impliquées***, par exemple dans le cas de la consultation d'un dictionnaire ou de l'analyse syntaxique d'une phrase en vue de l'application d'une règle.

⁵⁴ Martin (1984) pointe d'ailleurs l'enseignement du vocabulaire à l'aide de gloses ou de synonymes comme la source principale d'erreurs lexicales chez les apprenants avancés d'une L2.

⁵⁵ Ceci rejoint la notion d'erreur de performance.

Une autre source possible d'erreurs est un problème d'accès aux items lexicaux permettant d'exprimer un sens donné, ce que l'on pourrait appeler un « trou lexical ». Ce trou lexical peut être propre au lexique mental d'une personne, qui ne possède pas certains mots, ou encore inhérent à une langue particulière (paradigme défectif). Un tel problème d'accès à des UL pousse le locuteur à utiliser, plutôt que le terme propre, un terme sémantiquement proche mais inapproprié dans le contexte ou encore une paraphrase maladroite, ou *forme analytique* (Milićević, 2008); il peut aussi être à l'origine de collocations douteuses (Polguère, 2007). En d'autres mots, un tel problème d'accès pousse le locuteur à faire appel aux stratégies communicatives de James (1998) évoquées plus haut. Levenston et Blum (1976) se sont intéressés à la simplification lexicale, notamment dans le discours de locuteurs de L2, qui sont souvent confrontés à des « trous lexicaux ». Les auteurs dressent une liste de stratégies déployées dans pareilles situations; certaines sont ponctuelles (paraphrase, changement de langue, demande d'aide à l'interlocuteur, changement de sujet, etc.) alors que d'autres peuvent initier un processus qui mènera à une fossilisation, c'est-à-dire à un emploi lexical systématique et approximatif pour exprimer un sens précis (emploi d'un hyperonyme, d'un synonyme approximatif, d'un terme inventé, d'un conversif, etc.).

Ellis (1994) mentionne, en s'inspirant des travaux de Richards (1971) et de Lott (1983), quelques sources d'erreurs, qui correspondent en majorité au besoin d'économie mentionné par Frei (1929). Certaines erreurs interlinguales, par exemple, peuvent être attribuées à une mauvaise utilisation d'une lexie de la L2 en fonction d'une ressemblance avec une lexie de la L1 (analogie); il est par contre important de souligner que l'analogie formelle peut aussi être à la source d'erreurs intralinguales (cf. paronymie). Le fait qu'il n'existe pas toujours de correspondance univoque entre les UL de langues différentes peut aussi être à l'origine d'erreurs lexicales, si par exemple certaines distinctions sémantiques sont occultées dans l'une ou l'autre des langues en jeu. Ellis parle, en ce qui concerne les erreurs intralinguales, de surgénéralisation (cf. besoin d'économie);

ceci pourrait conduire à transposer certaines propriétés d'une UL à une autre à laquelle elles ne s'appliquent pas. L'application de règles à des contextes inappropriés rejoint aussi ce besoin d'économie, mais s'avère peut-être moins pertinente pour ce qui est des erreurs lexicales, à part peut-être dans le cas de formes erronées découlant de l'application fautive de règles de dérivation morphologique.

En résumé, l'analogie constitue une clé importante dans la compréhension des erreurs lexicales, tout comme l'idée d'*interférence interlinguistique* proposée par l'analyse contrastive. Avoir une bonne connaissance du parcours linguistique de l'apprenant (langues parlées, environnement linguistique) peut aussi orienter les hypothèses sur les sources possibles des erreurs. Nous ferons à la section V-4 un survol des sources d'erreurs possibles présentées plus précisément dans les travaux s'intéressant à l'erreur lexicale.

3.1.4 LE TRAITEMENT DES ERREURS

Le traitement des erreurs désigne « [la façon] dont l'enseignant réagit aux erreurs produites par les apprenants, à savoir si l'erreur doit être corrigée, quand, comment et par qui. » (Legendre, 2005: 1401). La capacité à traiter les erreurs fait partie des compétences fondamentales de l'enseignant (Corder, 1981).

Pour ce qui est des erreurs à corriger, les travaux sur l'analyse d'erreurs ne proposent pas de pistes précises et ont tendance à les considérer toutes. Malgré tout, certains auteurs suggèrent de corriger en priorité les erreurs directement en lien avec les objectifs de l'activité et celles dont la fréquence est élevée (Stirman-Langlois, 1995 ; Marquilló Larruy, 2003), surtout dans le cas où le nombre d'erreurs est grand. Bien entendu, il est aussi important de considérer le niveau scolaire des apprenants et de corriger principalement les aspects du code qu'ils sont censés maîtriser. Comme il n'existe pas de cursus clair en ce qui concerne le vocabulaire à maîtriser, nous sommes d'avis que les erreurs lexicales en production devraient toutes être relevées (mais pas nécessairement

sanctionnées), dans la mesure où le fait d'utiliser une UL en indique une maîtrise au moins partielle (Shaughnessy, 1977 ; Henriksen, 1999).

Pour les erreurs lexicales en production, le moment de la correction est obligatoirement après l'écriture, alors que l'enseignant a relevé les erreurs, soit de façon implicite – seulement par un trait –, soit plus explicite – par une annotation plus ou moins complète présentant ou non une correction. Les modalités de correction, ensuite, peuvent varier énormément⁵⁶, allant de la simple remise de la copie annotée à l'apprenant (en se fiant à sa bonne volonté pour pousser plus loin sa réflexion) à un travail collectif mené oralement en classe; Marquilló Larruy (2003: 117) suggère d'ailleurs de diversifier les méthodes d'intervention.

Une approche de correction issue des courants de l'analyse d'erreurs a été proposée par Lamy (1976) : la *pédagogie de la faute*. Cette approche s'adresse à des apprenants d'une L2 d'un niveau avancé⁵⁷ et propose de bâtir des interventions à partir d'erreurs commises par les élèves. Cette démarche, menée en grand groupe, peut être réduite à cinq étapes principales :

1. observation de l'énoncé fautif (soit en ciblant l'erreur, soit en la laissant « découvrir »)
2. tentative de correction
3. génération, si possible, de phrases correctes dans lesquelles est maintenu l'élément fautif (dans le but de voir s'il existe ailleurs dans la langue, s'il s'agit d'une *erreur relative*)
4. regroupement et contraste des phrases correctes et incorrectes pour dégager les conditions d'emploi de l'élément fautif
5. production de phrases correctes contextualisées contenant l'élément problématique

Lamy mentionne qu'il est important de ne pas travailler sur des fautes relevant d'un micro-système de la langue avec lequel l'élève est encore peu familier et souligne la nécessité de toujours signaler (par un astérisque par exemple) le caractère fautif d'un énoncé. Cette approche intéressante s'applique par contre difficilement à l'erreur lexicale, puisque rares sont les dimensions lexicales

⁵⁶ Voir Marquilló Larruy (2003) pour un tableau récapitulatif des approches de correction.

⁵⁷ Bien que Lamy (1984) relate aussi une expérience avec des apprenants débutants de FLE.

faisant réellement partie d'un microsystème qui se prêtent bien à un exercice de comparaison. Par contre, l'idée de contraster différentes phrases et celle de générer des phrases dans lesquelles l'item problématique devient acceptable pourraient être pertinentes pour le traitement d'erreurs sémantiques, notamment dans le cas d'une mauvaise sélection lexicale à l'intérieur d'un paradigme (ex. MAISON, CABANE, MANOIR, BICOQUE) ou de problèmes de cooccurrence, comme l'emploi de mauvais collocations auprès de synonymes (ex. **prendre un choix/prendre une décision*).

Une autre approche traitant spécifiquement des erreurs lexicales en L2 est mise de l'avant par Milićević et Hamel (2007) : la reformulation. Partant du principe que la compétence à paraphraser est au cœur de la compétence linguistique (Milićević, 2008), les auteures proposent l'utilisation par les apprenants d'un dictionnaire de reformulation qui, en mettant en lumière les relations entre les UL ainsi que certaines règles de paraphrasage⁵⁸, permettrait à ces derniers de transformer leurs énoncés erronés en phrases correctes. Si l'appareillage conceptuel nécessaire à l'utilisation d'un tel ouvrage nous permet de douter de son efficacité dans des classes de langues (ou du moins de la capacité des élèves à l'utiliser par eux-mêmes de façon spontanée), les principes qui en sous-tendent la conception nous paraissent d'excellentes pistes d'intervention pour le traitement des erreurs lexicales en production, et ce, autant pour les élèves de L1 que de L2. En mettant en évidence des patrons lexicaux récurrents ainsi que des « règles » lexicales productives, une approche qui s'apparente à celle proposée par Milićević et Hamel saurait sans doute développer des réflexes chez les élèves et leur permettre de mieux saisir les zones à risque dans l'utilisation du lexique. Il faut par contre être conscient qu'une telle approche pourra être efficace pour certains types d'erreurs (notamment sémantiques et collocationnelles), mais moins pour d'autres relevant des propriétés

⁵⁸ Par exemple le remplacement d'un verbe par le nom déverbal correspondant précédé d'un verbe support (*choisir* → *faire un choix*). Pour une introduction à la notion de règle de paraphrasage, voir Mel'čuk (1992).

idiosyncrasiques des UL comme le genre nominal ou les prépositions régies⁵⁹. Il est intéressant de noter que cette approche vise à rendre le traitement d'erreurs moins ponctuel et anecdotique en mettant en lumière certaines régularités lexicales, ce qui rejoint l'idée que l'enseignement des structures lexicales est plus important que celui des mots eux-mêmes (Polguère, 2004a). Dans un même ordre d'idées, Martin (1984) souligne la pertinence des fonctions lexicales dans l'enseignement du vocabulaire et dans le traitement de problèmes relatifs aux « dissonances collocationnelles ».

3.1.5 L'ÉVALUATION DES ERREURS

Voyons maintenant les pistes que nous donnent les travaux sur l'analyse d'erreurs sur leur évaluation, dimension cruciale pour tout enseignant de langue. Il est étrange de constater que la notion d'évaluation, pourtant très intimement liée à l'erreur, est très peu présente dans les travaux sur l'analyse d'erreurs (Johansson, 1978). Même ceux présentant des grilles d'analyse des erreurs ne donnent aucune indication quant à la façon de les opérationnaliser dans l'évaluation, ce qui nous laisse bien peu de pistes pour l'évaluation des erreurs lexicales. Comme le mentionne Tréville (2000), le vocabulaire dans un texte est souvent apprécié par un jugement global pour rendre compte de trois critères qualitatifs (justesse, précision, variété). Très rares sont les grilles de correction qui font mention des erreurs lexicales en distinguant différents types (impropriété, registre, emprunt, etc.); pourtant, tout porte à croire que la présence d'erreurs lexicales influence le score global attribué à une production écrite (Engber, 1995).

En bref, nous pouvons affirmer que les travaux sur l'analyse d'erreurs ne fournissent que peu de pistes pour l'évaluation des erreurs lexicales. Ellis (1994) souligne d'ailleurs que la recherche sur l'évaluation des erreurs ne permet pas à ce jour de proposer une échelle d'évaluation qui permettrait de les hiérarchiser. De façon générale, les erreurs devraient être évaluées d'abord par l'impact

⁵⁹ D'où l'intérêt d'un dictionnaire de reformulation, qui présenterait sans doute aussi ce type d'informations.

qu'elles ont sur la compréhension du correcteur, mais aussi en tenant compte d'autres critères comme la fréquence ou le degré d'irritation provoqué (Ellis, 1994); nous reviendrons sur les critères permettant d'évaluer la gravité d'une erreur au point suivant. En terminant, il est tout de même important de mentionner que certains travaux sur l'évaluation de la richesse lexicale d'un texte proposent des calculs qui tiennent compte des erreurs lexicales (Arnaud, 1992 ; Engber, 1995); par contre, ces travaux présentent des conceptions variables (voire réductrices) de l'erreur lexicale et, comme le souligne Read (2000), accordent le même poids à tous les types de problèmes lexicaux.

3.2 LA GRAVITÉ DES ERREURS

Un des rôles de l'évaluation est de permettre à l'enseignant de hiérarchiser les difficultés des élèves.

The main reason for evaluation [...] is to get our priorities right [...] to prevent obsession with trivial errors and give priority to the ones that really matter. Passing judgement of errors is not a matter of devaluation of learners and their language, but rather one of assigning relative values to errors. This may seem paradoxical : how can we give value to something that is undesirable? What it means is that this is an error that merits attention. (James, 1998: 204-205)

En effet, à la lecture d'un texte présentant des erreurs de langue, certaines paraissent spontanément plus « graves » ou « grosses », alors que d'autres sont jugées triviales, « petites » ou sans importance. Sur quelles bases reposent de tels jugements de gravité? C'est à cette question que nous tenterons de répondre en faisant un survol des différents facteurs pouvant influencer la perception de la gravité d'une erreur. Ce tour d'horizon s'inspire principalement de Johansson (1978) et de James (1998), que nous aborderons tour à tour.

3.2.1 LES CRITÈRES DE GRAVITÉ SELON JOHANSSON (1978)

Suivant l'idée que les erreurs les plus graves sont celles qui constituent une entrave à la communication (Burt et Kiparsky, 1972 ; Enkvist, 1973), Johansson place l'efficacité communicative comme principe de base de l'évaluation de l'erreur. L'auteur fait par contre bien ressortir que l'efficacité communicative ne se réduit pas à la compréhensibilité du message.

To be communicatively effective, the message must get across swiftly and unambiguously and without undue demands upon the receiver. Hence, the scheme of evaluation [...] includes both comprehensibility and degree of irritation. (Johansson, 1978: 6)

Johansson (1975) propose un système d'évaluation de la gravité des erreurs (schématisé ci-dessous) qui rend compte de l'efficacité communicative.

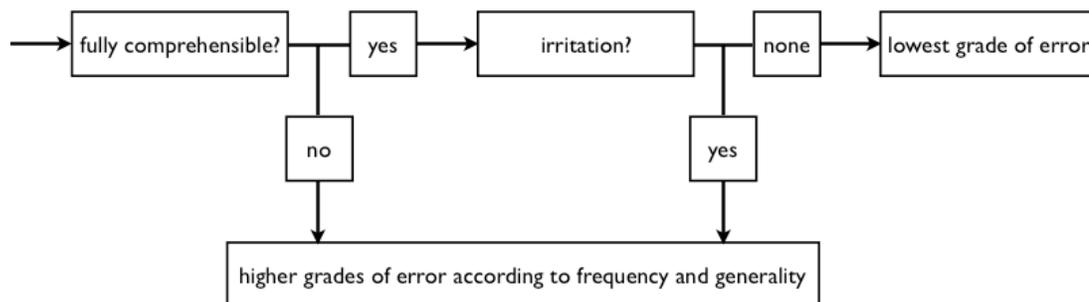


Figure 5. The evaluation of errors (Johansson, 1978: 5)

Cette façon de hiérarchiser les erreurs classe celles qui entravent la compréhension du message parmi les plus graves. Il est d'ailleurs intéressant de souligner qu'il semble plutôt rare qu'une erreur conduise à une mauvaise interprétation du message visé (donc à un échec de la communication). Johansson parle par exemple d'une étude de Juhász (1970) dans laquelle seulement 1,84 % des 34 erreurs grammaticales et sémantiques d'apprenants de l'allemand comme L2 présentées à des locuteurs natifs ont été qualifiées par ces derniers d'incompréhensibles. La quasi-totalité des énoncés erronés ont en effet été interprétés correctement.

Le deuxième critère de gravité est lié à l'irritation provoquée par l'erreur. Deux critères depuis longtemps utilisés dans les travaux sur l'évaluation de l'erreur – soit la fréquence et la généralité – permettent à Johansson d'établir des degrés de compréhensibilité et d'irritation. Par exemple, si une erreur entravant la compréhension est fréquente (*high frequency*) et qu'elle correspond à une règle ayant un champ d'application élevé dans la langue (*high generality*), elle sera jugée plus grave qu'une autre erreur tout aussi fréquente compromettant aussi la compréhension, mais relative à un aspect de la langue plus idiosyncrasique (par exemple la conjugaison d'un verbe irrégulier).

Les travaux de Johansson l'ont tout de même conduit à émettre certains doutes quant au poids du critère de généralité dans un jugement de gravité. Selon ce principe, les erreurs grammaticales – qui découlent généralement de la transgression de règles ayant un large champ d'application dans la langue – devraient être perçues comme plus graves que les erreurs lexicales, ce qui va à l'encontre des résultats des recherches de Johansson⁶⁰.

The fact that grammatical errors cause fewer problems of interpretation than do lexical ones is not surprising. By definition, grammar deals with the rule-governed, predictable aspects of language, the lexicon with the idiosyncratic, unpredictable aspects. A grammatical choice is part of a closed system, a lexical choice involves a selection from among members of an open class. In interpreting sentences with lexical errors the receiver is therefore faced with a more difficult task, since the possibilities of correction are much more varied and his knowledge of the language offers less help in predicting the correct form. (Johansson, 1978: 71)

Johansson souligne tout de même que le caractère plus grave des problèmes de nature lexicale n'est qu'une tendance forte, mais qu'il est possible de rencontrer certains problèmes grammaticaux perçus comme très graves. Nous ajoutons que les problèmes lexicaux inclus dans les travaux de Johansson sont essentiellement de nature sémantique. Bien sûr, en analysant des jugements de gravité portant sur un éventail de problèmes lexicaux aussi large que celui que couvre notre conception de l'erreur lexicale, nous observerions sans doute beaucoup plus de variation en fonction des types de problèmes⁶¹.

3.2.2 LES CRITÈRES DE GRAVITÉ SELON JAMES (1998)

Les premiers critères que mentionne James sont liés au fonctionnement du système linguistique. Est-ce que la transgression à l'origine de l'erreur concerne une « règle » générale de la langue ou est plutôt liée à des propriétés idiosyncrasiques, le régime verbal par exemple? D'un point de vue linguistique, les erreurs grammaticales seraient plus graves que les erreurs lexicales, puisque les règles grammaticales sont plus générales et leur application est beaucoup

⁶⁰ Johansson rapporte des conclusions similaires chez Engh (1971) et Olsson (1972, 1973).

⁶¹ Par exemple, dans les expériences de Johansson, les problèmes orthographiques apparaissent parmi les problèmes les moins graves.

plus fréquente; nous avons vu plus haut que l'importance de ce critère de gravité n'est pas démontrée dans les travaux de Johansson (1978). L'importance du domaine de l'erreur, c'est-à-dire la grandeur de la structure impliquée (voir plus bas, 3.2.3), peut aussi influencer le jugement de sa gravité, tout comme le fait que le phénomène linguistique en jeu soit récurrent dans la langue ou n'apparaisse au contraire que dans un nombre limité de contextes. Finalement, le potentiel de concomitance de l'erreur, c'est-à-dire sa capacité à entraîner d'autres erreurs (par un effet domino), est aussi un facteur linguistique pouvant influencer notre jugement de gravité.

La fréquence des erreurs joue aussi un rôle important dans notre jugement sur leur gravité. Il peut bien sûr s'agir du nombre d'occurrences élevé d'une même erreur, mais aussi de la densité des erreurs présentes dans un texte. La cohérence dans les erreurs commises constitue aussi un facteur dans l'appréciation des problèmes rencontrés. Un même item linguistique apparaît-il sous des formes différentes? Le même phénomène linguistique donne-t-il toujours lieu à des erreurs? Bien entendu, ces questionnements doivent tenir compte de la fréquence d'utilisation des différentes formes et constructions linguistiques; le fait que certaines structures soient sous-représentées pourrait révéler des stratégies d'évitement.

La deuxième famille de critères soulevée par James concerne le caractère compréhensible de l'erreur, qui est d'abord lié à l'impact de l'erreur sur l'intelligibilité du message, évidemment indissociable de la dimension lexicale. Par la « *comprehensibility* », l'auteur entend aussi la capacité du locuteur à transmettre son message de la bonne façon d'un point de vue pragmatique et connotationnel, sans laquelle l'énoncé produit sera perçu comme problématique.

Le degré auquel un problème linguistique est susceptible d'être repéré – ce que James appelle la « *noticability* »⁶² – peut être influencé par différents facteurs, par exemple la fréquence de l'erreur, mais aussi par l'attitude de l'interlocuteur. En effet, selon que la personne qui reçoit le message s'intéresse plus au sens ou à

⁶² Nous pourrions traduire ce terme par *saillance de l'erreur*.

la forme, différents types d'erreurs ressortiront de façon plus évidente; l'attitude de cette personne influencera donc ses jugements de gravité sur les erreurs relevées.⁶³

L'irritation provoquée par une erreur est certes difficile à mesurer, mais constitue aussi un facteur qui peut influencer notre jugement de gravité. Selon James, trois variables peuvent influencer l'irritation provoquée par une erreur. La première est le caractère prédictible de l'erreur, qui peut soit augmenter l'agacement face l'erreur commise ou au contraire susciter davantage de compréhension. La relation qu'entretiennent le locuteur et l'interlocuteur peut aussi jouer sur l'irritation; le fait que l'on soit moins bien disposé à l'égard de quelqu'un influencera défavorablement notre appréciation de ses productions linguistiques. Finalement, une erreur violant non seulement la norme linguistique, mais aussi la norme sociale, sera plus irritante; en effet, une erreur ayant pour origine un décalage culturel peut souvent être perçue comme un manque de politesse, ce qui peut augmenter son caractère irritant et nous la faire percevoir comme une erreur grave (Marquilló Larruy, 2003). Outre ces trois variables, James souligne aussi que la surutilisation de stratégies communicatives (par exemple l'emploi de périphrases pour « masquer » un manque d'accès au vocabulaire) peut aussi devenir irritante.

3.2.3 SURVOL DES TRAVAUX SUR L'ÉVALUATION DE LA GRAVITÉ DES ERREURS

James (1998) propose un survol des recherches s'étant intéressées à la question de la gravité des erreurs. Beaucoup des recherches mentionnées portent sur les différences de perception de gravité entre des locuteurs – principalement des enseignants – natifs et non natifs. Johansson (1978) consacre aussi le premier chapitre de son livre à un tour d'horizon des recherches sur la gravité des erreurs, qu'il regroupe par thématiques : les études portant sur la perception des accents étrangers, celles s'intéressant à l'intelligibilité entre dialectes, celles décrivant les réactions de locuteurs à des énoncés produits par des gens ayant

⁶³ C'est d'ailleurs une des raisons pour lesquelles les travaux sur la gravité font généralement appel à plusieurs locuteurs pour poser un jugement de gravité sur les erreurs plutôt qu'à un seul.

des troubles de la parole et les tests d'acceptabilité. Ellis (1994) fait quant à lui une recension des recherches qui se sont intéressées à l'effet que les erreurs produisent sur le correcteur (niveau de compréhension, perception de la gravité, « irritation » provoquée). Bien que la plupart des travaux présentés par ces auteurs ne s'avèrent pas particulièrement pertinents pour nous dans le cadre de cette thèse – d'autant plus que les recherches mentionnées portent toutes sur des erreurs commises par des apprenants d'une L2 –, quelques observations intéressantes en ressortent.

Tout d'abord, différentes études ont démontré une hiérarchisation similaire des types d'erreurs en ce qui a trait à leur gravité, indépendamment des groupes à l'origine des jugements (enseignants ou non, adolescents ou adultes, locuteurs natifs ou non)⁶⁴; McCretton et Rider (1993) émettent l'hypothèse d'une « hiérarchie universelle des erreurs ». Sans aller aussi loin, nous pouvons tout de même remarquer que les erreurs lexicales semblent occuper le rang des erreurs les plus graves dans une forte majorité de recherches, même chez les non-enseignants (Poltzer, 1978). Il ressort aussi que les erreurs dites « globales » – celles qui affectent la structure complète d'une phrase – semblent influencer davantage la compréhension que les erreurs « locales » – qui n'ont de répercussion qu'à l'intérieur d'un constituant – et auraient donc un plus grand impact sur le jugement des correcteurs. Une piste de recherche très intéressante est fournie par Lennon (1991a), qui propose de définir plus clairement l'opposition erreurs globales ~ locales en l'opérationnalisant en deux critères beaucoup plus précis. Le « domaine de l'erreur » (*error domain*), correspond à l'unité linguistique minimale que l'on doit considérer comme contexte pour que l'erreur devienne apparente (le mot seul dans le cas d'une erreur d'orthographe lexicale, au moins le verbe et son sujet dans le cas d'un mauvais accord sujet-verbe). L'« étendue de l'erreur » (*error extent*) correspond quant à elle à l'unité linguistique minimale que l'on doit enlever, remplacer, réordonner ou enrichir

⁶⁴ Il semble par contre que le poids relatif des facteurs influençant la perception de la gravité varie d'un groupe de correcteurs à un autre; les locuteurs natifs accordent généralement davantage d'importance à la compréhension du message, alors que le jugement des locuteurs non natifs est surtout influencé par une adéquation formelle à la langue (Ellis, 1994).

pour rendre correct l'énoncé (Lennon, 1991a: 191). Le domaine de l'erreur est toujours supérieur ou égal à son étendue. Ces critères, beaucoup moins subjectifs que les étiquettes « globale » et « locale », pourraient même constituer un autre axe de description des erreurs; Lennon démontre d'ailleurs leur utilité pour distinguer différents types d'erreurs lexicales et pour analyser des énoncés dans lesquels plusieurs erreurs se superposent.

Malgré les observations intéressantes qu'elles ont permises, les méthodologies associées aux recherches portant sur la gravité des erreurs en L2 ont fait l'objet de certaines critiques. Rifkin et Roberts (1995) leur reprochent le manque d'authenticité des échantillons de textes sur lesquels sont posés les jugements de gravité, ces échantillons étant souvent des condensés d'erreurs recueillis dans des productions ou alors des phrases retravaillées pour ne présenter qu'une seule erreur. Dans le même ordre d'idées, Johansson (1978) pointe une faille méthodologique dans les travaux qui présentent les erreurs hors contexte, ce qui évidemment fausse les jugements de gravité posés par les informateurs, puisqu'ils ne peuvent alors s'appuyer sur les éléments contextuels pour rétablir le message visé par la personne à l'origine de l'erreur. Les assises théoriques de plusieurs travaux sont aussi remises en cause par Rifkin et Roberts (1995), notamment la définition même du concept d'erreur. Les auteurs pointent également la difficulté de ces travaux à mettre en lumière le caractère multidimensionnel des jugements de gravité, qui – comme nous l'avons vu – reposent tant sur des critères linguistiques que sociaux et contextuels.

Par contre, cette dernière limite des travaux sur l'évaluation des erreurs n'a peut-être pas une incidence si grande leur fiabilité que l'on pourrait le croire à priori. En effet, comme l'indique Johansson (1978), il existe deux types de mesures de la gravité : les mesures subjectives et les mesures objectives. Les premières consistent à demander à l'informateur de porter un jugement sur un item linguistique problématique. Les secondes impliquent que l'informateur réalise une tâche (par exemple, répéter l'énoncé problématique); la gravité des problèmes est alors révélée indirectement par les comportements de

l'informateur lors de la réalisation de sa tâche. Comme les informateurs sont conscients de poser un jugement dans le cas des mesures subjectives, on aurait tendance à penser que les résultats produits par les études employant ce type de méthodologie seraient moins fiables que celles privilégiant les mesures objectives. Or, les recherches effectuées par Johansson viennent étayer les résultats d'autres travaux ayant mené à la conclusion que les jugements de gravité posés de façon subjective correspondent très fortement à ceux issus de mesures objectives.

The agreement between subjective et objective techniques means that the problem of establishing error gravity is reduced to manageable proportions, as the study could be based mainly on judgement tests. Such tests are easy to construct, administer and score, whereas the objective experiments require elaborate préparation, are time-consuming and difficult to evaluate. (Johansson, 1978: 48)

Toujours d'un point de vue méthodologique, Johansson rapporte les résultats d'études menées par Spencer (1973) et Quirk et Svartvik (1966) selon lesquelles des informateurs n'ayant pas un haut niveau de sophistication linguistique (*non-linguists*) fournissent des jugements d'acceptabilité plus convergents que d'autres ayant une formation linguistique plus approfondie, chez qui les jugements varient davantage d'un informateur à l'autre, peut-être parce qu'ils davantage tiraillés entre différentes normes, dont un locuteur « naïf » est plus ou moins conscient.

Avant de conclure cette section sur la gravité des erreurs, nous croyons important de faire une courte remarque terminologique. En effet, l'emploi du terme *gravité*, négativement connoté, pourrait paraître choquant dans une recherche sur l'erreur qui tente, dans la mesure du possible, de garder une position neutre par rapport aux écarts observés. Nous aurions pu envisager l'utilisation d'un terme plus neutre, comme nous l'ont suggéré certains, comme *poids de l'erreur*. Si nous prenons le parti de conserver le terme *gravité*, c'est parce que nous nous faisons l'héritier de toute une tradition de recherche sur *l'error gravity* et que nous jugeons approprié d'aligner notre terminologie à celle des chercheurs qui ont pavé la voie à ce type de recherche, dont ceux cités plus

haut. Bien entendu, il faut comprendre que lorsque nous parlons de *gravité de l'erreur*, nous faisons en fait plutôt référence à la *perception de la gravité de l'erreur*; une erreur n'est donc pas grave en soi, mais peut être perçue comme plus ou moins grave selon la personne qui y est confrontée.

Nous retenons donc de cette discussion sur l'évaluation des erreurs qu'il est en effet possible de hiérarchiser les erreurs en fonction de différents critères; outre certains critères linguistiques comme la fréquence et le type d'erreurs, des critères liés à l'effet que produisent les erreurs sur le destinataire de l'énoncé erroné influencent les jugements de gravité. Les recherches ont par contre démontré que les méthodes subjectives d'évaluation des erreurs sont aussi efficaces pour établir gravité relative de celles-ci que les méthodes objectives, plus lourdes.

CONCLUSION

Bien que les travaux sur l'analyse d'erreurs ne se soient que peu intéressés à l'erreur lexicale en tant que telle, nous pouvons conclure qu'ils fournissent certaines pistes intéressantes, notamment en ce qui concerne la démarche d'analyse ainsi que la caractérisation de l'erreur lexicale, que nous aborderons au chapitre suivant. Le traitement didactique de l'erreur semble la dimension la moins explorée jusqu'à maintenant, ce qui explique peut-être que les enseignants « demeurent [...] démunis face aux erreurs » (Marquilló Larruy, 2003: 118). En ce qui concerne les travaux ayant une visée plutôt psycholinguistique, le fait qu'ils fassent à peine mention de l'erreur lexicale découle sans doute du caractère très ponctuel de ce type de problèmes, difficile à intégrer à un système qui révélerait un réel cheminement dans les processus d'apprentissage. Malgré tout, l'analyse des erreurs lexicales, bien qu'elle n'y suffise pas à elle seule, constitue un outil important dans la description et la compréhension de la compétence lexicale d'un apprenant (Granger et Monfort, 1994), puisqu'elle peut faire ressortir des zones problématiques dans l'emploi du lexique.

4. LES DIFFÉRENCES ENTRE LES ERREURS EN L1 ET EN L2

La très grande majorité des travaux que nous avons consultés, notamment ceux proposant des typologies d'erreurs lexicales (que nous aborderons dans le prochain chapitre), traitent des problèmes de vocabulaire chez des apprenants d'une langue seconde, la plupart du temps l'anglais. En quoi ces recherches peuvent-elles réellement nous être utiles si nous nous intéressons aux problèmes lexicaux d'élèves de langue française ayant toujours été scolarisés en français? Est-ce que les erreurs d'apprenants d'une L1 sont les mêmes que ceux qui apprennent une L2? Les prochains paragraphes nous permettront de mieux comprendre ce qui rapproche et ce qui différencie les erreurs commises dans ces deux contextes.

Selon Arabski (1979), les erreurs en L1 et en L2 se rapprochent sur différents points, dont certains sont directement liés à la définition même de l'erreur, qui demeure la même en L1 et en L2. Tout d'abord, toutes les erreurs sont relatives, et un emploi n'est considéré comme fautif qu'en relation avec une norme, qui peut varier selon les contextes; Arabski donne l'exemple de formulations jugées tout à fait correctes en anglais américain qui étaient condamnées dans les classes polonaises d'anglais dans les années 60, où l'anglais britannique faisait office de norme. Ensuite, les erreurs des natifs, comme celles d'apprenants d'une L2, présentent une certaine systématisme, systématisme sans laquelle des interventions didactiques pour les prévenir seraient impossibles. Finalement – et c'est là le point particulièrement intéressant soulevé par Arabski –, il se peut que des erreurs produites en langue maternelle par des locuteurs natifs aient en quelque sorte une source « interlinguistique », lorsqu'elles peuvent être expliquées non pas par le contact de deux langues, mais par celui de deux registres, par exemple le français oral familier et le français standard.

Selon Ringbom (1987: 73), toute comparaison entre des corpus de textes écrits par des apprenants de L1 et de L2 de niveaux scolaires comparables et de niveaux de compétence équivalents révélera un nombre d'erreurs en L2 beaucoup plus élevé et une variété d'erreurs plus grande que chez les natifs.

Pour tenter de cerner plus clairement les différences entre les erreurs en L1 et en L2, l'auteur a comparé une cinquantaine de textes écrits par des élèves anglophones de 15 ans du nord de l'Angleterre avec le même nombre de textes écrits par des apprenants finnois de l'anglais de 19 ans. Cette étude, bien que plutôt exploratoire considérant le nombre peu élevé de rédactions étudiées, a permis d'intéressantes constatations en ce qui concerne les problèmes de nature lexicale.

Premièrement, pour ce qui est du nombre d'erreurs d'orthographe, le chercheur n'a constaté aucune différence importante entre les deux groupes qu'il a étudiés. Par contre, les sources présumées des problèmes d'orthographe lexicale des sujets varient selon que ceux-ci s'expriment en L1 ou en L2. Chez les natifs, ces problèmes orthographiques découlent souvent d'une difficulté à transposer à l'écrit une forme qu'ils maîtrisent à l'oral, difficulté amplifiée par l'orthographe opaque de l'anglais. Chez les apprenants d'une L2, qui apprennent souvent les formes orale et écrite d'un mot de façon simultanée, les problèmes de transposition d'une forme à l'autre peuvent intervenir dans les deux sens : ou la forme orale conduit l'apprenant à commettre une erreur orthographique ou encore, c'est la forme graphique du mot qui influence sa prononciation à l'oral. Une erreur orthographique en L2 peut aussi être commise sous l'influence d'une forme similaire en L1 ou dans une autre L2.

Deuxièmement, en ce qui concerne les autres types d'erreurs lexicales, Ringbom constate très peu de différences entre ses deux groupes. Comme les apprenants d'une L2, beaucoup de locuteurs natifs possèdent un vocabulaire limité, ce qui les conduit à confondre des mots éloignés sémantiquement sur la base de similarités formelles, à faire une mauvaise utilisation des suffixes dérivationnels, voire à utiliser des mots inexistantes en anglais. L'auteur établit aussi un parallèle entre les scripteurs faibles de L1 avec les scripteurs moyens ou faibles de L2 en soulignant l'utilisation de termes vagues et imprécis et la surutilisation de certains verbes communs (Ringbom, 1987: 77). Dans un même ordre d'idées, Hyltenstam (1988) n'a pas observé de différences significatives entre des

apprenants avancés (*near-native learners*) du suédois et des locuteurs natifs en ce qui a trait à la taille, la variété et la sophistication du vocabulaire utilisé en production écrite.

Finalement, Ringbom vient étayer notre idée selon laquelle les erreurs en L1 et en L2 ne se distinguent pas tant par leur forme que par leur source en tirant la conclusion suivante :

The difference between L1-writers and L2-writers is reflected not only in the slightly differing types of errors that they make, but also in their arrival at the same linguistic product via different process routes. (Ringbom, 1987: 78)

Les conclusions auxquelles arrivent les quelques chercheurs mentionnés dans cette section pourraient donc être ainsi résumées : d'un point de vue purement descriptif, les erreurs en L1 et en L2 seront à peu près les mêmes, mais se distinguent par leur fréquence et leur distribution. Les grandes différences entre les deux types d'erreurs apparaissent lorsque l'on s'intéresse à leur explication et les hypothèses posées varient selon qu'une erreur est commise en contexte de langue première ou de langue seconde. Par exemple, la question du transfert à partir de la L1, centrale dans les travaux en L2 (et à la base de l'analyse contrastive), ne nous permettra d'expliquer qu'une très faible proportion des problèmes que nous rencontrerons dans notre corpus (chez les élèves parlant plus d'une langue), encore que l'hypothèse d'Arabski (1979) mentionnée plus haut nous amènera peut-être à parler de transfert à partir de la langue orale.

Évidemment, l'objectif de notre recherche n'est pas de comparer des erreurs lexicales commises en L1 et en L2. Par contre, nous avons démontré que le recours à des travaux sur l'analyse d'erreurs en L2, comme ceux auxquels nous faisons référence tout au long de notre cadre conceptuel, est tout à fait pertinent pour éclairer des problèmes de nature lexicale en langue maternelle. Toutes les typologies d'erreurs lexicales analysées dans le chapitre suivant nous ont permis d'alimenter notre réflexion sur la description des erreurs et, dans une moindre mesure, sur les sources possibles des problèmes. Bien entendu, les résultats

relatifs à la fréquence ou à la distribution des erreurs présentés dans ces travaux nous sont quant à eux assez peu utiles.

Nous croyons qu'en distinguant clairement les niveaux de description linguistique et d'explication psycholinguistique des erreurs, il sera beaucoup plus facile de mesurer l'apport relatif de chacun des travaux analysés à notre propre recherche sur les problèmes lexicaux en langue maternelle. Selon nous, une grille d'analyse des erreurs lexicales purement descriptive devrait être universelle et s'appliquer autant à des problèmes relevés dans des corpus de L1 ou de L2. Évidemment, ce ne sera pas le cas pour le volet explicatif : les processus menant les locuteurs natifs à produire des erreurs, que nous désirons mettre en lumière dans cette thèse, diffèrent sans doute de ceux qui émergeraient si nous nous intéressions plutôt à ceux d'apprenants d'une L2.

CONCLUSION

Le présent chapitre nous a permis de mieux cerner la notion d'erreur linguistique, centrale dans notre travail sur l'erreur lexicale. Le survol des différents courants de recherche sur l'analyse d'erreurs a quant à lui fait ressortir les éléments intéressants que nous lèguent ces travaux pour l'étude d'un type d'erreur linguistique spécifique : l'erreur lexicale. C'est à ce concept que sera consacré le chapitre suivant.

V. L'ERREUR LEXICALE

À la lumière des informations sur le lexique présentées en introduction de ce cadre conceptuel et de notre définition de l'erreur linguistique, nous voilà prêt à aborder le cœur de notre thèse : l'erreur lexicale. Nous présenterons d'abord notre définition de l'erreur lexicale avant de faire un survol des principaux travaux s'étant intéressés à ce type d'erreurs de façon à faire ressortir les éléments à retenir pour constituer les outils descriptif et explicatif que nous utiliserons dans notre propre analyse.

1. DÉFINITION DE L'ERREUR LEXICALE

Comme le souligne Agustín Llach (2005a: 12), l'erreur lexicale est rarement définie dans les travaux s'y intéressant et les quelques définitions qui en sont proposées reposent sur des critères différents. Cette section vise à critiquer quelques définitions apparaissant dans la littérature afin de formuler celle que nous retiendrons dans le cadre de cette thèse.

1.1 L'ERREUR LEXICALE SELON AGUSTÍN LLACH (2005A)

Dans sa revue critique de la terminologie utilisée dans les travaux sur l'erreur lexicale, Agustín Llach (2005a) propose sa propre définition :

... the incorrect use of a lexical item (content word or idiom) in a specific context as a result of confusion between two words, owing to formal or semantic similarity and induced by mother tongue (L1) or target language (L2) influence. (Agustín Llach, 2005a: 16)

L'auteure ajoute que « the concept of 'incorrect' is a convention in language teaching that refers to those utterances that somehow differ from those of a native speaker of the L2 (cf. Corder, 1973). » (p.16). En ce sens, cette partie de sa définition se rapproche de la conception de l'erreur que nous avons décrite à la section III-1.4, d'autant plus que l'auteure y mentionne aussi l'importance du contexte.

Par contre, la caractérisation du concept d'erreur lexicale fournie par Agustín Llach soulève deux problèmes importants. Premièrement, elle est clairement orientée vers un contexte d'apprentissage d'une L2, même s'il y est fait mention d'erreurs dont la source serait intralinguale (« *induced by monther tongue [...] influence* »). Or, nous voudrions que la définition d'erreur lexicale retenue pour notre recherche soit universelle, c'est-à-dire qu'elle puisse désigner tous les problèmes de nature lexicale, peu importe que ceux-ci interviennent dans le discours de locuteurs natifs ou d'apprenants d'une L2.

Deuxièmement – et c'est là la principale limite de la définition d'Agustín Llach –, elle définit l'erreur par rapport au processus y ayant mené, processus qui serait le même pour toutes les erreurs lexicales. Cette caractérisation de l'erreur lexicale repose sur deux aprioris :

1. Toutes les erreurs lexicales découlent d'une confusion entre deux UL.
2. Toute confusion entre des UL est attribuable à des similarités formelles ou sémantiques entre ces UL.

Si la seconde affirmation nous semble acceptable, il en va tout autrement de la première. En effet, il est tout à fait possible qu'une erreur lexicale ne soit pas la conséquence d'une confusion entre deux UL; nous n'avons qu'à penser à l'emploi d'UL appartenant à un registre de langue inapproprié, à l'utilisation d'une locution déformée ou au pléonasma.

En bref, la définition de l'erreur lexicale proposée par Agustín Llach (2005a) n'est pas satisfaisante pour les raisons mentionnées plus haut. Nous trouvons d'ailleurs un peu surprenant que l'auteure propose une telle caractérisation de ce type de problèmes alors qu'elle critique les typologies d'erreurs lexicales dans lesquelles se chevauchent les niveaux descriptif et explicatif. Nous sommes pourtant d'accord avec elle lorsqu'elle affirme :

A clear definition of lexical error is needed, a definition that comprises all aspects pointed out by different authors and that serves all perspectives and analyses of errors in vocabulary. (Agustín Llach, 2005a: 16)

Voyons maintenant si les définitions proposées par d'autres auteurs permettront de répondre à ces critères.

1.2 DES DÉFINITIONS PLUS FORMELLES

Nous avons choisi de regrouper ici trois définitions de l'erreur lexicale qui se ressemblent en ce qu'elles caractérisent l'erreur en fonction des dimensions de l'UL qu'elle peut mettre en jeu. Chacune des définitions proposées présente certains problèmes, que nous soulèverons plus bas.

Résultat de la transgression de la norme concernant une propriété d'une unité lexicale (sens, forme, propriété de combinatoire) (Anctil, 2008: 351)

Tout emploi inapproprié d'une lexie L ayant comme cause des connaissances insuffisantes de ses propriétés sémantiques, formelles et/ou de cooccurrence. Un emploi est jugé inapproprié s'il mène à l'agrammaticalité, mais aussi s'il résulte en une maladresse. (Hamel et Milićević, 2007: 9)

[...] erreurs qui sont imputables à une mauvaise connaissance des propriétés inhérentes au mot, que celles-ci soient d'ordre logico-sémantique (le sens dénotatif), stylistique (le registre), collocationnel (la combinatoire lexicale) ou lexico-grammatical (les propriétés morpho-syntaxiques du mot). (Granger et Monfort, 1994: 58-59)

Le principal défaut de la caractérisation que nous avons nous-même proposée (2008) en lien avec notre recherche de maîtrise (Anctil, 2005) est d'être construite autour de la notion de norme qui, nous l'avons vu, n'est pas toujours très claire en ce qui regarde le lexique. La définition d'*erreur* que nous avons privilégiée nous permet de ne pas faire de référence directe à la norme et constitue sans doute un meilleur genre prochain pour une définition de l'erreur lexicale; c'est d'ailleurs la voie qu'ont choisie Granger et Monfort (1994). L'utilisation d'*erreur* comme noyau de notre définition permettra aussi d'éviter la question de la gravité des problèmes lexicaux dont Hamel et Milićević (2007) font mention dans la deuxième partie de leur définition.

La dimension explicative que nous avons reprochée à Agustín Llach (2005a) est aussi présente dans les définitions de Hamel et Milićević (2007) et de Granger et Monfort (1994). Ces auteurs attribuent les erreurs lexicales à des connaissances insuffisantes ou erronées sur des UL, ce qui peut en effet être à la source

d'erreurs, comme le mentionne James (1998)⁶⁵. Par contre, il existe sans doute d'autres raisons pouvant expliquer des erreurs de nature lexicale, notamment les confusions dont parle Agustín Llach. Une définition de l'erreur lexicale ne devrait pas, selon nous, poser d'hypothèses sur les sources d'erreurs ou, du moins, ne devrait pas présenter ces hypothèses comme les seules envisageables.

La définition de Granger et Monfort (1994) est à notre avis la plus intéressante parce que le terme *erreur* en constitue le noyau et parce qu'elle est la plus explicite sur les différentes facettes de l'UL sur lesquelles peut porter l'erreur. Outre la dimension explicative que nous venons d'évoquer, deux détails méritent d'être soulevés. Premièrement, le choix du terme *mot* ne nous semble pas le plus heureux pour les raisons présentées dans notre section sur l'unité lexicale (voir II-3). Deuxièmement, Granger et Monfort ont choisi de ne pas considérer les erreurs de forme parmi les problèmes lexicaux, ce qui était tout à fait justifié dans le cadre de leurs travaux, qui s'intéressaient spécifiquement à ce que les auteurs appellent « erreurs lexicales *stricto sensu* » (voir 2.3). Par contre, si nous voulons proposer une définition « universelle » de l'erreur lexicale, il sera important d'y inclure les problèmes formels.

1.3 NOTRE DÉFINITION DE L'ERREUR LEXICALE

Maintenant que nous avons fait ressortir les points problématiques des définitions retenues, voici le temps de proposer notre propre définition de l'erreur lexicale.

Erreur lexicale : *erreur linguistique qui concerne une propriété inhérente à l'unité lexicale, que celle-ci soit d'ordre formel, sémantique, stylistique, collocationnel ou lexico-grammatical.*

Une telle définition, lorsque comprise en relation avec la conception de l'erreur linguistique présentée à la section IV-1.4, permet une caractérisation exhaustive de l'erreur lexicale, qui peut s'appliquer autant dans un contexte de L1 que de

⁶⁵ Cet auteur établit d'ailleurs une distinction intéressante entre « *not knowing* » et « *knowing wrong* ». Selon lui, ce n'est qu'au-delà de ces deux niveaux que l'apprenant pourra produire des « *mistakes* » (« *getting it wrong but knowing how to put it right* »).

L2. Elle présente aussi l'avantage de témoigner d'une conception large du lexique (cf. Richards, 1976 ; Nation, 2001), qui permet de séparer les problèmes purement grammaticaux de ceux ayant pour origine la méconnaissance de propriétés lexicales en mettant en évidence les liens unissant lexique et grammaire (Granger et Monfort, 1994).

2. SURVOL DE QUELQUES TRAVAUX SUR L'ERREUR LEXICALE

La présente section propose un survol des principaux travaux s'étant intéressés au classement des erreurs lexicales afin de les contextualiser. Ce tour d'horizon nous conduira à détailler au point 3 les différents éléments à prendre en compte dans la construction d'une typologie descriptive des erreurs lexicales, éléments qui serviront d'assises à l'outil descriptif que nous proposerons dans le cadre de cette thèse (voir VII-3.1).

Dans notre revue critique des principaux travaux sur l'erreur lexicale, nous nous attarderons d'abord sur les typologies des travaux portant sur la L1, avant de critiquer certains classements issus de recherches en L2, dont certaines visaient le calcul de la richesse lexicale de textes d'apprenants. Comme quelques études consultées ne fournissaient qu'une ou deux pistes pertinentes pour notre thèse, nous avons choisi de ne pas toutes les présenter en détail, mais nous en mentionnerons les points intéressants à la section 3.

2.1 PATRIS ET VANSNICK (1992)⁶⁶

Parmi les typologies consultées, elle est une des seules à avoir été développée dans le cadre d'analyse d'erreurs lexicales commises par des élèves de L1. Assez complète quant aux types de problèmes lexicaux présentés, cette typologie offre notamment des pistes intéressantes en ce qui concerne la description et la compréhension de problèmes relatifs aux expressions idiomatiques (addition/suppression d'un élément, substitution d'un terme, passivisation, etc.). L'opposition *erreur* ~ *maladresse* proposée par les auteurs est aussi intéres-

⁶⁶ Voir Tableau 25 de l'annexe 1.

sante, mais nous ne croyons pas que les types de problèmes lexicaux puissent être associés de façon systématique avec l'un de ces jugements de gravité, ce que suggèrent les grilles présentées (l'une regroupant les fautes, l'autre les maladresses); il s'agit selon nous de deux axes d'analyse distincts. La plus grande réserve que nous avons à propos de cette typologie est le fait qu'elle soit « construite autour d'une même idée-maîtresse : l'analogie » (Patris et Vansnick, 1992: 50). En effet, si l'analogie est sûrement une source importante d'erreurs lexicales, nous sommes convaincu qu'elle n'est pas la seule (voir point 4); la typologie proposée élimine donc d'emblée tous les autres scénarios d'explication des erreurs analysées. À noter aussi que cette typologie n'est pas descriptive, mais explicative, puisqu'elle prend comme point de départ la source des erreurs. Les problèmes de nature orthographique n'y sont pas mentionnés.

2.2 ANCTIL (2005)

Développée durant notre maîtrise dans le but de décrire les problèmes lexicaux d'étudiants universitaires, cette typologie⁶⁷ a pour fondements linguistiques les concepts utilisés par la *Lexicologie explicative et combinatoire* (voir II-1.1). La prise en compte de toutes les facettes de l'UL (sens, forme, combinatoire grammaticale et lexicale), telles qu'elles sont conceptualisées par la Théorie Sens-Texte, y est donc bien présente.

Trois grands regroupements structurent cette typologie: les problèmes découlant du non-respect d'une propriété inhérente à l'UL, ceux liés à la transgression de contraintes plutôt contextuelles ou stylistiques et les cas de création d'une lexie fictive, c'est-à-dire une mauvaise association signifié ~ signifiant. Nous remettons aujourd'hui ces regroupements en question, pour différentes raisons. Tout d'abord, le fait de regrouper les cas de création d'une lexie fictive (impropriété, mauvaise dérivation morphologique, création d'une nouvelle acception d'un vocable) nous conduit à répartir les problèmes formels et sémantiques dans différents endroits de la grille, ce qui en

⁶⁷ Voir Tableau 26 de l'annexe 1.

embrouille la structure. Par exemple, il serait à notre avis plus intuitif que l'impropriété figure avec les autres problèmes de sens. Par ailleurs, le regroupement des « usages douteux en fonction du contexte » présentait à la fois des problèmes lexicopragmatiques (ex. registre de langue inapproprié) et des problèmes liés davantage à des aspects stylistiques, comme la répétition ou l'imprécision. Si nous sommes convaincu que les problèmes lexicopragmatiques doivent faire partie d'une étude sur les difficultés lexicales des élèves, nous ne sommes pas certain que ce soit le cas pour les problèmes stylistiques, qui peuvent difficilement être quantifiés et sont soumis à une très grande subjectivité (voir point 3.6).

Malgré le caractère opérationnel de notre grille, révélé par sa mise à l'essai sur un corpus de près de 400 problèmes lexicaux – dont plus de 99% ont pu être codés –, certaines observations nous avaient poussé à proposer des modifications à y apporter. L'un des principaux problèmes constatés était le fait que certaines classes d'erreurs (par exemple « impropriété » ou « mauvaise collocation ») regroupaient des problèmes de nature somme toute assez différente et auraient gagné à être subdivisées de façon à fournir une description plus fine. Par contre, parmi les subdivisions proposées, certaines relevaient plutôt de l'explication des erreurs que de la description pure, et leur intégration telles quelles à la grille aurait conduit à un chevauchement des niveaux d'analyse.

L'utilisation de notre typologie pour le codage des erreurs d'étudiants a aussi révélé certains problèmes d'ordre méthodologique, le plus frappant étant sans doute la difficulté à classer les problèmes de sens à l'aide de nos classes d'erreurs sémantiques. Ces classes, pourtant solidement justifiables d'un point de vue lexicologique – et à première vue intéressantes dans une perspective d'intervention didactique – se sont avérées peu opérationnelles pour départager les problèmes rencontrés dans notre corpus. Nous reviendrons sur ce point à la section 3.2.

Par ailleurs, le fait que la grille proposée dans notre mémoire de maîtrise contienne une classe d'erreurs propre aux locutions (« non-respect du caractère figé d'une locution ») nous semble aujourd'hui problématique; ce traitement particulier n'est pas justifié dans la mesure où les locutions sont aussi des UL qui possèdent des caractéristiques similaires à celles des lexèmes. Cette classe d'erreurs regroupait notamment les problèmes relatifs à la déformation de la locution (par exemple par le remplacement d'un terme de l'expression par un synonyme), que nous classerions maintenant parmi les problèmes formels avec les barbarismes. Les autres problèmes étiquetés « non-respect du caractère figé d'une locution » étaient liés à la transgression de propriétés de combinatoire qui peuvent aussi s'appliquer aux lexèmes : invariabilité, impossibilité de la passivation, non-respect du genre nominal, etc. Par conséquent, nous croyons qu'il n'est pas pertinent de traiter à part les problèmes de locutions, même s'il sera bien sûr intéressant de vérifier si certaines transgressions sont plus fréquentes avec ce type d'UL.

Finalement, la question du double codage, c'est-à-dire le fait pour une erreur de pouvoir appartenir à plus d'une des classes de la typologie, avait aussi attiré notre attention et sera traitée dans notre méthodologie (section VII-3.1.2).

2.3 GRANGER ET MONFORT (1994)

Granger et Monfort (1994) veulent démontrer que l'analyse d'erreurs, mise de côté par les chercheurs au profit d'une orientation psycholinguistique, « reste un outil très utile pour décrire la compétence lexicale de l'apprenant [lorsqu']elle repose sur une base empirique suffisante et des catégories bien définies. » (p.55-56) Développée pour analyser des productions d'apprenants d'une L2, la typologie proposée dans cet article contient un nombre très restreint (5) de classes d'erreurs assez larges, ceci dans le but de réduire le risque de classements arbitraires. Le classement repose principalement sur la distinction entre les erreurs lexicogrammaticales – celles découlant de « violations des propriétés morpho-syntaxiques inhérentes au lexème employé » (p.63) – et les erreurs lexicales *stricto sensu* (sémantiques, stylistiques et syntagmatiques).

Tableau 2. Catégorisation des erreurs lexicales selon Granger et Monfort (1994)

TYPE D'ERREUR		EXEMPLES	
Erreurs lexico-grammaticales		<i>le livre sterling [la] ... dont les priorités sont différentes que celles de la France [de] their parents, girlfriends and wives [wives]</i>	
Erreurs lexicales <i>stricto sensu</i>	Lexème « inexistant » en L2	<i>Les conservatives britanniques [conservateurs] an unnegligible opportunity [significant]</i>	
	Lexème existant en L2	Erreurs logico-sémantiques ⁶⁸	<i>elle est positive dans son engagement [ferme, résolue] the late dispute between France and Belgium</i>
		Erreurs stylistiques	<i>Il y a ceux qui ont des sous [de l'argent] Europeans (sic) politicians will have a bone to pick with each other on various issues [have many issues to resolve]</i>
		Erreurs syntagmatiques	<i>... une collaboration proche avec les autres pays [étroite] Efforts are done in this direction [made]</i>

Les auteurs mentionnent que les catégories proposées sont susceptibles d'être subdivisées pour raffiner la description, mais sans en présenter les sous-catégories. Selon nous, une telle subdivision est pourtant essentielle pour que le classement puisse mener à des observations intéressantes; la seule lecture des exemples fournis suggère de toute façon des sous-classes possibles. Par ailleurs, on remarque les chercheurs ne considèrent pas les erreurs d'orthographe dans leur typologie et mentionnent qu'elles appartiennent à une classe à part.

Un point intéressant de la démarche de Granger et Monfort est de distinguer explicitement la description des erreurs de leur explication en utilisant deux grilles différentes. Par contre, tout aussi schématique que la grille descriptive, la grille explicative ne pousse pas très loin les hypothèses sur les sources d'erreurs : dans une perspective de L2, on se contente d'établir la distinction entre les erreurs interlinguales et intralinguales.

En bref, les grands principes qui sous-tendent les typologies (descriptive et explicative) de Granger et Monfort – soit la distinction nécessaire des niveaux d'analyse et la conception large de l'erreur lexicale – sont tout à fait pertinents pour assurer une analyse uniforme des erreurs et découlent d'une

⁶⁸ Cette classe inclut aussi les erreurs touchant à la connotation du mot.

argumentation solide. Par contre, les catégories constitutives de chacune des typologies nous semblent trop larges pour permettre des observations ayant une réelle valeur didactique. Ce classement ne constitue selon nous que le squelette de ce que doit être un outil d'analyse des erreurs lexicales. Le raffinement de cette structure à l'aide des classes proposées par les autres auteurs devrait conduire à une analyse plus susceptible de produire des résultats révélateurs des problèmes lexicaux en production.

2.4 MILIĆEVIĆ ET HAMEL (2007)

Proposée dans le cadre d'un projet de recherche visant le développement d'un dictionnaire de reformulation pour les apprenants du français, la typologie de Milićević et Hamel⁶⁹ se rapproche de celle d'Anctil (2005) sur deux points majeurs : 1) elle repose sur le même cadre théorique⁷⁰ : la Théorie Sens-Texte (Mel'čuk, 1997) et 2) elle adopte une démarche résolument descriptive qui s'appuie sur une caractérisation des erreurs en fonction de la dimension de l'UL impliquée : le signifié, le signifiant et les propriétés de combinatoire.

L'intérêt principal de cette typologie est la rigoureuse démarche de description qui y est associée. L'étape de reconstruction mentionnée par Corder (1981) y est systématiquement présentée et la comparaison des deux énoncés (erronné et restitué) y est faite de façon explicite et même décrite, pour certains types d'erreurs, à l'aide de formalismes – notamment les *fonctions lexicales* (Mel'čuk et coll., 1995). Cette typologie s'avère particulièrement intéressante pour l'analyse des erreurs liées au signifié, qu'elle appréhende non pas en tentant d'isoler des composantes de sens particulières de l'UL problématique – approche qui s'est révélée peu opérationnelle chez Anctil (2005) –, mais en évaluant la relation sémantique entre l'énoncé analysé et sa reconstruction (quasi-synonyme, terme

⁶⁹ Voir Tableau 27 de l'annexe 1.

⁷⁰ C'est d'ailleurs la seule typologie, parmi celles que nous avons rencontrées, pour laquelle une théorie linguistique lui servant de fondement est explicitement mentionnée. Corder (1981: 36) souligne d'ailleurs l'importance d'un modèle théorique adéquat et préconise le choix d'une théorie orientée vers le sens.

générique, mot de sens proche, etc.). Cette approche davantage holistique va dans le sens des observations faites par Zimmermann (1986).

2.5 ZIMMERMANN (1986)

La typologie proposée par Zimmermann (1986)⁷¹, destinée à décrire les erreurs lexicales d'étudiants germanophones de l'anglais, vise une plus grande différenciation des problèmes lexicaux, généralement regroupés dans des classes très larges comme « problème de sélection lexicale », peu utiles pour la correction de textes et offrant peu de pistes d'intervention didactique. Zimmermann souligne aussi le caractère problématique de la mixité des typologies proposées jusqu'alors, dans lesquelles se chevauchent la description et l'explication des erreurs; il propose de s'en tenir en un premier temps à la description des problèmes rencontrés. Bien que l'auteur fasse mention de plusieurs types d'erreurs lexicales (problèmes de collocations, redondance, problèmes dans la formation de mots, etc.), le classement qu'il propose s'intéresse surtout aux problèmes d'ordre sémantique et la démarche mise de l'avant est claire :

I am interested in a description of the meaning difference between the interlanguage lexeme chosen and the target language lexeme intended, which should be as exact as possible. (p.32)

Nous reviendrons en détail sur les propositions de classement des problèmes de sens faites par Zimmermann à la section 3.2.1. Nous nous contenterons pour l'instant de dire que nous retenons principalement des travaux de Zimmermann la tendance naturelle à privilégier un traitement holistique dans la description des problèmes sémantiques plutôt qu'une analyse par composantes de sens.

2.6 AGUSTÍN LLACH (2004)

Consciente du besoin de développer une typologie des erreurs lexicales qui soit à la fois non ambiguë et généralisable, Agustín Llach (2004) propose une classification qui vise à rendre compte des erreurs selon trois critères :

⁷¹ Voir Tableau 28 de l'annexe 1.

linguistique, psycholinguistique et pédagogique. Sa description linguistique regroupe les erreurs en fonction de la dimension touchée, c'est-à-dire en quatre grandes classes (erreurs sémantiques, syntaxiques, phonologiques/orthographiques et pragmatiques). Le critère psycholinguistique s'intéresse quant à lui au processus ayant mené à l'erreur, dans une perspective d'interférence avec la L1 ou la L2. Partant de l'idée que certains mots sont plus difficiles à apprendre que d'autres (Laufer, 1990), Agustín Llach a jugé pertinent d'introduire un troisième critère de classification, qu'elle qualifie de « pédagogique », car il sert selon elle à orienter les pratiques. Ce critère consiste essentiellement à spécifier l'UL en jeu dans l'erreur, de façon à observer si certains mots, ou certaines classes de mots, sont plus susceptibles d'occasionner des erreurs.

L'application de cette classification aux erreurs tirées d'un corpus de 19 rédactions d'apprenants hispanophones de l'anglais lui a permis des observations intéressantes. Tout d'abord, les noms et les verbes sont les classes de mots les plus affectées par l'erreur lexicale. Ensuite, les erreurs de collocations (classées parmi les problèmes syntaxiques par l'auteure) découlent la plupart du temps d'interférences avec la L1. Finalement, les erreurs pragmatiques sont les moins nombreuses, mais cela est sans doute attribuable au type de textes constitutifs du corpus analysé (textes produits dans un cours d'anglais des affaires).

La classification d'Agustín Llach est intéressante parce qu'elle tient compte à la fois de la description des erreurs et de leur explication, tout en distinguant explicitement ces deux niveaux d'analyse. Elle permet aussi d'établir clairement la source intra ou interlinguale des problèmes analysés. Par contre, d'un point de vue linguistique, les classes proposées demeurent très larges (erreurs sémantiques, syntaxiques, phonologiques/orthographiques, pragmatiques) et ne permettent pas une grande différenciation des erreurs. De plus, le critère « pédagogique », bien que tout à fait pertinent quant aux observations qu'il permet sur les « zones à risque » de l'utilisation du vocabulaire, relève selon

nous de la description linguistique et ne devrait pas constituer un troisième niveau d'analyse. En effet, le fait de dire que l'erreur concerne telle UL, qui est un nom ou un verbe, fait selon nous partie de la description de l'erreur.

2.7 ENGBER (1995)

Engber (1995) vise à déterminer l'impact du vocabulaire présent dans un texte sur le score général qui sera attribué à ce texte par un correcteur. Pour ce faire, elle confronte les résultats obtenus avec quatre formules de calcul de la richesse lexicale aux scores effectivement attribués. Comme les formules proposées tiennent compte des erreurs lexicales présentes dans les textes, Engber a dû définir ce qu'elle entend par « erreurs lexicales en production de textes »; elle en propose une typologie, reproduite en fin de document (voir Tableau 29 de l'annexe 1).

Ce classement se fonde sur une division entre les problèmes de « choix lexical » et ceux de « forme lexicale ». Nous trouvons un peu étrange d'effectuer une telle bipartition, puisque « forme » et « choix » sont d'ordre différent : la forme est une des dimensions de l'UL, alors que le choix lexical est un mécanisme inhérent au processus de production. Par ailleurs, beaucoup des sous-catégories d'erreurs proposées par Engber nous semblent nébuleuses et les exemples fournis pour les illustrer ne nous permettent de nous en faire une idée plus claire. Finalement, l'auteure n'a pas comptabilisé les erreurs relevant de la combinatoire syntaxique parmi les erreurs lexicales et elle ne fait aucune mention des autres types de problèmes de combinatoire grammaticale.

Malgré tout, cette typologie offre des idées intéressantes relativement à certains types d'erreurs, notamment sur la nature formelle de la confusion entre paronymes (souvent considérée comme un problème sémantique). Elle présente aussi certaines catégories d'erreurs absentes de la plupart des typologies, par exemple la forme des verbes irréguliers.

2.8 ARNAUD (1984 ; 1992)

Dans ces deux travaux, Arnaud aborde aussi la question de l'erreur lexicale, mais cette fois en tentant de déterminer la validité de tests de vocabulaire par la mise en relation de résultats obtenus à ces tests par des apprenants de L2 et d'indices de richesse lexicale pour des textes produits par ces mêmes apprenants.

Dans sa recherche de 1984, Arnaud distingue six types d'erreurs lexicales (voir tableau ci-dessous), mais les catégories proposées relèvent à la fois de la description (ex. *minor spelling mistake*) et de l'explication (ex. *confusion between two lexemes*) et ne concernent que la forme et le sens : les erreurs de combinatoire n'en font pas partie.

Tableau 3. Types d'erreurs lexicales selon Arnaud (1984)

Minor spelling mistakes	<i>*personnal</i>
Major spelling mistakes	<i>*scholl</i>
Derivation mistakes	<i>*comparate, *sucedded</i>
Faux-amis	<i>They should be *prevented that it is difficult.</i>
Interference from another language	<i>*to spare money</i>
Confusion between two lexemes	<i>The teacher *learn them maths.</i>

Son article de 1992 témoigne d'une conception de l'erreur lexicale beaucoup plus large – et donc plus proche de la nôtre – tenant compte de la dimension collocationnelle et des prépositions régies, mais ne présente pas explicitement de typologie des erreurs lexicales.

Les travaux d'Arnaud ne sont donc pas très aidants pour le classement d'erreurs en soi, mais contiennent des pistes de réflexion à explorer. Pour le départage des erreurs lexicales et grammaticales, notamment, l'auteur suggère de ne pas considérer comme lexicales les erreurs qui concernent des UL appartenant à des classes fermées (pronoms, déterminants, conjonctions, prépositions). Cette piste nous a semblé intéressante à première vue, mais nous semble maintenant trop contraignante; à l'instar de Granger et Monfort (1994: 58), nous croyons que l'opposition ensemble ouvert/ensemble fermé n'est pas satisfaisante pour départager le lexical du grammatical, puisque beaucoup d'UL appartenant aux

classes fermées « ont un comportement éminemment idiosyncrasique ». En effet, il est tout à fait possible d'imaginer des erreurs concernant des lexies des classes fermées que nous voudrions voir parmi les erreurs lexicales, par exemple, des locutions prépositives ou conjonctives déformées (**au dépend de*) ou des erreurs portant sur le sens d'une préposition non régie (*#à l'instar de* au sens de 'contrairement à').

Une deuxième idée intéressante soulevée par Arnaud (1992) est de considérer les problèmes liés à la forme des verbes irréguliers parmi les erreurs lexicales. Nous croyons par contre que cette position devrait s'appliquer à toutes les UL présentant des flexions particulières, qui peuvent aussi conduire à des erreurs formelles (ex. *three little *mouses, She has long *hairs.*).

3. ÉLÉMENTS POUR UNE TYPOLOGIE DESCRIPTIVE DES ERREURS LEXICALES

Malgré les pistes intéressantes proposées par les travaux présentés, le classement d'erreurs lexicales à l'aide de typologies demeure difficile (Patris et Vansnick, 1992 ; Anctil, 2005). Maintenant que nous avons fait ressortir les forces et les limites des principales typologies d'erreurs lexicales présentes dans la littérature, nous ferons une synthèse des points à considérer dans la construction de la typologie la plus opérationnelle possible. Nous traiterons d'abord de différents aspects liés à la structure de la typologie, avant de nous attarder à la question du classement des problèmes de nature sémantique, puis finalement au classement des autres types de problèmes lexicaux.

3.1 LA STRUCTURE DE LA TYPOLOGIE

Les sections qui suivent s'intéresseront à différents aspects relatifs à la construction d'une typologie d'erreurs lexicales et devraient nous permettre de faire certains choix concernant la structure de la grille que nous proposerons dans notre méthodologie (voir VII-3.1).

3.1.1 DISTINCTION DES NIVEAUX D'ANALYSE

Un des reproches principaux adressés aux différents classements proposés dans le cadre des travaux sur l'analyse d'erreurs est, comme nous l'avons vu, le chevauchement des étapes de description et d'explication. L'utilisation d'une grille d'analyse d'erreurs multidimensionnelle, ou d'une série de grilles, est préconisée par de nombreux auteurs (James, 1998 ; Granger, 2003) dans le but d'éviter les problèmes de chevauchement. Il sera donc crucial que notre outil d'analyse distingue explicitement, à la façon de celui de Granger et Monfort (1994), ces différents niveaux d'analyse.

Nous devons aussi nous méfier d'un second type de chevauchement, présent dans le classement de Patris et Vansnick (1992) : la division des types de problèmes lexicaux en deux grandes classes, celle des *erreurs* et celle des *maladresses*. S'il existe sans doute certaines corrélations entre des types d'erreurs et le fait qu'on les perçoive comme plus ou moins graves – par exemple, l'utilisation d'une UL de registre familier à l'écrit apparaît souvent maladroite alors qu'une erreur d'orthographe paraît plus grave –, notre section sur l'évaluation de la gravité des erreurs (voir IV-3.2) a bien établi que les variables en jeu dans un jugement de gravité sont nombreuses. Par conséquent, il serait dangereux d'établir à priori des rapprochements entre des types d'erreurs et des jugements de gravité. Nous écartons donc une correspondance du type de celle qui structure le classement de Patris et Vansnick (1992). La prise en compte nécessaire du niveau de gravité des problèmes observés constitue selon nous une autre facette de la description des erreurs⁷² et devra être faite indépendamment de la caractérisation structurale des problèmes, donc à l'aide d'une grille indépendante. Évidemment, une correspondance entre les types d'erreurs et les jugements que l'on pose sur ces erreurs pourra être faite à posteriori, si l'analyse d'un nombre important de problèmes lexicaux fait ressortir des tendances fortes; par exemple, dans notre corpus de maîtrise

⁷² Cette distinction des niveaux d'analyse est d'ailleurs illustrée par la Figure 3 présentée à la section IV-2.2.

(2005), la totalité des 79 problèmes relatifs à l'orthographe ont reçu le statut d'« erreur ».

3.1.2 LES GRANDES DIVISIONS D'UNE TYPOLOGIE DES ERREURS LEXICALES

L'observation de différentes typologies d'erreurs lexicales a fait ressortir le désir des chercheurs d'effectuer des regroupements parmi les types d'erreurs recensés. En effet, rares sont les typologies présentées sous forme de listes, comme celle de Zughoul (1991) ou d'Arnaud (1984). Dans la plupart des typologies, de grandes divisions permettent de réunir des problèmes partageant des caractéristiques communes. Quelles sont les divisions les plus pertinentes dans le cas d'erreurs lexicales?

Un type de classement, notamment utilisé par Dulay, Burt et Krashen (1982), caractérise les erreurs en fonction des altérations faites à l'énoncé cible pour le rendre erroné. Les catégories sont généralement les suivantes : omission, addition, mauvaise sélection et mauvais positionnement syntaxique⁷³. Il est bien sûr évident que l'apprenant qui fait une erreur n'altère pas un énoncé correct pour produire une phrase incorrecte, mais ces quatre catégories d'erreurs permettent tout de même de caractériser l'écart entre l'énoncé produit et l'énoncé cible. Par ailleurs, un classement de ce type serait peu intéressant dans une perspective d'analyse d'erreurs lexicales, puisque la grande majorité des erreurs devraient se situer par les problèmes de sélection.

Lennon (1991b) propose de classer les erreurs lexicales selon les classes de mots auxquelles appartiennent les UL utilisées de façon problématique. Comme le soulignent Granger et Monfort (1994), un tel classement peut mener à des observations intéressantes (par exemple en faisant ressortir le nombre important d'erreurs touchant les prépositions), mais « présente le net désavantage d'être très éloigné des pratiques didactiques qui [...] se basent sur les différents types de "connaissances lexicales" » (p.60). Par ailleurs, cette

⁷³ Nous constatons que ces quatre types de modifications à l'énoncé cible correspondent aux quatre manipulations syntaxiques utilisées par la grammaire moderne (effacement, ajout, remplacement, déplacement).

approche pourrait aussi compliquer le classement des erreurs dans le cas de problèmes qui découlent de la cooccurrence de deux UL (cf. erreurs de collocation, d'incompatibilité sémantique, pléonasme), puisqu'il faudrait déterminer quelle classe de mots, parmi celles des deux UL en jeu, devrait être retenue. Si nous croyons que l'information à propos de la classe des mots des items impliqués dans les erreurs doit être considérée dans une analyse des erreurs lexicales, nous ne jugeons pas qu'elle doive être centrale au point de structurer le classement des problèmes observés.

James (1998: 144) croit que la façon la plus pratique de classer les erreurs lexicales est justement de le faire en fonction des types de connaissances que les gens ont sur les mots et suggère sept catégories correspondant aux dimensions de la connaissance d'un mot selon Richards (1976).

- i. sa morphologie
- ii. son comportement syntaxique
- iii. les contraintes fonctionnelles ou contextuelles auxquelles il est soumis
- iv. ses différentes acceptions (son potentiel dénotatif)
- v. sa charge connotationnelle
- vi. les mots auxquels il peut s'associer
- vii. sa fréquence d'emploi⁷⁴

Toujours selon James (1998: 145), la dichotomie *erreurs formelles* ~ *erreurs sémantiques* doit être centrale dans la classification.

Il semble que cette idée de classer les erreurs en fonction des dimensions de l'UL auxquelles elles sont liées soit partagée par plusieurs chercheurs. Par exemple, Granger et Monfort (1994) proposent une typologie opposant les erreurs lexicogrammaticales et les erreurs lexicales *stricto sensu*, c'est-à-dire une classe d'erreurs large qui regroupe les problèmes lexicaux sémantiques, stylistiques et syntagmatiques ; par contre, les erreurs orthographiques sont absentes de leur classement. Les typologies qui reflètent le plus explicitement les principes énoncés par James sont celles dont les grandes divisions figurent dans le tableau suivant.

⁷⁴ Notre traduction.

Tableau 4. Grandes divisions des principales typologies

Agustín Llach (2005c)	Milićević et Hamel (2007)	Anctil (2005)
semantic	ayant trait au signifié	altération du sens
syntactic	ayant trait au syntactique	altération d'une propriété de combinatoire grammaticale
orthographic	ayant trait au signifiant	altération de la forme
pragmatic	ayant trait au syntactique (régime)	usage douteux en fonction du contexte

Même si leur typologie se veut davantage explicative que descriptive, Patris et Vansnick (1992) effectuent aussi des regroupements du même type, en séparant les erreurs concernant la combinatoire lexicale de celles relative à la rection (cf. régime) et en distinguant les problèmes ayant pour source une confusion formelle, sémantique ou « de construction ».

Évidemment, il serait aussi possible d'effectuer de grandes divisions entre les erreurs selon les mécanismes y ayant conduit (par exemple l'analogie), mais une telle position se révélerait peu satisfaisante dans une typologie à visée descriptive. Nous reviendrons sur la pertinence de ce type de regroupement lorsque nous nous intéresserons à l'étape d'explication des erreurs (voir 4).

Nous retenons donc de cette discussion qu'il serait pertinent de structurer, comme nous l'avons fait en 2005, notre typologie des erreurs lexicales autour des différentes dimensions de l'UL. Pour cette raison, et pour être en accord avec notre conception de l'unité lexicale et de l'erreur lexicale, nous croyons que les problèmes d'orthographe lexicale doivent aussi apparaître dans la typologie, ce qui nous empêche de partager la position de Granger et Monfort (1994) d'en exclure les problèmes formels.

3.1.3 LE DEGRÉ DE RAFFINEMENT DES CLASSES D'ERREURS PROPOSÉES

La question du degré de raffinement des classes d'erreurs utilisées pour notre analyse est cruciale. Évidemment, plus la description visée est fine, plus le nombre de classes d'erreurs est élevé, ce qui a comme conséquence de rendre plus lourde l'utilisation de la typologie proposée et d'augmenter « le risque de classements arbitraires » (Granger et Monfort, 1994: 61). Par contre, une

caractérisation plus précise des problèmes analysés permet des observations plus riches ayant davantage de potentiel didactique. Il s'agit donc de trouver un équilibre entre le raffinement de la description et le caractère opérationnel de la démarche d'analyse. En effet, si une description fine des problèmes rencontrés est souhaitable, il ne faut pas oublier que c'est la quantité d'erreurs observées qui donne de la validité aux observations faites. Par conséquent, il faut que le travail d'étiquetage des problèmes lexicaux se fasse assez aisément et rapidement pour permettre le classement d'un nombre de problèmes assez important pour justifier un traitement quantitatif subséquent, susceptible de faire émerger des observations intéressantes.

Il ne faut pas oublier non plus que les groupes d'erreurs constitués à l'aide de la typologie ne sont que l'aboutissement d'une première étape de l'analyse; le classement proposé doit servir au traitement initial des données, qui seront ensuite analysées de façon plus qualitative et, possiblement, redivisées en sous-classes. Les classes d'erreurs retenues dans la typologie doivent donc être assez précises pour permettre un départage intéressant dès le départ sans trop alourdir le processus d'étiquetage des erreurs.

À notre avis, des classes d'erreurs aussi larges que celles proposées par Granger et Monfort (1994) – erreurs logico-sémantiques, stylistiques, syntagmatiques, etc. – ne permettent pas un traitement quantitatif des erreurs ayant un très fort potentiel didactique, puisque chaque classe regroupe un éventail de problèmes ayant sans doute des caractéristiques très différentes; un reclassement des problèmes de chacune des classes sera donc nécessaire pour pouvoir tirer des observations intéressantes. Zimmermann (1986) souligne d'ailleurs la nécessité d'une plus grande différenciation des erreurs lexicales. Nous croyons que la position adoptée par Debyser *et coll.* (1967), soit de fournir une description des erreurs en deux étapes – d'abord à l'aide de classes larges et ensuite par un raffinement à l'aide de sous-classes – est plus susceptible de produire des données rentables du point de vue de l'analyse. Ainsi, le premier niveau de classement pourrait correspondre aux grandes divisions de la typologie retenue

et le second, présenter les classes d'erreurs plus précises. Par ailleurs, nous ne croyons pas que ces deux étapes doivent être nécessairement séparées dans le temps; seulement, l'analyste devrait pouvoir se réserver le droit de classer des erreurs au premier niveau seulement s'il juge qu'aucune des sous-classes proposées ne correspond au problème à étiqueter. Ceci éviterait les « classements arbitraires » contre lesquels Granger et Monfort (1994) font une mise en garde. Les problèmes n'ayant pu faire l'objet d'un classement de deuxième niveau pourraient ensuite être analysés plus à fond et, possiblement, conduire à l'ajout de nouvelles sous-classes d'erreurs à la typologie.

Parmi les typologies présentées plus haut, le nombre de classes d'erreurs varie de 5 (chez Granger et Monfort, 1994) à 19 (chez Anctil, 2005). Ayant expérimenté l'étiquetage d'erreurs avec une typologie contenant près de 20 classes de départ, nous pouvons affirmer que ce nombre en apparence élevé est tout à fait « gérable » et qu'il n'alourdit pas le classement des erreurs dans la mesure où les classes sont bien définies et maîtrisées par la personne qui effectue le travail d'analyse. Des données classées avec ce degré de raffinement permettent d'ailleurs des observations intéressantes, puisqu'elles donnent l'accès à des observations quantitatives plus précises et, de par le nombre plus limité d'éléments dans chacune des classes, se prêtent bien à une analyse plus qualitative.

Bien entendu, une grille contenant un nombre aussi élevé de classes d'erreurs s'avère pertinente dans le cadre d'une recherche comme la nôtre, mais pourrait ne pas être appropriée pour une utilisation telle quelle par un enseignant pour diagnostiquer les erreurs de ses élèves.

3.2 L'ANALYSE DES PROBLÈMES DE SENS

Les problèmes sémantiques sont sans aucun doute les plus difficiles à diagnostiquer (Martin, 1984 ; Anctil, 2005). Parmi les typologies présentées plus haut, trois proposent un classement plus détaillé de ce type de problèmes : celles de Zimmermann (1986), d'Anctil (2005) et de Milićević et Hamel (2007). Si nous

nous limitons précisément à ces travaux, c'est qu'ils demeurent les plus descriptifs dans leur façon de traiter les problèmes de sens; nous avons donc écarté, pour ce volet de notre discussion, les classements qui impliquent une hypothèse sur la source des erreurs sémantiques⁷⁵. Nous verrons dans les paragraphes qui suivent de quelle façon sont décrites les erreurs sémantiques dans ces travaux et quels sont les problèmes soulevés par les classements proposés.

3.2.1 LES PROBLÈMES DE SENS CHEZ ZIMMERMANN (1986)

Zimmermann divise les erreurs lexicales sémantiques en deux grandes classes : celles dont le classement repose sur une conception holistique du sens (*holistic conception of meaning*) et celles qui sont décrites en faisant référence à des composantes de sens particulières (*feature errors*). La première classe contient à son tour deux grandes familles d'erreurs. Les erreurs appartenant à la première famille découlent de la confusion entre deux UL liées sémantiquement (*sense relation errors*), par exemple par une relation d'hyponymie/hyperonymie, d'hétéronymie, de synecdoque (partie ~ tout) ou de métonymie (effet ~ cause)⁷⁶. La deuxième famille englobe ce que Zimmermann appelle les « *field errors* », c'est-à-dire les cas pour lesquels c'est le champ sémantique auquel appartient l'UL erronée qui a été mal identifié; l'exemple fourni par l'auteur est l'utilisation de **memorial*⁷⁷ (= 'monument'), qui appartient au champ sémantique des « immeubles publics », à la place de *memorandum* (= 'mémoire'), qui appartient au champ sémantique de la « communication écrite publique ». Le tableau qui suit résume le classement des erreurs sémantiques proposé par Zimmermann.

⁷⁵ Par exemple *analogie sémantique* (Patris et Vansnick, 1992) ou *faux amis* (Arnaud, 1984).

⁷⁶ Voir Tableau 28 de l'annexe 1 pour des exemples de chaque type.

⁷⁷ Les textes étudiés par Zimmermann ont été écrits en anglais par des étudiants allemands.

Tableau 5. Les différents types d'erreurs sémantiques selon Zimmermann (1986)

<i>Holistic conception of meaning</i>	Sense relation errors : use of a wrong supernym (underspecification) use of a wrong hyponym according to the context confusion between co-hyponyms, part/whole, result/process, etc.
	Field errors
<i>Small differences in meaning</i> <i>(componential approach of meaning)</i>	Feature errors

Cette répartition soulève quelques problèmes.

Tout d'abord, la bipartition entre les erreurs dont le classement repose sur une conception holistique du sens et celles pour lesquelles des composantes de sens spécifiques sont identifiées ne témoigne pas selon nous de caractéristiques inhérentes aux problèmes sémantiques analysés, mais rend plutôt compte de la capacité de l'analyste à les expliquer plus ou moins précisément. Zimmermann l'admet lui-même : « [...] the classification as feature or field errors depends largely on whether a semantic feature or meaning component lends itself to description. » (p.36) Cette opposition entre les deux types d'erreurs n'est donc pas justifiée du point de vue de la description linguistique et est plutôt soumise à des considérations d'ordre méthodologique (capacité ou incapacité de rendre compte des problèmes dans une perspective d'analyse componentielle), elles-mêmes liées à la difficulté de proposer une description rigoureuse du sens.

La question du chevauchement des différentes classes d'erreurs est aussi soulevée par Zimmermann et met en lumière certaines limites du classement proposé. L'auteur s'interroge entre autres sur la pertinence de classer les problèmes de confusion entre des hétéronymes/cohyponymes⁷⁸ parmi les *sense*

⁷⁸ Zimmermann distingue d'ailleurs les hétéronymes (ex. WASH ~ SCRUB) des cohyponymes (ex. ERADICATE ~ EXTERMINATE), les premiers ayant pour caractéristique d'avoir des sens mutuellement exclusifs. Cette distinction, illustrée de façon peu convaincante par l'auteur, nous semble plutôt difficile à saisir et trop pointue pour être opérationnelle dans la description des problèmes de choix lexical.

relation errors, alors que ces derniers peuvent souvent être considérés comme des quasi-synonymes qui se distinguent par un nombre très limité de composantes de sens, ce qui les rapproche des *feature errors*. Encore une fois, le classement effectué dépendra de la capacité à isoler les composantes de sens distinctives des UL en jeu. Départager les *field errors* des *feature errors* s'avère aussi difficile, et ce, malgré les principes de classements proposés par l'auteur (dans le cas des *field errors*, possibilité d'associer l'UL erronée et l'UL cible à des sous-champs sémantiques distincts, composantes sémantiques distinctives hiérarchiquement plus élevées).

Le classement des *field errors* nous paraît aussi problématique dans la mesure où certains des champs sémantiques évoqués par l'auteur dans ses exemples semblent établis de façon ad hoc (ex. DETACHED appartient au sous-champ « *lack of dedication* », alors que *DESINTERESTED fait plutôt partie du sous-champ « *objective attitude* »). Pour qu'un tel classement soit réellement opérationnel, il faudrait spécifier un répertoire des champs sémantiques de référence; à notre connaissance, un tel répertoire n'existe pas.

La mise à l'essai que Zimmermann a faite de sa typologie sur un corpus d'environ 400 problèmes lexicaux tirés de textes écrits par des apprenants germanophones de l'anglais permet certaines constatations intéressantes relativement aux erreurs lexicales sémantiques, qui constituent plus de 60 % des problèmes analysés. Tout d'abord, les *sense relation errors* représentent 20 % des erreurs, mais selon l'auteur, une bonne partie de ces problèmes auraient aussi pu être classés parmi les *feature errors*. Les *field errors* constituent le tiers de tous les problèmes analysés, alors que les *feature errors* ne s'élèvent qu'à 10 %. Zimmermann conclut de cette répartition une tendance à préférer une description holistique des problèmes de sens. Il souligne aussi que plus de 40 % des erreurs appartiennent aux deux catégories les plus difficiles à décrire : les *field errors* et les *feature errors*.

Dans un article écrit l'année suivante, Zimmermann (1987) affirme que, pour être fécondes dans une perspective d'analyse du langage, les notions d'« analyse

par composantes sémantiques » (*semantic feature analysis*) et de « champ sémantique » (*semantic field*) doivent être prises dans leur sens large. Selon lui, l'analyse componentielle en jeu dans la description de problèmes lexicaux doit se limiter aux composantes « nécessaires et suffisantes » à la comparaison de l'UL erronée avec l'UL visée, et non pas prétendre à l'exhaustivité recherchée par les lexicographes visant une décomposition sémantique complète. Quant à la notion de « champ sémantique », elle doit être élargie de façon à rendre compte d'associations entre UL qui transcendent les relations sémantiques étroites, par exemple sur la base de liens conceptuels beaucoup plus éloignés (ex. JUSTICE et BALANCE) ou sur la base de liens métaphoriques. Zimmermann parle de cadres sémantiques (*semantic frames*) ou de schèmes cognitifs. L'auteur démontre que la prise en compte de ces schèmes cognitifs permet d'expliquer certains problèmes dans les choix lexicaux.

Les travaux de Zimmermann sont donc très intéressants en ce qu'ils montrent la forte tendance des analystes d'erreurs à privilégier un traitement holistique dans la description des problèmes sémantiques. L'enrichissement d'une description holistique des écarts sémantiques par une approche componentielle se limitant aux composantes de sens pertinentes sans aspirer à une décomposition sémantique exhaustive nous semble aussi pertinent dans une perspective d'intervention didactique. Par contre, nous croyons qu'une typologie présentant un classement holistique des problèmes de sens risque d'être plus opérationnelle qu'une autre fondée sur une analyse componentielle du sens des UL en jeu. À cet effet, voyons les problèmes rencontrés par Anctil (2005), qui a adopté une telle approche.

3.2.2 LES PROBLÈMES DE SENS CHEZ ANCTIL (2005)

La typologie que nous avons proposée dans le cadre de notre maîtrise (2005) contient quatre types de problèmes sémantiques⁷⁹, répartis en deux endroits

⁷⁹ Nous ne tenons pas compte ici des problèmes de mélange de codes, notamment l'emploi d'anglicismes, qui peuvent parfois être considérés comme des problèmes sémantiques (ex. *Le film #prend place dans la localité de Beaufort High.*), mais pas toujours (ex. *Il est satisfait avec son emploi.*).

dans la grille. Le tableau ci-dessous présente une réduction de cette typologie avec seulement les problèmes de sens.

Nous constatons que deux de ces classes d'erreurs – celles relevant du non-respect d'une composante de sens – reposent sur une analyse componentielle du sens des UL en jeu. En effet, pour associer une erreur à l'une ou l'autre de ces classes, il est nécessaire de cibler la composante de sens de l'UL erronée à la source du problème; par exemple, la composante 'dormir à **demi**' de SOMNOLER, qui entre en conflit avec le modificateur *profondément* présent dans le deuxième exemple présenté dans le Tableau 5. Les cas de création d'une nouvelle acception d'un vocable (B.1) demandent aussi une analyse des composantes de sens pour pouvoir identifier le patron de polysémie à la source de l'erreur; si un étudiant utilise BALLON-PANIER pour désigner un panier de basketball, c'est sur la base d'un lien de métonymie (Anctil, 2005: 126). Dans le cas de l'impropriété, par contre, le jugement est posé sur le problème de façon holistique: on considère que l'on est en présence d'une impropriété lorsque le sens du terme erroné et celui de l'UL visée nous semblent trop éloignés pour, justement, expliquer le problème par une analyse componentielle.

Tableau 6. Les différents types d'erreurs sémantiques selon Anctil (2005)

Cas			Exemples
A. Altération d'une dimension de l'unité lexicale	Sens	Non-respect d'une composante de sens	Non-respect du typage des actants. <i>Il ne faut pas déplacer un blessé pour ne pas #l'endommager.</i>
		Non-respect d'une différence spécifique	<i>Elle n'a pas entendu la sonnette, car elle #somniait profondément.</i>
B. Utilisation d'une lexie fictive	1. Par la création d'une nouvelle acception d'un vocable selon un patron de polysémie productif.		<i>Il l'attendait sur la #rondelle de la station Berri-UQAM. (= 'banc de forme arrondie')</i>
	2. Impropriété : association d'un sens avec le signifiant d'une lexie dont le signifié n'entretient aucun lien sémantique avec la nouvelle lexie créée.		<i>Ils ont pris un repas #frugal. = 'abondant, copieux'</i>

Nous avons conclu, après une mise à l'épreuve de notre typologie sur un corpus de textes d'étudiants universitaires, que si les classes d'erreurs sémantiques

proposées sont valides théoriquement, elles sont peu opérationnelles pour le classement des problèmes réels. En effet, des 93 problèmes sémantiques relevés dans le corpus et répartis dans les quatre classes présentes dans le Tableau 6, 58,2 % ont été considérés comme des impropriétés alors que, de notre propre aveu, certains problèmes auraient dû appartenir aux autres classes.

Il se peut donc que plusieurs problèmes aient été codés « impropriété » faute de pouvoir déterminer avec précision la dimension du sens transgressée. La classe « impropriété » a donc en quelque sorte servi de classe fourretout aux problèmes de sens qui nous laissaient perplexes. (Anctil, 2005: 122)

Il est à noter que la classe d'erreurs la plus problématique est celle relative au non-respect d'une différence spécifique, qui correspond d'ailleurs aux *feature errors* de Zimmermann (1986). Selon nous, la difficulté à cerner les composantes de sens problématiques, qui rend extrêmement ardu le classement des problèmes sémantiques, peut être expliquée par deux facteurs principaux. Tout d'abord, l'étape de restitution du sens visé par la personne ayant commis l'erreur n'est faite qu'à partir d'un nombre très limité d'occurrences (souvent une seule) de l'UL problématique, ce qui fait qu'il n'est pas toujours facile de mettre en lumière le sens de l'UL erronée tel que le conçoit la personne qui est à l'origine de l'erreur. Par ailleurs, les dictionnaires qui servent de norme de référence dans le diagnostic des problèmes lexicaux ne fournissent pas de façon systématique les décompositions sémantiques exhaustives qui permettraient de cibler aisément toutes les composantes du sens des UL.

Les difficultés rencontrées dans le classement de problèmes sémantiques sur la base des composantes de sens transgressées constituent donc un argument de plus en faveur d'un classement fondé sur une appréciation plus holistique des problèmes de sens. Même si nous jugeons que le fait de pouvoir pointer la composante de sens problématique soit pertinent d'un point de vue d'intervention didactique, nous pensons aujourd'hui qu'une typologie des problèmes sémantiques serait plus opérationnelle en appréhendant ces derniers dans une perspective holistique.

À la lumière des problèmes constatés dans notre mémoire de maîtrise, certains aménagements avaient été suggérés à la typologie, dont certains affectaient les problèmes sémantiques. La suggestion la plus importante en cette matière était l'intégration à la typologie des sous-classes suivantes parmi les impropriétés.

Tableau 7. Sous-classes des impropriétés proposées par Anctil (2005)

IMPROPRIÉTÉS	Accidentelle : pour lesquelles la confusion nous paraît difficilement explicable (ex. EXPRESSIF pour signifier 'turbulent')
	Anglicisme : l'unité lexicale erronée a été utilisée sous l'influence de l'anglais
	Paronymes : l'erreur découle d'une confusion entre deux unités lexicales présentant des similarités formelles

Certaines de ces sous-classes nous semblent aujourd'hui problématiques dans le cadre d'une typologie descriptive des erreurs, puisqu'elles nous paraissent relever de l'explication. C'est notamment le cas de la classe « paronyme », qui pose clairement une hypothèse sur la source de l'erreur, alors que rien ne prouve que la personne ayant commis cette dernière a effectivement une connaissance des deux paronymes possiblement en jeu. La classe « accidentelle », en fondant clairement le classement sur la capacité de l'analyste à formuler une hypothèse sur la source du problème, se situe résolument du côté de l'explication et n'apporte rien à la description structurale de l'erreur. Ces deux sous-classes ne devraient donc pas selon nous faire partie d'une typologie descriptive des erreurs lexicales.

La sous-classe « anglicisme », par contre, nous paraît toujours pertinente d'un point de vue descriptif. En effet, les erreurs appartenant à cette classe ne sont pas nécessairement produites sous l'influence de l'anglais – ce qui alors constituerait une explication de leur source –, mais sont considérées comme problématiques en français parce qu'elles ont fait leur apparition dans cette langue sous l'influence de la langue anglaise alors que le français disposait déjà d'un terme équivalent (ou, dans le cas des anglicismes syntaxiques, d'une construction équivalente). Ainsi, il se pourrait tout à fait qu'un francophone

monolingue utilise l'anglicisme DÉFINITIVEMENT (au sens de 'certainement') simplement parce que ce mot est fréquemment utilisé en ce sens par un nombre important de locuteurs francophones au Québec; le fait de qualifier ce terme d'anglicisme permet donc simplement de préciser le type d'impropriété en jeu dans cette erreur.

Nous retenons donc de notre mémoire qu'une analyse componentielle des problèmes sémantiques n'est pas toujours très efficace et ne permet pas une description homogène des problèmes sémantiques, puisque les ouvrages de référence ne décrivent pas tous les sens avec la même finesse et que la décomposition sémantique nécessaire à une analyse componentielle ne peut être faite de façon également intuitive pour toutes les UL.

3.2.3 LES PROBLÈMES DE SENS CHEZ MILIĆEVIĆ ET HAMEL (2007)

Milićević et Hamel accordent sans aucun doute la place la plus importante à une approche holistique dans le traitement des problèmes sémantiques. En effet, leur classement de ce qu'elles appellent les erreurs sémantiques affectant la nomination (voir Tableau 8) repose précisément sur une appréciation globale de la « distance sémantique » entre le terme erroné et l'UL visée.

Ce traitement des problèmes de nomination permet donc de placer les erreurs observées dans différentes classes sans avoir à cibler des composantes de sens particulières. Les auteures proposent une distinction entre les mauvais quasi-synonymes et les lexies de sens proches : les quasi-synonymes doivent être substituables dans au moins quelques contextes, alors que « deux lexies de sens proche ne sont jamais substituables en contexte, même si elles peuvent appartenir à un même champ sémantique. » (p.154) La classe « sens fictif » correspond aux cas de figure où aucun lien sémantique évident ne lie le terme erroné et l'UL visée. Cette gradation à trois niveaux nous paraît tout à fait pertinente, puisqu'elle permet de rendre compte de relations sémantiques plus ou moins fortes tout en restant assez intuitive. Par ailleurs, notre participation au travail de Milićević et Hamel à l'automne 2006 nous a permis de constater

que les classes d'erreurs sémantiques proposées permettent un classement des problèmes sémantiques plus aisé que, par exemple, la typologie d'Antil (2005).

Tableau 8. Les erreurs sémantiques chez Milićević et Hamel (2007)

	Affectant la nomination	Affectant la combinaison
Erreurs ayant trait au signifié	1. Mauvais quasi-synonyme ex. #C'est aussi une bonne occasion pour regarder (⇒ observer) la nature.	5. Incompatibilité sémantique des sens ex. #La ville n' encourage pas les bicyclettes (⇒ n' encourage pas l' usage des bicyclettes).
	2. Mauvais terme générique ex. #La bicyclette est un véhicule (⇒ moyen de transport).	
	3. Lexie de sens proche ex. #La bicyclette permet de se rendre à beaucoup d' espaces (⇒ endroits).	6. Incompatibilité pragmatique (du sens et de la situation) ex. #Vous pouvez chercher (⇒ trouver) la plupart des ingrédients au supermarché.
	4. Sens fictif ex. #... la pluie et les bicyclettes sont des sûretés (⇒ faits) de la vie belge.	

Nous avons par contre une certaine réserve relativement à la classe d'erreurs « mauvais terme générique »; l'exemple fourni ne nous convainc pas de la pertinence de cette dernière. En effet, cette classe nous informe de la relation d'hyponymie/hyperonymie qui existe entre le terme erroné et une autre UL présente dans la phrase (entre BICYCLETTE et MOYEN DE TRANSPORT dans l'exemple). Or, la caractérisation des problèmes sémantiques doit généralement reposer – et c'est le cas pour les autres classes proposées par les auteures – sur une évaluation de la proximité sémantique du terme erroné et de l'UL visée (entre VÉHICULE et MOYEN DE TRANSPORT dans l'exemple). Ainsi, pour l'exemple fourni, la classe « mauvais quasi-synonyme » aurait été appropriée, puisque VÉHICULE et MOYEN DE TRANSPORT sont effectivement des quasi-synonymes. Évidemment, il sera important de pointer la relation sémantique d'hyponymie/hyperonymie existant entre BICYCLETTE et MOYEN DE TRANSPORT pour expliquer l'erreur, mais il nous semble méthodologiquement problématique de le faire à l'étape de classement des problèmes si nous voulons que la démarche d'analyse des erreurs soit la même pour tous les types de problèmes.

La classe « incompatibilité pragmatique (du sens et de la situation) » repose aussi sur une conception holistique du sens et nous semble intéressante dans la

mesure où elle permet de rendre compte des cas où l'énoncé produit est tout à fait correct en apparence, mais qu'il ne correspond pas à l'intention de l'auteur. Par contre, la distinction entre les cas relevant de cette classe et les cas d'utilisation d'un mauvais quasi-synonyme n'est pas toujours évidente⁸⁰. Par conséquent, nous croyons qu'il est peut-être plus prudent de ne pas multiplier les classes possibles et de s'en tenir à la description holistique correspondant aux problèmes de nomination.

La classe « incompatibilité sémantique des sens » est la seule du classement de Milićević et Hamel (2007) qui implique une analyse plus componentielle du sens; elle correspond en quelque sorte à la classe « non-respect du typage sémantique des actants » présente dans notre typologie (Anctil, 2005). Pour pouvoir étiqueter un problème de ce type, il faut être en mesure d'identifier certaines contraintes sémantiques qui limitent les possibilités de combinaison de l'UL en jeu⁸¹. Malgré cette approche componentielle du sens, cette classe d'erreurs est opérationnelle, puisqu'une similaire s'est révélée productive chez Anctil (2005).

Le classement des problèmes de sens proposé par Milićević et Hamel s'avère donc le plus intéressant des trois observés dans cette section. Malgré les quelques réserves soulevées, nous considérons que l'approche holistique du sens qui sous-tend les classes d'erreurs proposées constitue un bon compromis entre la précision de la description des erreurs et l'opérationnalité de la typologie.

3.3 L'ANALYSE DES PROBLÈMES DE NATURE FORMELLE

Dans les typologies étudiées, les problèmes relatifs à la forme sont traités de façon beaucoup plus homogène que ce que nous avons constaté pour les

⁸⁰ Un des cas représentatifs de cette difficulté est la phrase suivante, produite par une étudiante anglophone du français : *Puis, moi-même #j'écoute un son extraordinaire; un crie (sic) d'un bébé, mon bébé [...]* Ici, le terme visé était ENTENDRE, un quasi-synonyme d'ÉCOUTER qui est incompatible avec la situation.

⁸¹ Les auteures mentionnent le non-respect du typage sémantique des actants comme un cas possible d'incompatibilité sémantique des sens. Par contre, les seuls exemples fournis illustrent ce cas précis et nous voyons mal quelles sont les autres possibilités.

problèmes sémantiques. Les classes d'erreurs formelles y sont généralement assez larges et peu nombreuses, probablement parce que ces erreurs ne sont pas celles qui intéressent le plus les chercheurs travaillant sur l'erreur lexicale.

Étrangement, les erreurs d'orthographe, qui constituent sans doute le cas le plus typique de problèmes formels, ne sont pas présentes dans tous les travaux considérés. Pour Granger et Monfort (1994: 65), par exemple, « elles constituent une catégorie à part, distincte des erreurs grammaticales et des erreurs lexicales ». Si cela nous paraît peu cohérent avec notre conception de l'erreur lexicale – la forme étant l'une des facettes de la connaissance d'une UL – et que nous adopterons une position différente, nous voyons certains arguments en faveur de l'exclusion des erreurs d'orthographe d'une typologie des erreurs lexicales.

Tout d'abord, il est vrai que dans la tradition pédagogique, les erreurs d'orthographe ont toujours été traitées comme un ensemble distinct des erreurs de vocabulaire; elles sont généralement quantifiées dans la même façon que les erreurs d'accords grammaticaux et font l'objet de plus d'attention didactique. Ceci s'explique sans doute par le fait que l'orthographe se plie davantage à un enseignement structuré et est soumis à une norme plus claire que les autres dimensions du vocabulaire. Ensuite, les erreurs d'orthographe sont si fréquentes et si nombreuses qu'elles nécessiteraient à elles seules une typologie très étoffée pour que l'on puisse rendre compte de la variété des distorsions possibles (choix de graphèmes erroné, inversion de phonèmes, épenthèse, mauvais emploi de majuscules, d'accents, oubli de consonnes doubles ou de lettres étymologiques, etc.); l'intégration d'un nombre très élevé de classes d'erreurs formelles à une typologie globale des erreurs lexicales compromettrait son caractère opérationnel en la rendant difficilement utilisable. Par ailleurs, l'intégration d'une seule catégorie regroupant tous les types d'erreurs d'orthographe (lexicale) mènera nécessairement à un ensemble d'erreurs hétéroclite peu intéressant d'un point de vue didactique. Pour ces raisons, nous comprenons que certains chercheurs aient pu choisir de ne pas inclure les erreurs d'orthographe

lexicale à leur typologie. Bien sûr, il existe des travaux portant spécifiquement sur les problèmes orthographiques, dont les plus importants sont sans contredit ceux de Nina Catach (voir entre autres Catach *et coll.*, 1986).

Malgré tout, la plupart des classements étudiés contiennent ce type de problèmes et établissent une distinction entre les « simples » erreurs d'orthographe et les erreurs affectant le signifiant de façon plus importante (Anctil, 2005), les secondes affectant généralement la prononciation de l'UL déformée (cf. les barbarismes, chez Anctil, 2005). Il est intéressant de souligner que certains auteurs incluent parmi ces problèmes formels les déformations affectant les locutions (Anctil, 2005).

La grande majorité des typologies développées dans un contexte de L2 mentionnent l'utilisation d'UL (lexèmes ou locutions) inexistantes dans la langue cible; certaines regroupent tous les cas dans une même classe (Granger et Monfort, 1994), alors que d'autres, comme Zimmermann (1986), distinguent différents cas de figure (erreurs dans la formation de mots, emprunts directs à la L1, formes ininterprétables). Du côté des typologies s'intéressant à la L1 (Patris et Vansnick, 1992 ; Anctil, 2005), nous retrouvons les problèmes relatifs à une mauvaise dérivation morphologique (**emplacer*, **obsessionner*), mais l'emploi d'une forme inexistante n'apparaît pas, sans doute parce qu'il est moins susceptible de survenir dans un contexte de L1, l'emprunt direct à une autre langue étant moins fréquent⁸².

Un autre type d'erreur formelle souvent mentionné par les auteurs est l'emploi d'une paraphrase « illicite » (Milićević et Hamel, 2007), c'est-à-dire une reformulation peu idiomatique et maladroite, au lieu d'une UL précise. Milićević et Hamel nomment ce type de problème « forme analytique », puisque la paraphrase produite correspond en quelque sorte à une définition analytique (cf. décomposition sémantique) plus ou moins complète et habile de l'UL visée. Un problème que nous entrevoyons avec cette classe d'erreurs est la distinction

⁸² La classe « mélange de codes » de notre typologie (2005) permet tout de même de rendre compte des emprunts directs à une autre langue, mais pas des cas où la forme utilisée est tout simplement inexistante.

entre les erreurs de « forme analytique » réelles et les problèmes découlant de l'emploi de périphrases. Les premières témoignent réellement d'un problème d'accès à une forme existante dans la langue et impliquent l'apparition d'une UL précise dans la correction; Milićević et Hamel donnent l'exemple *une bicyclette qui ne bouge pas* pour STATIONNAIRE. Les problèmes de périphrase présentent aussi une formulation maladroite, mais ils relèvent plutôt du style que de la méconnaissance d'une UL précise. Leur correction est plus difficile, car ils ne peuvent être aisément remplacés par un seul terme; par exemple, // *dit des tas d'idées et des raisonnements pour se défendre...* (Patris et Vansnick, 1992).

Les derniers types de problèmes formels que nous retrouvons dans les travaux consultés sont les problèmes relatifs à la flexion de verbes irréguliers, principalement mentionnés par les chercheurs s'intéressant au calcul de la richesse lexicale (Arnaud, 1992 ; Engber, 1995 ; Laufer et Paribakht, 1998). Nous jugeons cette classe d'erreurs très intéressante, mais comme nous l'avons mentionné plus haut, croyons qu'elle doit être élargie à toutes les UL présentant des flexions particulières.

3.4 L'ANALYSE DES PROBLÈMES LIÉS À LA COMBINATOIRE RESTREINTE

Les problèmes liés à la combinatoire restreinte des UL sont sans contredit les moins bien représentés dans les travaux consultés. Cela a sans doute à voir avec le fait que, même s'ils reconnaissent aux UL des propriétés variées autres que le sens et la forme, les auteurs délimitent rarement quelles sont ces propriétés; le clivage entre les travaux reposant sur un cadre théorique clair (Anctil, 2005) et les autres est marqué en ce qui concerne la façon de traiter les problèmes de combinatoire restreinte. En effet, les travaux de ces auteurs proposent un inventaire assez complet des propriétés des UL, inventaire qui se reflète à des degrés variables dans les classements proposés. Les exemples d'erreurs proposés par Granger et Monfort (1994) témoignent d'une conception des erreurs lexico-grammaticales très proche de la nôtre. Par contre, ces auteurs ne proposent qu'une classe d'erreurs large qui regroupe tous ces types de problèmes et ne permet par conséquent aucune différenciation, ce qui est tout à

fait compréhensible vu les objectifs de leurs travaux (description des erreurs lexicales *stricto sensu*, voir 2.3). Les typologies proposées par Anctil (2005) et Milićević et Hamel (2007) reposant sur le même cadre théorique, elles contiennent des classes similaires d'erreurs liées à la combinatoire restreinte et présentent un degré de raffinement comparable.

Les problèmes de collocation sont sans doute le type d'erreurs de combinatoire le plus récurrent dans les typologies consultées, même si une certaine confusion semble régner entre les problèmes de collocation, de combinatoire sémantique et de figement des locutions dans certains classements (notamment chez Engber, 1995). Beaucoup de classements incluent aussi des erreurs de prépositions régies⁸³ ou de régime verbal, qui constituent les seules propriétés lexicogrammaticales généralement considérées.

3.5 L'ANALYSE DES PROBLÈMES LEXICOPRAGMATIQUES

Les problèmes liés à l'emploi d'UL appartenant au registre de langue familier sont présents dans la plupart des typologies d'erreurs et portent souvent l'étiquette « erreur stylistique » (Zimmermann, 1986 ; Granger et Monfort, 1994 ; Milićević et Hamel, 2007). Milićević et Hamel incluent ces problèmes parmi les problèmes de combinatoire restreinte, ce que nous jugeons tout à fait pertinent en regard de notre conception de l'erreur lexicale.

L'utilisation de régionalismes dans des contextes où ils pourraient ne pas être compris est aussi liée à la dimension pragmatique de l'utilisation du vocabulaire et apparaît seulement chez Patris et Vansnick (1992) et Anctil (2005).

Un dernier type d'erreurs, que nous croyons lié aux dimensions pragmatiques, est mentionné par Zimmermann (1986) : les erreurs de connotation, qui surviennent lorsqu'une UL est utilisée sans que le locuteur y ayant recours semble conscient du poids connotatif de celle-ci. Nous pourrions envisager de classer ce type d'erreurs avec les problèmes de sens, comme le font (Granger et

⁸³ Granger et Monfort (1994) et nous-même (2005) mentionnons aussi le cas des conjonctions régies, par exemple la conjonction *QUE* régie par l'adjectif *MEILLEUR*.

Monfort, 1994) – la connotation étant un contenu informationnel véhiculé par une UL⁸⁴ –, mais il nous semble que ce contenu informationnel devient problématique seulement dans des contextes précis d'énonciation. C'est pourquoi nous jugeons préférable de laisser ce type de problèmes parmi les erreurs lexicopragmatiques.

3.6 UN MOT SUR LES PROBLÈMES DE STYLE

Certaines des typologies étudiées contiennent des classes d'erreurs liées aux dimensions stylistiques de l'emploi du vocabulaire. Patris et Vansnick (1992) évoquent notamment les problèmes de répétition et ceux d'imprécision (emploi de périphrases ou de mots « passepartouts » comme *CHOSE* ou *ÇA*). Nous avons nous-même (2005) repris ces classes en y ajoutant les problèmes de pléonasmes, dont nous croyons aujourd'hui qu'ils devraient plutôt faire partie des problèmes de sens, un pléonasma découlant de la cooccurrence sémantique de deux sens dont l'un inclut l'autre. Par ailleurs, nous nous interrogeons sur la pertinence d'intégrer les problèmes stylistiques mentionnés par Patris et Vansnick (1992) dans une typologie des problèmes lexicaux.

Nous avons établi à la section 3.1.2 qu'une organisation des classes d'erreurs en fonction des dimensions de l'UL serait la plus pertinente pour structurer une typologie des problèmes lexicaux. Or, le choix d'un tel classement nous pousserait à ne pas inclure dans notre typologie les problèmes liés aux aspects plus stylistiques de l'utilisation du vocabulaire, notamment la répétition et l'imprécision; en effet, de tels problèmes ne sont pas liés à des propriétés inhérentes aux UL ni à des incompatibilités entre diverses UL. De plus, il serait très difficile d'évaluer l'imprécision dans la perspective quantitative qui caractérise notre approche. Pour ce qui est de la répétition, quant à elle quantifiable, il serait aussi difficile d'établir dans quels cas elle est acceptable et dans quels cas elle ne l'est pas : en effet, il est tout à fait normal qu'un texte comporte un certain nombre de répétitions, que ce soit du fait du thème exploité

⁸⁴ Granger et Monfort notent (1994) le traitement inégal que font les dictionnaires de la dimension connotative, qui est parfois présentée comme faisant partie intégrante du sens.

– on a qu'à penser au nombre de fois que l'expression erreur lexicale est employée dans cette thèse! – ou de la présence d'UL fréquentes dans la langue (préposition, verbes auxiliaires, collocatifs fréquents). Comment pourrait-on décider des items lexicaux qui peuvent être répétés et de ceux dont la répétition est à éviter sans tomber une fois de plus dans la subjectivité.

L'analyse des problèmes d'imprécision et de répétition est donc soumise à une trop grande subjectivité pour qu'il soit envisageable de les observer de la même façon que les autres types d'erreurs lexicales présentés jusqu'à maintenant. Si ces problèmes lexicostylistiques ne sont certes pas triviaux et constituent sans doute une préoccupation pour les enseignants, nous croyons qu'ils débordent l'analyse des erreurs lexicales telle que nous la concevons dans cette thèse et faisons le choix de ne pas en tenir compte dans notre analyse.

CONCLUSION

La présente section nous a permis de caractériser plus finement la notion d'erreur lexicale et de poser les grands principes qui devraient structurer une typologie descriptive des erreurs lexicales, soit son caractère résolument descriptif, son organisation en fonction des dimensions de l'UL et le haut degré de raffinement des classes d'erreurs. C'est en nous appuyant sur les classements présentés et critiqués ici que nous proposerons notre propre grille d'analyse des erreurs lexicales (voir VII-3.1). Mais pour l'instant, intéressons-nous à un autre axe d'analyse des erreurs lexicales : leur explication.

4. L'ANALYSE DES SOURCES DES ERREURS : VERS UNE TYPOLOGIE EXPLICATIVE DES ERREURS LEXICALES

Comme nous l'avons établi, l'analyse d'erreurs doit bien sûr s'intéresser à la description des erreurs, mais aussi à leur explication. Nous avons aussi mentionné qu'une des meilleures façons de ne pas confondre ces deux axes d'analyse est d'utiliser des grilles d'analyse distinctes pour la description et l'explication des erreurs. La section 3 nous a permis de déterminer quels seraient les types d'erreurs lexicales les plus pertinents à inclure dans une typologie descriptive de ce type de problèmes. Nous nous attarderons maintenant sur les éléments que pourrait contenir une grille explicative des problèmes lexicaux. Pour ce faire, nous ferons un survol de certaines typologies explicatives, principalement proposées dans le cadre de travaux en L2; comme ces travaux sont généralement assez peu éclairants pour l'explication d'erreurs en L1, nous ne ferons ressortir que les éléments que nous jugeons pertinents dans le cadre de cette thèse.

Nous avons mentionné plus haut la place prépondérante qu'occupe l'opposition erreurs interlinguales / erreurs intralinguales dans les travaux sur l'analyse d'erreurs et souligné l'impossibilité, dans bien des cas, de classer avec assurance les erreurs dans l'une ou l'autre des catégories. Pour rendre compte de ce problème, Granger et Monfort (1994) ont proposé de diviser les erreurs interlinguales en deux classes : celles qui ne s'expliquent que par le transfert et celles pour lesquelles le transfert ne constitue qu'une explication parmi d'autres. Malgré cette tripartition des erreurs, le classement explicatif de Granger et Monfort n'est pas suffisant, puisqu'il ne contient pas d'explications réelles des problèmes; le fait de dire qu'une erreur est intralinguale ne nous renseigne pas sur la cause de l'erreur en tant que telle. Tout comme leur typologie descriptive, les grandes divisions proposées demanderaient à être raffinées. En effet, le simple fait de placer une erreur dans l'une des trois grandes classes suggérées implique que l'on formule à propos de cette erreur une hypothèse d'explication;

or, ces hypothèses ne figurent pas dans le classement de Granger et Monfort (1994).

Certains travaux sur l'erreur lexicale formulent des hypothèses plus précises sur les sources des erreurs observées. Voici un survol de ceux que nous jugeons les plus pertinents.

4.1 WARREN (1982)

Warren suggère des pistes très intéressantes pour l'explication de certains problèmes lexicaux. Sa typologie des erreurs lexicales a été construite de façon empirique, à partir de l'observation systématique des erreurs produites par ses étudiants à l'université, des étudiants suédois dont l'anglais est la L2. Le classement proposé par l'auteure se situe résolument du côté de l'explication et distingue quatre classes majeures d'erreurs lexicales.

1. **Les erreurs d'équivalence** : Ces erreurs, les plus fréquentes dans le corpus recueilli par Warren, interviennent lorsque la correspondance entre une UL de la L1 et une lexie de la L2 n'est pas univoque, c'est-à-dire que cette correspondance peut fonctionner dans certains contextes, mais pas dans tous. Warren inclut aussi dans cette classe les faux amis ainsi que les problèmes relatifs au comportement syntaxique (régime, invariabilité, etc.). Toutes les erreurs regroupées ici découlent d'interférence avec la L1 et constituent donc des erreurs interlinguales⁸⁵; par conséquent, elles seraient moins susceptibles de survenir dans un contexte de L1.
2. **Les confusions conceptuelles** : Les confusions conceptuelles concernent des paires d'UL de la L2 qui partagent des composantes de sens centrales, mais qui présentent des différences marquées dans leurs conditions d'utilisation (ex. WEDDING ~ MARRIAGE, SEPARATE ~ DIVIDE). Ce sont des erreurs intralinguales.
3. **Les erreurs de dérivation** : Les erreurs de dérivation que présente l'auteure ne découlent pas normalement d'interférences avec la L1; elles sont donc intralinguales. Il en existe plusieurs types :

⁸⁵ Il semble en effet que Warren classe les erreurs lexicogrammaticales parmi les erreurs d'équivalence, ce qui voudrait dire que ces problèmes découlent d'une interférence avec une unité lexicale de la L1. Ce choix nous paraît discutable, puisqu'il est tout à fait possible que l'apprenant suppose une équivalence des propriétés grammaticales ou syntaxiques entre deux unités lexicales de la L2; il s'agirait dans ce cas d'une erreur intralinguale.

- a. **Forme inexistante** : Ces UL fictives sont formées par l'application de patrons de dérivation par ailleurs existants dans la langue.
 - b. **Confusion entre les acceptions** : Dans ce cas, la forme dérivée existe réellement, mais elle n'a de sens qu'en relation avec une autre acception du vocable en jeu que celle qui était visée par la personne ayant commis l'erreur. L'auteure donne l'exemple de la phrase *#The meaning of the plan was to...* dans laquelle MEANING a été utilisé par le scripteur pour exprimer 'intention'; le fait qu'une des acceptions du verbe [to] MEAN corresponde au nom MEANING (*Magnanimous means 'generous'. = The meaning of magnanimous is 'generous'*) a poussé le scripteur à utiliser MEANING pour nominaliser une autre acception du verbe, qui cette fois correspond plutôt à INTENTION (*He meant to hurt Gary. = His intention was to hurt Gary.*).⁸⁶
 - c. **Confusion entre les différents dérivés** : Dans un tel cas, la confusion entre deux UL découle du fait que les deux termes sont dérivés d'une même base, par exemple ECONOMIC et ECONOMICAL.
 - d. **Confusion entre les bases** : Dans certains cas, beaucoup plus rares, Warren a constaté que les étudiants avaient de la difficulté à associer un mot dérivé à sa base. Elle donne comme exemple l'adjectif DEDUCTIBLE qui serait associé au verbe [to] DEDUCE (plutôt qu'au verbe [to] DEDUCT).
4. **Confusion phonétique (ou incertitude formelle)** : Les formes produites dans de tels cas partagent avec les formes visées une parenté phonétique plus ou moins grande. Il peut arriver que la forme produite soit inexistante ou alors qu'elle corresponde à celle d'un autre mot-forme de la langue, auquel cas Warren suggère que la forme visée soit non seulement rétablie, mais que le sens de l'UL correspondant à la forme erronée soit aussi expliqué à l'étudiant. Les erreurs de ce type sont intralinguales⁸⁷.

Malgré les quelques réserves que nous venons de formuler et le fait qu'elle ne couvre pas l'ensemble des problèmes lexicaux, mais seulement ceux présents dans le corpus étudié⁸⁸, la typologie explicative de Warren (1982) propose des

⁸⁶ En termes propres à la Théorie Sens-Texte, nous pourrions dire qu'il s'agit d'un problème de dérivation sémantique (voir II-4.4).

⁸⁷ Encore une fois, l'assimilation de ce type d'erreurs à une source intralinguale nous paraît hasardeuse, puisqu'il est tout à fait envisageable qu'une distortion formelle intervienne sous l'influence d'une UL de la L1, ce qui mène à une erreur interlinguale.

⁸⁸ Sauf les erreurs d'orthographe, qui ont été ignorées, et les problèmes mettant en jeu des expressions idiomatiques, qui ne sont qu'effleurés, mais sans être intégrés à l'une ou l'autre des classes d'erreurs de la typologie.

pistes intéressantes pour l'explication de certaines erreurs. La discussion à propos des erreurs de dérivation, notamment, devrait s'avérer utile pour l'explication de certains problèmes de sens. En ce qui concerne les problèmes de confusion conceptuelle – catégorie qui mériterait selon nous d'être raffinée –, nous constatons que Warren privilégie, tout comme Zimmermann (1987), une analyse sémantique visant une mise en lumière des composantes de sens distinctives plutôt qu'une décomposition sémantique exhaustive dans l'explication d'erreurs lexicales.

4.2 ZUGHOUL (1991)

La typologie de Zughoul a été élaborée de façon empirique à partir de la correction de 3 textes rédigés par 128 étudiants arabophones apprenants de l'anglais. L'intérêt de la démarche de Zughoul réside dans le fait que la correction des productions a été faite lors de miniconférences avec les étudiants, de façon à ce qu'ils puissent formuler ce qu'ils voulaient dire dans les passages erronés. Une telle démarche diminue énormément la part de subjectivité lors de l'étape de restitution et permet en outre d'établir avec beaucoup plus d'assurance la source des erreurs rencontrées.

Six-cent-quatre-vingt-onze erreurs lexicales ont été identifiées, puis regroupées par Zughoul dans 13 classes d'erreurs. Les principaux problèmes découlaient d'une confusion entre des termes sémantiquement proches (considérés comme synonymes par l'apprenant) ou présentant une certaine parenté formelle, d'une traduction littérale, d'une mauvaise dérivation ou relevaient du phénomène de collocation.

Bien que la typologie de Zughoul se veuille explicative, plusieurs des catégories qu'il propose relèvent selon nous de la description (paraphrase, collocation, etc.). Certaines classes semblent aussi se recouper; les exemples présentés pour les problèmes d'idiomaticité pourraient dans bien des cas être classés parmi les erreurs de collocation, de traduction littérale ou d'influence de l'arabe. Par ailleurs, son classement est résolument orienté vers la description d'erreurs

lexicales en L2 et s'avère peu utile pour les problèmes en L1; il démontre par contre la prépondérance des problèmes de confusion entre des termes « synonymes », problèmes auxquels s'est précisément intéressée l'auteure dont nous parlerons maintenant.

4.3 MARTIN (1984)

Martin constate que le nombre d'erreurs de vocabulaire commises par des apprenants d'un niveau avancé augmente avec leur compétence dans la L2 et que la majorité de ces erreurs sont intralinguales. Selon elle, la plupart des erreurs découlent de fausses hypothèses par rapport à l'UL erronée et sont une conséquence de l'enseignement du vocabulaire à l'aide de « synonymes » ou de gloses. L'auteure distingue quatre types de problèmes.

The four dissonances that interfere with encoding by advanced learners are stylistic (dunk the chicken in the beaten egg mixture); syntactic (the author purports that tobacco is harmful); collocational (I used to be a large smoker); and semantic (I was badly damaged in the accident). (Martin, 1984: 130)

Toutes ces dissonances peuvent, selon l'auteure, être expliquées par le fait que l'apprenant transpose au mot qu'il utilise les propriétés stylistiques, syntaxiques, collocationnelles et sémantiques d'un autre terme qu'on lui a présenté comme un synonyme (ou qu'il considère comme tel); cela rejoint en quelque sorte l'idée d'analogie (Frei, 1929) évoquée plus haut et autour de laquelle est construite la typologie de Patris et Vansnick (1992). Cette façon d'expliquer certains problèmes de collocation rappelle aussi la notion de *greffe collocationnelle* proposée par Polguère (2007), que nous verrons au point 4.5.

En ce qui concerne les dissonances de nature sémantique, Martin (1984) donne une liste d'oppositions (reproduite ci-dessous) qui peut s'avérer extrêmement précieuse pour cibler la composante de sens problématique à l'origine de l'erreur.

Tableau 9. Oppositions involved in semantic dissonances (Martin, 1984)

OPPOSITIONS	EXAMPLES
SUPERORDINATE/HYPONYM	<i>correct ~ rectify</i>
PHYSICAL/NON-PHYSICAL	<i>wobble ~ fluctuate</i>
STATE/EVENT	<i>exist ~ happen</i>
ANIMATE/INANIMATE	<i>injure ~ damage</i>
PLANNED/UNPLANNED	<i>spectator ~ onlooker</i>
± SOURCE-ORIENTED	<i>take back ~ return</i>
DEGREE OF INTENSITY	<i>surprise ~ flabbergast</i>
VOLITIONAL/NONVOLITIONAL	<i>deter ~ prevent</i>
± APPROBATION	<i>persevere ~ continue ~ persist</i>

4.4 LAUFER (1992)

Laufer s'est intéressée particulièrement aux problèmes de confusions formelles. Elle appelle ces « paronymes » des *synforms* et distingue différents liens sur lesquels s'appuie cette parenté formelle (même nombre de syllabes, même début, phonèmes communs, etc.). Pour que l'on puisse parler de *synform*, il faut que le mot utilisé existe réellement; les faux amis correspondent au cas où le lien de ressemblance existe entre un mot de la L1 et un de la L2. Laufer identifie dix catégories de *synforms*, qu'elle regroupe en quatre grands types communs :

1. Le type « suffixe » (*consider<able>*, *consider<ate>*)
2. Le type « préfixe » (*<com>press*, *<sup>press*)
3. Le type reposant sur des voyelles (*seat/set*, *manual/menial*)
4. Le type reposant sur des consonnes (*prize/price*, *ledge/pledge*)

Dans le même ordre d'idées, Zimmermann (1987) propose d'élargir la conception de « *structural slip* » de Hotopf (1983) pour établir une échelle de similarité structurelle à trois niveaux, reproduite ci-dessous.

strong	same word structure, unit-size correspondence (morphemes, syllables, clusters)	<i>retirement - retiring</i>
middle	only one of the two above (i.e. roughly <u>Hotopf</u>)	<i>retirement - resign</i>
weak	purely quantitative; including isolated phonemes	<i>retirement - resignation</i>

Figure 6. Levels of structural similarity (Zimmermann, 1987: 56)

Des classements comme ceux de Laufer (1992) et Zimmermann (1987) permettent d'expliquer avec précision la source des problèmes de confusion formelle.

4.5 POLGUÈRE (2007)

Polguère s'intéresse quant à lui à la description de problèmes de collocation, entendus principalement dans des émissions de radio. Il s'appuie sur un ensemble de concepts propres à la théorie Sens-Texte pour proposer un appareillage théorique permettant une modélisation, à l'aide des fonctions lexicales, et une explication de ce type de problèmes. La greffe collocationnelle est définie ainsi par Polguère (2007: 260) :

Une greffe collocationnelle est une expression produite à l'oral ou à l'écrit qui présente les caractéristiques suivantes.

1. *elle est constituée d'au moins deux éléments lexicaux A_1 (par ex., chaleur) et B_2 (par ex., de plomb);*
2. *elle fait penser à une collocation A_2+B_2 bien formée (par ex., soleil de plomb);*
3. *ce n'est pas une collocation bien formée;*
4. *on a des raisons de penser qu'elle a été produite par interférence inter- ou intralinguistique mettant en jeu une collocation valide A_1+B_1 (par ex., chaleur insoutenable), initialement visée par le locuteur, et une collocation également valide A_2+B_2 (soleil de plomb) ayant parasité la production de A_1+B_1 .*

Les greffes collocationnelles analysées par Polguère témoignent exclusivement d'interférences intralinguistiques. Le patron de greffe collocationnelle typique anticipé par Polguère – soit la greffe d'un collocatif erroné à une base sous l'influence d'une base synonyme – se retrouve effectivement dans les données observées, mais les cas étudiés font aussi émerger d'autres patrons possibles, notamment la « greffe inversée », dans laquelle l'élément greffé erroné est la base plutôt que le collocatif (ex. **prendre toutes les déférences** qu'il faut au lieu de **prendre toutes les précautions**). La méthode d'analyse proposée dans cet article s'avère tout à fait pertinente à la fois pour la description et l'explication des problèmes de collocation et pourrait s'avérer précieuse pour dégager des sous-

groupes en fonction des partons de greffes parmi les problèmes que nous recueillerons dans notre corpus.

4.6 PATRIS ET VANSNICK (1992)

L'article de Patris et Vansnick donne des pistes intéressantes pour l'explication de certains types de problèmes lexicaux. Les auteurs mentionnent notamment que plusieurs erreurs découlent d'une ellipse, qui correspond au besoin d'économie linguistique de Frei (1929). L'exemple fourni est *#Les photos marquées d'une barre étaient **décédées** récemment.*, dans lequel l'incompatibilité sémantique entre *photos* et *décédées* peut être expliquée par le fait que l'élève ayant produit cette phrase a utilisé l'adjectif DÉCÉDÉ non pas pour qualifier les photos elles-mêmes, mais bien les personnes représentées sur ces photos.

En ce qui concerne les problèmes liés à l'emploi de locutions, Patris et Vansnick mentionnent aussi différentes sources d'erreurs, notamment l'amalgame de deux expressions (ex. *dès leur jeune enfance*), l'addition/suppression d'un élément (ex. *Il se mit à pleurer toutes les larmes.*) et la substitution d'un terme à un autre (ex. *La situation est sans sortie.*).

Bien entendu, c'est l'analogie qui occupe la plus grande place dans les travaux de Patris et Vansnick et elle permet d'expliquer tant les problèmes de confusion entre des paronymes que des confusions sémantiques ou des problèmes de régime (ex. *On instruit **aux enfants** trop de choses inutiles.*).

4.7 QUELLE STRUCTURE POUR UNE TYPOLOGIE EXPLICATIVE DES ERREURS?

Il est important de rappeler que l'étape d'explication des erreurs est celle qui laisse la plus grande place à la subjectivité dans le processus d'analyse d'erreurs. Ceci est particulièrement vrai dans le cas où les hypothèses d'explication sont formulées seulement à partir de productions linguistiques contenant des erreurs et sans avoir recours à la personne les ayant commises. Plusieurs scénarios d'explication seront envisageables dans bien des cas et seul le fait d'interroger le locuteur à l'origine de l'erreur pourrait nous éclairer sur celui à privilégier.

Le survol des recherches présentées nous permet de constater que certaines pistes ont été avancées pour l'explication des erreurs lexicales, mais que nous disposons de balises moins précises pour l'explication des erreurs que pour leur description. Les sources des erreurs intralinguales ont aussi été beaucoup moins étudiées, et les travaux consultés nous laissent avec moins d'indices sur les sources des problèmes en L1.

De façon générale, l'analogie semble servir de toile de fond à beaucoup des scénarios d'explication avancés par les auteurs. Elle pourrait donc servir de classe englobante pour différents scénarios d'explication et constituer l'une des grandes divisions d'une typologie explicative des erreurs lexicales. Certaines sources que l'on pourrait qualifier d'extralinguistiques – par exemple la surcharge cognitive, les erreurs d'inattention –, et qui ne sont pas propres aux erreurs lexicales, pourraient aussi être regroupées. Par contre, les autres grandes divisions qui pourraient structurer une telle typologie ne nous apparaissent pas clairement; ainsi, il nous est impossible, ici, de jeter les bases d'une typologie explicative idéale comme pour l'avons fait plus haut pour une grille descriptive.

CONCLUSION

Ce chapitre nous a permis de bien délimiter le concept d'erreur lexicale et de mettre en place les éléments qui nous serviront à développer les outils que nous comptons utiliser dans l'analyse des erreurs lexicales de notre corpus de textes. Nous présenterons notre démarche d'analyse et nos outils au chapitre VII.

VI. LE RAPPORT À L'ERREUR LEXICALE DES ENSEIGNANTS

La notion de représentation intéresse aujourd'hui plusieurs sciences sociales. En sciences de l'éducation, notamment, l'étude des représentations et des attitudes constitue un champ de recherche important depuis que l'on considère, notamment à la suite des travaux de Giordan et de Vecchi (1987), que les représentations des apprenants peuvent constituer des obstacles à l'apprentissage. En didactique du français, on s'intéresse aux représentations des élèves, puisque l'on sait que celles qu'ils ont sur la langue « influencent largement les procédures et les stratégies qu'ils développent et mettent en œuvre pour apprendre cette langue et en user. » (Moore, 2004b: 9) Parce qu'elles orientent leurs pratiques et leur enseignement (Gilly, 1989 ; Garcia-Debanc et Plane, 2004), les représentations des enseignants font aussi l'objet de plusieurs travaux.

La première partie de ce chapitre visera à définir le concept de représentation et à le mettre en relation avec d'autres termes souvent utilisés dans les travaux sur les représentations, soit ceux de conception et d'attitude. Nous verrons ensuite pourquoi nous jugeons important d'aller au-delà des simples représentations pour éclairer le rapport d'un sujet à un objet en nous intéressant à la notion de « rapport à », que nous illustrerons à l'aide de travaux sur le rapport à l'écrit. Nous caractériserons finalement une notion centrale de notre thèse : le rapport à l'erreur lexicale des enseignants.

1. REPRÉSENTATIONS, CONCEPTIONS ET ATTITUDES

La section qui suit propose un bref détour du côté de la psychologie sociale, d'où provient le concept de *représentation*, avant d'explorer quelques termes qui y sont rattachés. Nous présenterons en conclusion les définitions de deux de ces termes, tels que nous les entendons dans cette thèse.

1.1 LES REPRÉSENTATIONS SOCIALES

Le concept de *représentation sociale* a été proposé à la fin du XIX^e siècle par le sociologue français Émile Durkheim (1898). Réintroduite et raffinée par Moscovici (1961) dans le cadre d'une recherche sur les représentations de la psychanalyse, cette notion a depuis investi le champ de nombreuses sciences sociales (sociologie, psychologie, anthropologie, sociolinguistique, etc.) et donné lieu à une quantité impressionnante de recherches. Notion fuyante et difficile à cerner, ayant une « position "mixte", au carrefour d'une série de concepts sociologiques et d'une série de concepts psychologiques » (Moscovici, 1976: 39), la représentation sociale n'en influence pas moins nos comportements, nos rapports aux autres, notre façon de penser.

Différentes disciplines s'étant approprié la notion de représentation sociale, il est difficile d'en proposer une définition universelle et transversale. Nous tenterons donc simplement de faire ressortir les éléments communs à différentes caractérisations.

Jodelet (1989: 53) définit la représentation sociale comme une « forme de connaissance, socialement élaborée et partagée, ayant une visée pratique et concourant à la construction d'une réalité commune à un ensemble social ». Il s'agit en quelque sorte de schémas de connaissances que l'on se construit, en dehors des cadres de référence initiaux (ex. théorie scientifique), par des informations glanées auprès de diverses sources (médias, conversations, livres, etc.), et qui nous permettent de fonctionner au quotidien (Reuter *et coll.*, 2007). Lüdi et Py (2002: 98) parlent de « microthéories socialement partagées et prêtes à l'emploi, suffisamment vagues pour faciliter un large consensus et une application étendue. »

Ces définitions font ressortir le caractère systémique de la représentation sociale, qui consiste en une articulation d'informations, plus ou moins en accord avec les savoirs scientifiquement admis, partagée par un groupe d'individus et qui leur permet d'appréhender les réalités de la vie quotidienne, c'est-à-dire penser, communiquer et agir. Abric (2003) s'intéresse précisément à la structure

de ce « système socio-cognitif présentant une organisation spécifique » (p.59). C'est ainsi qu'il propose sa théorie du noyau central : ce regroupement de quelques éléments constitue le cœur de la représentation autour duquel gravitent les autres éléments. Cette façon d'envisager la représentation sociale permet non seulement de dégager le contenu d'une représentation, mais aussi d'en saisir l'organisation par la prise en compte du poids relatif des différents éléments qui la composent.

Une autre caractéristique importante des représentations sociales semble faire l'unanimité : elles sont intimement liées au discours (Lüdi et Py, 2002). Non seulement c'est à travers lui qu'elles émergent, mais c'est par le discours qu'elles prennent forme et se transforment (Gajo et Serra, 2000 ; Cavalli *et coll.*, 2004). Ce point fait aussi ressortir le fait que les représentations sociales ne sont pas directement observables, elles doivent être inférées à partir « [d']observables médiatisés par les "dires" des enquêtés » (Billiez et Millet, 2004: 34) : comportements, productions, verbalisations, etc. (Delcambre et Reuter, 2002).

Les représentations sociales remplissent plusieurs fonctions au sein des groupes qui les partagent. Abric (1994a) en distingue quatre :

1. *fonctions de savoir*, puisqu'elles permettent de comprendre et d'expliquer la réalité;
2. *fonctions identitaires*, en ce qu'elles permettent l'élaboration d'une identité sociale (et sa sauvegarde);
3. *fonctions d'orientations*, parce qu'elles guident les comportements et les pratiques;
4. *fonctions justificatrices*, comme elles servent de justification à posteriori des prises de position et comportements.

1.2 LES REPRÉSENTATIONS EN SCIENCES DE L'ÉDUCATION

Depuis l'avènement du constructivisme, l'apprentissage est envisagé comme un processus de restructuration de systèmes explicatifs déjà existants plutôt que d'être perçu comme la simple intégration de nouvelles connaissances. Il est donc primordial, pour le monde de l'éducation, de s'intéresser aux représentations présentes chez les apprenants, puisqu'elles font partie intégrante de cette

restructuration visée par l'enseignement et jouent un rôle central dans son succès. En effet, tour à tour des obstacles ou des points d'appui (Halté, 1992), elles influencent grandement les processus d'enseignement-apprentissage. C'est ainsi que les sciences de l'éducation se sont aussi saisies du concept de représentation, ouvrant ainsi la porte à un nouveau courant de recherche, dont l'un des précurseurs fut Migne (1969).

Dans les travaux en éducation, la représentation est la plupart du temps appréhendée d'un point de vue individuel plutôt que social; on s'intéresse aux représentations et connaissances⁸⁹ qu'ont, de façon individuelle, les apprenants⁹⁰.

Dans ces usages dans les différentes sciences sociales, la notion de représentation est reprise dans une acception libre, dans laquelle le caractère social n'est pas toujours affiché ni étudié, voire explicitement non pris en compte. Ce qui est repris, et en particulier dans les quelques investigations qui portent précisément sur les représentations de l'écriture, c'est le sens commun du terme représentation, à savoir que la représentation est une construction, ou une reconstruction, du réel par le sujet. (Barré-de Miniac, 2000: 58)

Ainsi, la représentation n'est plus nécessairement envisagée comme une construction sociale, mais comme un mode de pensée propre à un individu⁹¹. L'acception première du terme *représentation*, en éducation, correspond donc davantage à l'usage qu'on en fait en psychologie et dénote des informations stockées dans la mémoire permanente (Clément, 1994) et qui constituent autant d'éléments du système de pensée de l'apprenant.

Reuter *et coll.* (2007: 197) définissent avec une efficace simplicité ce que les chercheurs en didactique entendent généralement par *représentation*: « En didactique, la notion de représentation a été définie pour parler des connaissances qu'un sujet mobilise face à une question ou à une thématique, que

⁸⁹ Pour une distinction entre les notions de *représentation* et de *connaissance*, voir 1.3.

⁹⁰ Bien entendu, les chercheurs s'intéresseront aussi à dégager les représentations récurrentes chez plusieurs apprenants, mais le terme de *représentation sociale* ne semble pas utilisé en ce sens.

⁹¹ Il est tout de même important de souligner qu'il existe aussi des recherches en éducation qui s'intéressent aux représentations sociales, notamment en lien avec l'apprentissage d'une L2. Voir par exemple les différents travaux présentés dans Moore (2004b).

celle-ci ait fait l'objet d'un enseignement ou pas.» Évidemment, ces connaissances pourront correspondre ou non aux savoirs scientifiques (ou, plus largement, aux savoirs propres aux différents cadres de référence) relatifs à cette question ou thématique; il se peut d'ailleurs que des représentations erronées cohabitent avec des connaissances savantes (Barré-de Miniac, 2007). C'est donc principalement le caractère non social qui distingue la représentation entendue en ce sens de la représentation sociale, puisqu'elle remplit aussi des fonctions de compréhension ~ explication, d'orientation et de justification.

Par ailleurs, différentes représentations, contradictoires lorsqu'on les analyse objectivement, peuvent cohabiter chez un même individu (Clément, 1994). Cette concurrence des représentations peut par exemple découler du fait qu'un individu appartient à la fois à plusieurs groupes sociaux distincts ayant chacun ses représentations sociales propres; des représentations différentes sont donc sollicitées en fonction des situations auxquelles fait face l'individu. Par conséquent, les recherches s'intéressant aux représentations ont donc pour limite de ne décrire qu'une partie des représentations des sujets : celles qui sont sollicitées dans la réalisation de la tâche à accomplir. Clément (1994: 21) nomme *conceptions conjoncturelles* ces représentations « qui sont mobilisées (en mémoire de travail) dans [une] situation précise (dialogue, apprentissage, réalisation d'une tâche...) ». L'auteur mentionne aussi que le fait de ne plus voir émerger, dans la mémoire de travail, des conceptions conjoncturelles erronées témoigne qu'il y a eu restructuration du système conceptuel dans la mémoire permanente de l'apprenant, donc apprentissage.

1.3 REPRÉSENTATIONS ET CONCEPTIONS : DU PAREIL AU MÊME?

À cause du caractère très polysémique de *représentation* et du fait que ce terme soit utilisé par plusieurs sciences humaines pour désigner des concepts différents, certains auteurs privilégient l'emploi de *conception* (Giordan et de Vecchi, 1987 ; Clément, 1994). Selon Giordan et de Vecchi (1987), le fait d'utiliser plutôt le terme *conception* permet de prendre une certaine distance par rapport aux travaux sur les représentations sociales, qui tentent de découvrir les

dénominateurs communs d'un groupe social plutôt que de comprendre le système de pensée propre à un individu particulier (ce système étant évidemment influencé par les représentations sociales, mais ne pouvant y être réduit).

Au-delà de ces choix terminologiques, il ne semble pas qu'il y ait de différences fondamentales entre les représentations, telles qu'entendues en didactique, et les conceptions, que Giordan et de Vecchi (1987: 77) définissent comme des « structures mentales mises en œuvre face à des situations-problèmes particulières. »

À notre connaissance, seuls Coquin-Viennot et Gaonac'h (1995) tentent de distinguer les conceptions des représentations en proposant une caractérisation de la conception qui en fait une sous-classe des représentations. Selon eux, une conception est une représentation qui respecte les trois conditions suivantes :

1. **communauté**, c'est-à-dire qu'elle est partagée par un nombre non négligeable d'apprenants d'un même niveau.
2. **efficacité**, c'est-à-dire qu'elle permet de résoudre un certain nombre de problèmes dans une situation donnée. Évidemment, son application dans de nouvelles conditions pourra conduire à des erreurs.
3. **stabilité**, c'est-à-dire que, sans être immuables, les conceptions ne sont pas non plus des « représentations construites dans l'instant, liées au contexte et qui s'effacent aussi vite qu'elles s'élaborent » (p.302).

Ces critères, bien qu'intéressants, ne semblent pas faire l'objet d'un consensus et la distinction entre *conception* et *représentation* demeure difficile à cerner. La difficulté de distinguer les deux termes ressort de manière évidente si l'on souligne que les critères proposés par Coquin-Viennot et Gaonac'h pour définir la conception ont aussi été utilisés pour définir la notion de *représentation sociale*. Par conséquent, plusieurs auteurs continuent à utiliser les deux termes de manière assez équivalente (Nadeau et Fisher, 2006: 154). Il semble donc que la recherche sur les représentations en didactique s'est jusqu'à ce jour accommodée de cette (quasi-)synonymie entre *conception* et *représentation*.

Bien que nous n'ayons pas nous-même de critères distinctifs à proposer, quelques observations émanent de nos lectures.

Tout d'abord, il semble que l'objet de la représentation oriente en quelque sorte le choix de l'un ou l'autre des termes, du moins dans certains cas. En effet, même dans les textes d'auteurs chez qui cohabitent *conception* et *représentation* (la plupart), certains emplois semblent plus contraints, le terme *conception* apparaissant plus approprié lorsque l'objet de la représentation est réellement un concept. Par exemple, dans une étude sur les représentations de l'enseignement du lexique par des enseignants du primaire, alors qu'elle utilise davantage *représentation*, Dreyfus (2004) parle de *conception du sens lexical*. Le terme *représentation*, quant à lui, s'accorde plus spontanément à des dimensions plus sociales ou à des éléments que l'on hésiterait à qualifier de concepts (*représentation du rôle de l'élève/de l'enseignement/de la tâche à accomplir*, etc.), même si son champ d'application semble plus large que celui de *conception*.

Ensuite, le champ de recherche semble aussi influencer le choix des termes utilisés. Le terme *conception*, défendu notamment par Giordan et de Vecchi (1987), semble plus caractéristique des travaux en didactique des sciences et des mathématiques. La didactique des langues, quant à elle, privilégie l'emploi de *représentation*, peut-être parce que la dimension culturelle (et donc sociale) est plus fortement liée à la langue. Bien sûr, ceci témoigne d'une tendance observée et nous n'affirmons pas que l'un des termes soit associé exclusivement à un champ de recherche particulier. D'ailleurs, nous avons fréquemment vu le terme *conception* utilisé en didactique des langues (notamment par Nadeau et Fisher, 2006), Migne (1994) parle de représentations relativement à des concepts de physique... et nous parlerions plus spontanément de *conception de l'erreur lexicale* que de *représentation de l'erreur lexicale*! Nous jugeons donc que les termes *conception* et *représentation* sont équivalents dans une perspective didactique, mais que leur emploi peut varier selon le domaine de recherche.

Une autre opposition présente dans certains travaux est celle entre les termes *conception* et *connaissance*, l'emploi du premier étant réservé aux connaissances

qui ne correspondent pas aux savoirs scientifiques, ce que les anglophones appellent « *misconceptions* ». Clément (1994) s'oppose à cette idée pour deux raisons. Premièrement, les connaissances scientifiques ne sont pas scientifiques par essence, mais bien parce qu'une communauté scientifique les reconnaît comme telles, ce qui en fait donc des conceptions socialement admises. Deuxièmement, nos structures mentales ne séparent pas spontanément ce qui est scientifiquement fondé de ce qui ne l'est pas. Ainsi, « nos conceptions forment un tout où se mêlent connaissances scientifiques, croyances, idéologies, fonctionnalités sociales, dimensions rationnelles et esthétiques, émotionnelles, affectives, ... » (Clément, 1994: 28). Un dernier argument pourrait nous convaincre de ne pas opposer *conception* et *connaissance* : le fait que la conception, comme la représentation, est en fait un ensemble structuré de connaissances.

1.4 LES ATTITUDES

La notion d'*attitude* provient de la psychologie sociale et ne fait pas non plus l'objet d'une définition consensuelle (Grawitz, 2001 ; Lüdi et Py, 2002). Une caractéristique des attitudes semble généralement admise : tout comme les représentations, elles ne sont pas directement observables et se traduisent par les comportements et les discours qu'elles génèrent (Lüdi et Py, 2002 ; Moore, 2004a). La définition d'*attitude* que propose Legendre (2005: 138) reprend les éléments généralement soulevés par les autres auteurs :

État d'esprit (sensation, perception, idée, conviction, sentiment, etc.), disposition intérieure acquise d'une personne à l'égard d'elle-même ou de tout élément de son environnement (personne, chose, situation, évènement, idéologie, mode d'expression, etc.) qui incite à une manière d'être ou d'agir favorable ou défavorable.

Cette définition met bien en lumière le fait que les attitudes préexistent aux comportements et contribuent à les influencer (Billiez et Millet, 2004). Elle mentionne aussi le caractère acquis des attitudes; elles peuvent être acquises à la suite d'expériences en relation avec leur objet, mais aussi être influencées par (voire empruntées à) des membres de l'entourage (Lüdi et Py, 2002).

Moore (2004a) mentionne que l'on distingue généralement trois composantes dans l'attitude : la composante cognitive, la composante affective et la composante conative, qui détermine les intentions d'action et affecte les comportements. De ce point de vue, nous pourrions concevoir les attitudes comme un élément charnière entre les représentations sociales et les comportements, puisque ces représentations semblent relever des composantes cognitive et affective des attitudes. Billiez et Millet (2004) y voient pourtant deux concepts davantage imbriqués, position partagée par Doise (1989: 57), qui affirme qu'une étude en profondeur des attitudes « implique toujours qu'on les étudie aussi comme des représentations sociales. » Le fait qu'« [attitudes] et croyances fonctionnent comme justifications mutuelles » (Lüdi et Py, 2002: 97) renforce cette idée d'imbrication; Moscovici lui-même considérait l'attitude comme une des dimensions constitutives de la représentation (Gajo et Serra, 2000).

Les concepts de *stéréotype* et d'*opinion* sont aussi voisins de celui d'*attitude* avec lequel ils apparaissent souvent. Encore une fois, la démarcation demeure difficile. Lüdi et Py (2002) rapportent une distinction terminologique proposée par Kolde (1981), mais, comme le souligne Doise (1989), ce type de distinction n'est pas repris dans la recherche sur les attitudes, et *attitude* est utilisé comme terme général pour désigner aussi les opinions et les stéréotypes.

1.5 QUELQUES DÉFINITIONS

Pour clore cette section sur la distinction entre les termes *représentation*, *conception*, et *attitude*, nous proposons des définitions « maison » de ces termes, inspirées par nos lectures. Malgré la difficulté de distinguer, dans les faits, les représentations des attitudes, nous croyons que les définitions proposées sont opérationnelles. Nous ne présentons bien sûr que les acceptions pertinentes dans le contexte de cette thèse.

Représentation/conception : ensemble de connaissances (en accord ou non avec les savoirs scientifiques, ce qui peut en faire des aides ou des obstacles à l'apprentissage) ou d'idées mobilisées face à une question ou à une thématique, que celle-ci ait fait l'objet d'un enseignement ou pas.

Attitude : disposition psychique à réagir favorablement ou défavorablement à une classe d'objets (au sens large).

2. UN CONCEPT UNIFICATEUR : LE « RAPPORT À »

Comme nous l'avons constaté, la caractérisation des termes *représentation*, *conception* et *attitude* n'est pas une mince affaire, compte tenu de la polysémie de certains de ces termes et du manque de consensus qui les entoure. Si nous tentons d'intégrer à ce réseau de concepts les termes périphériques de *connaissance*, *opinion*, *perception*, *image*, *croyance*⁹², *préconception*, *stéréotype*, etc., alors l'entreprise devient carrément hasardeuse! Pourtant, malgré cette apparente confusion, un nombre important de recherches très éclairantes continuent d'être menées dans le champ des représentations/conceptions et attitudes. Nous pensons par exemple à la recherche de Millet, Lucci et Billiez (1990) sur les représentations de l'orthographe et les attitudes qui y sont liées, dans laquelle les auteurs décrivent en détail le rapport des sujets à l'orthographe sans jamais définir explicitement les termes *représentation* ou *attitude*, ni en tracer clairement les limites. Et c'est le cas de nombreux travaux sur les représentations en didactique, du moins en didactique du français (Fisher, 2004).

Si cette absence de cloisonnement entre les différents aspects du rapport à l'objet étudié n'est pas sentie comme un manque de rigueur scientifique, c'est peut-être parce que justement le rapport d'un individu à un objet ne peut être éclairé qu'en en considérant toutes les facettes, dans leur imbrication et leurs interrelations. C'est pourquoi nous avons aujourd'hui tendance à appréhender les représentations/conceptions et les attitudes comme des éléments constitutifs du rapport d'un individu à un objet. En didactique, une notion plus

⁹² Nous avons en effet rencontré le terme *croyance* comme terme central d'une recherche portant sur ce qui nous semble être les représentations d'enseignants de maternelle sur les habiletés métaphonologiques et la connaissance des lettres (Brodeur, Deaudelin, Bournot-Trites, Siegel et Dubé, 2003).

large, qui permet d'englober les différents aspects évoqués plus haut en envisageant les contenus scolaires du point de vue de l'apprenant, est aujourd'hui proposée : celle de « rapport à »⁹³.

Nous jugeons que le concept de « rapport à » légitime en quelque sorte la recherche sur les représentations/conceptions et attitudes, même dans la mesure où ces éléments sont plus ou moins clairement délimités, en permettant de les envisager comme des composantes interreliées d'une relation unissant un individu à un objet, donc en mettant en lumière la nécessité de les considérer toutes malgré les difficultés liées à leur caractérisation. Mais quelles sont donc les autres dimensions du « rapport à »? C'est la question à laquelle nous tenterons de répondre à la section suivante.

2.1 LE RAPPORT AU SAVOIR

La notion de « rapport à » provient de celle de « rapport au savoir », qui a commencé à germer dans les sciences sociales il y a plusieurs décennies. Déjà dans les années 30, Bachelard y faisait référence, sans pourtant lui accoler l'étiquette de « rapport au savoir » (Charlot, 1997). Dans les années 60 et 70, l'expression fait son apparition dans les travaux des psychanalystes (notamment du français Jacques Lacan), puis des sociologues de l'éducation. Les deux figures de proue de la recherche sur le rapport au savoir sont Bernard Charlot et l'équipe E.SCOL, qui ont éclairé cette notion du point de vue de la sociologie de l'éducation, ainsi que Jacky Beillerot et le groupe de recherche *Savoir et rapport au savoir* du CREF (Centre de recherches éducation et formation, Université de Paris 10), qui l'ont plutôt appréhendée sous l'angle de la psychanalyse.

Partant du principe que le savoir n'existe qu'en relation à un individu, qu'il s'agit d'information appropriée par un sujet et qu'il ne peut être compris ou décrit qu'en relation avec ce sujet, Charlot (1997: 94) définit le rapport au savoir comme « l'ensemble des relations qu'un sujet entretient avec un objet, un

⁹³ Reuter *et al.* (2007) incluent d'ailleurs le « rapport à » dans leur *Dictionnaire des concepts fondamentaux des didactiques*.

“contenu de pensée”, une activité, une relation interpersonnelle, un lieu, une personne, une situation, une occasion, une obligation, etc.» Cet ensemble de relations cognitives, mais aussi socio-psycho-affectives (Reuter *et coll.*, 2007: 91), ne s’applique donc pas qu’aux savoirs à proprement parler (c’est-à-dire à des contenus de pensée), mais à différentes « figures de l’apprendre », qui peuvent être aussi des savoir-faire ou des dispositifs relationnels (Charlot, 1997); pour cette raison, Charlot juge que l’expression « rapport à l’apprendre » serait plus éloquente, même si c’est l’expression « rapport au savoir » s’est répandue dans la littérature.

Si, selon Charlot (1997), Giordan est le premier chercheur à avoir utilisé l’expression « rapport au savoir » dans un contexte didactique en 1977, c’est à Yves Chevallard, didacticien des mathématiques, que l’on doit véritablement la transposition de la notion « rapport à » au domaine de la didactique à la fin des années 80. Cette notion, bien qu’encore ouverte et non stabilisée (Barré-de Miniac, 2008), donne depuis lieu à de nombreuses recherches en didactique, dont un nombre important concernent le rapport à l’écrit.

2.2 UN EXEMPLE D’OPÉRATIONNALISATION DE LA NOTION DE « RAPPORT À » : LE RAPPORT À L’ÉCRIT

La notion de « rapport à » présente sans aucun doute une piste de recherche intéressante, mais demeure peu fructueuse telle quelle. Nous avons en effet vu dans la section sur la définition que pour décrire le sens d’une UL, il faut tenir compte, notamment, de ses participants sémantiques. Or, RAPPORT étant une UL contrôlant deux actants sémantiques (*rapport de X à Y*), il est primordial d’actualiser ces variables pour pouvoir bien en définir le sens et rendre compte de façon satisfaisante de la notion de « rapport à ». Comme Barré-de Miniac (2008), nous ne croyons pas qu’il existe un *rapport au savoir* en général, dont les invariants seraient applicables à tous les domaines ; pour caractériser un rapport à un objet, il faut tenir à la fois compte de l’objet et du sujet qui entretient un rapport avec celui-ci.

C'est d'ailleurs ce qu'a fait Barré-de Miniac (2000) pour caractériser la notion⁹⁴ de rapport à l'écriture des élèves, qui « désigne des conceptions, des opinions, des attitudes, de plus ou moins grande distance, de plus ou moins grande implication, mais aussi des valeurs et des sentiments attachés à l'écriture, à son apprentissage et à ses usages. » (p.13). L'auteure en propose une opérationnalisation en quatre dimensions :

1. *L'investissement de l'écriture*, c'est-à-dire l'intérêt affectif pour l'écriture la quantité d'énergie qui y est consacrée, les aspects de l'écriture sur lesquels porte cet intérêt, ainsi que le sens de la force d'investissement (positif ou négatif);
2. *Les opinions et les attitudes*, c'est-à-dire les sentiments et les valeurs accordés à l'écriture et les attentes à son égard pour la réussite;
3. *Les conceptions de l'écriture et de son apprentissage*, appréhendées à travers leurs manifestations;
4. *Les modes d'investissement*, c'est-à-dire la manière dont les apprenants parlent de l'écriture, ainsi que les procédures et démarches qui guident tant leurs activités d'écriture que leurs activités d'apprentissage de l'écriture.

Comme le souligne la chercheuse (Barré-de Miniac, 2008: 118), le fait qu'elle propose une opérationnalisation de la notion de rapport à l'écriture en quatre dimensions ne doit pas nous faire perdre de vue que ces dernières sont « articulées » et ne doivent pas être considérées indépendamment les unes des autres. Elle met aussi en garde contre l'idée que cette liste de dimensions est close. Pour Barré-de Miniac, le rapport à l'écriture n'est pas un concept achevé en soi, mais une notion euristique pour la recherche en didactique (Barré-de Miniac, 2007). Cet avis partagé par Chartrand et Blaser (2008: 127) :

Les notions de [rapport à l'écriture] et de [rapport à l'écrit] ont permis de mettre en lumière des aspects longtemps ignorés par l'institution scolaire et peu théorisés par la didactique du français, mais il faut les opérationnaliser dans des modèles plus complexes, tâche urgente de la communauté des chercheurs.

Dans le cadre des travaux menés par l'équipe de recherche *Scriptura* portant sur les activités de lecture et d'écriture dans les classes de sciences et d'histoire du

⁹⁴ Nous choisissons, avec Barré-de Miniac (2008), de parler de notion plutôt que de concept pour désigner le rapport à l'écriture. Nous adopterons la même position relativement au « rapport à l'erreur lexicale des enseignants ». Pour une argumentation de ce choix terminologique, voir Barré-de Miniac (2008).

secondaire, Chartrand et ses collaborateurs ont proposé une caractérisation du rapport à l'écrit⁹⁵ (RÉ) qui s'appuie sur la conception de Barré-de Miniac, mais s'en distingue sur un point principal : l'intégration des pratiques de lecture-écriture comme éléments constitutifs du rapport à l'écrit. Chartrand et Blaser (2008: 113) considèrent en effet « que les activités de lecture et d'écriture des individus ne peuvent être considérées comme de simples manifestations ou émanations du RÉ; elles contribuent à le forger. » Le rapport à l'écrit est donc opérationnalisé en quatre dimensions : affective, axiologique, conceptuelle et praxéologique.

- *la dimension affective du RÉ concerne les sentiments et les émotions, voire les passions du sujet concernant l'écrit. Elle se manifeste par l'investissement en temps, en fréquence et en énergie que le sujet dit déployer en menant des activités de lecture et d'écriture (dans différents contextes) et l'intérêt affectif porté envers les différents genres de textes lus, écrits;*
- *la dimension axiologique concerne les valeurs que le sujet attribue à l'écrit pour vivre et s'épanouir dans son milieu et, dans le cas des élèves, pour réussir à l'école comme dans chaque discipline;*
- *la dimension conceptuelle (idéelle) renvoie aux conceptions, aux idées, aux représentations du sujet à propos de la nature de l'écrit, de sa place dans la société, de sa ou ses fonctions dans l'apprentissage en général, dans l'apprentissage scolaire comme dans chaque discipline scolaire ainsi que des processus de lecture et d'écriture qu'ils mettent en œuvre;*
- *la dimension praxéologique a trait aux activités des sujets en matière d'écriture et de lecture : ce qu'ils lisent et écrivent, le contexte, la manière (processus et outils), le moment et le temps investi dans ces activités. (Chartrand et Blaser, 2008: 111)*

Cette caractérisation du rapport à l'écrit nous semble particulièrement intéressante, précisément parce qu'elle permet, au-delà des dimensions cognitive et psychoaffective, de prendre en compte les pratiques comme élément constitutif du rapport à un objet, ce qui nous semble tout à fait pertinent dans un contexte didactique, surtout si l'on considère que le RÉ inclut aussi les représentations et les valeurs que le sujet associe à ces pratiques. Cette inclusion des pratiques dans la définition du rapport à l'écrit correspond d'ailleurs à la

⁹⁵ Ce rapport concerne donc tant la lecture que l'écriture, contrairement au rapport à l'écriture de Barré-de Miniac (2000). Reuter *et al.* (2007: 193) semble d'ailleurs mettre en doute la pertinence d'unifier la lecture et l'écriture sous un même « rapport à », « les caractéristiques du rapport à l'écriture ne [trouvent] pas à s'appliquer directement au rapport à la lecture ».

position adoptée par Legendre (2005) et Dabène (2008). Il est aussi à noter que, de la même façon que Dabène (1987) affirmait que les représentations, motivations et attitudes d'un sujet par rapport à l'écriture faisaient partie de sa compétence scripturale, les chercheurs s'intéressant aujourd'hui au rapport à l'écrit(ure) considèrent qu'il est une composante de la littéracie (Barrède Miniac, 2007).

Voyons maintenant de quelle façon cette caractérisation de la notion de rapport à l'écrit peut nous aider à définir le rapport à l'erreur lexicale, qui constitue l'objet de notre discussion.

3. LE RAPPORT À L'ERREUR LEXICALE DES ENSEIGNANTS

Nous avons souligné plus haut que le rapport d'un sujet à un objet devait être caractérisé non pas seulement en fonction de cet objet, mais aussi en tenant compte du sujet. C'est pour cette raison que nous choisissons de parler ici de la notion de « rapport à l'erreur lexicale des enseignants » plutôt que simplement de « rapport à l'erreur lexicale ». En effet, comme le mentionne Schneuwly (2008), le rapport à un même savoir diffèrera nécessairement selon que l'on s'intéresse aux apprenants ou aux enseignants. Nous pourrions même pousser plus loin cette idée en définissant le « rapport à l'erreur lexicale d'enseignants de français du secondaire », puisque le rapport à l'erreur lexicale d'autres enseignants risque d'être différent, voire inexistant; il se pourrait tout à fait qu'un enseignant de mathématiques, par exemple, n'ait pas de représentations relatives à l'erreur lexicale ni de pratiques s'y rapportant, puisqu'il est très rarement confronté à ce type de problèmes et n'a par conséquent pas pris le temps d'y réfléchir. Pour alléger notre propos, nous continuerons malgré tout de parler de « rapport à l'erreur des enseignants ».

Par ailleurs, nous jugeons pertinent de caractériser notre notion autour d'un objet précis et bien délimité – l'erreur lexicale – plutôt que relativement à un objet plus large comme « l'erreur linguistique », voire « l'erreur » en général. En effet, le fait de bien circonscrire l'objet du rapport aura une incidence sur les

dimensions de ce rapport que nous retiendrons dans sa caractérisation. Par exemple, la dimension affective qui pourrait sans doute caractériser le rapport à l'erreur en général ne sera pas pertinente pour la définition du rapport à l'erreur lexicale (voir plus bas). Comme le mentionnent Reuter *et coll.* (2007: 194), « les transferts ne sont pas mécaniques » et les caractéristiques du rapport d'un sujet à un contenu d'enseignement ne seront pas nécessairement les mêmes que celles liées à un autre. Par ailleurs, il se peut tout à fait que des représentations relatives à des objets plus larges (l'erreur linguistique, l'erreur) soient des éléments constitutifs du rapport à l'erreur lexicale. Comme le souligne Charlot (1997: 98) :

Le « rapport à » inclut des représentations qui ne sont pas nécessairement celles de ce dont le rapport est « à ». Ainsi, le rapport à l'école peut mettre en jeu des représentations de l'école, mais aussi de l'avenir, de la famille présente et future, du travail et du chômage dans la société de demain, des technologies modernes, etc. »

De la même façon, le rapport à l'erreur lexicale d'un enseignant pourrait être teinté de conceptions relatives à d'autres types d'erreurs linguistiques, voire à sa conception de la norme, du français québécois, de la langue, etc.

Comme la notion de « rapport à l'écrit » est sans doute le « rapport à » ayant fait l'objet d'un plus grand effort d'opérationnalisation dans un contexte didactique, c'est celui sur lequel nous nous appuyerons pour développer la notion de « rapport à l'erreur lexicale des enseignants ». En nous inspirant fortement de la définition de « rapport à l'écrit » fournie par Legendre (2005), nous pourrions donc définir ainsi le rapport à l'erreur lexicale des enseignants :

Rapport à l'erreur lexicale des enseignants : ensemble des valeurs, représentations, attitudes et pratiques propres à un enseignant en relation avec l'erreur lexicale.

Trois des dimensions du rapport à l'écrit, tel que défini dans le cadre de la recherche *Scriptura*, nous semblent particulièrement pertinentes pour la délimitation de la notion de « rapport à l'erreur lexicale des enseignants ». Ce sont en effet les trois axes qu'il serait intéressant d'explorer dans le but d'éclairer le rapport à l'erreur lexicale d'enseignants du secondaire.

La dimension axiologique concerne la valeur que l'enseignant attribue à l'erreur lexicale et, plus largement, à la maîtrise du vocabulaire. Selon lui, est-ce important de relever les erreurs lexicales et de travailler ce type de problèmes en classe? Quelle importance l'enseignant accorde-t-il à la maîtrise du vocabulaire? Croit-il que la maîtrise d'un vocabulaire étendu est importante pour la réussite scolaire et sociale des élèves? De façon plus globale, cette dimension a aussi trait à la façon dont l'enseignant perçoit l'erreur en général; est-ce un indice des connaissances en construction chez l'élève ou plutôt une preuve de l'échec des enseignements reçus jusqu'à maintenant? Il est évident que la dimension axiologique du rapport de l'enseignant à l'erreur lexicale orientera ses attitudes et ses pratiques en lien avec ce type d'erreurs.

La dimension conceptuelle est liée aux représentations de l'enseignant à propos de l'erreur lexicale et de la maîtrise du vocabulaire. Qu'est-ce qu'une erreur de vocabulaire? Pourquoi commet-on de telles erreurs? Qu'est-ce que « connaître un mot »? Les conceptions qui constituent cette dimension du rapport à l'erreur peuvent aussi concerner des aspects plus didactiques de la question. Comment apprend-on le vocabulaire? Les élèves ont-ils un bon vocabulaire? Quels sont leurs principaux problèmes lexicaux? Comment peut-on corriger une erreur lexicale? Comment le dictionnaire peut-il aider à prévenir les erreurs de vocabulaire? Est-ce que cela aide les élèves lorsqu'on corrige leurs erreurs de vocabulaire?

La dimension praxéologique a trait aux pratiques des enseignants relativement à l'erreur lexicale et, plus largement, à l'évaluation du vocabulaire en production de texte. Corrigent-ils les problèmes lexicaux? Si oui, de quelle façon? Font-ils des retours avec leurs élèves sur ce type de problèmes? Comment utilisent-ils eux-mêmes le dictionnaire lorsqu'ils éprouvent des doutes sur un emploi lexical? Comment tiennent-ils compte de la dimension lexicale dans leur évaluation?

Nous n'avons pas retenu la dimension affective présente dans les descriptions du rapport à l'écriture de Barré-de Miniac (2000) et du RÉ de Chartrand et Blaser (2008) parce qu'elle nous paraissait moins pertinente pour un objet aussi précis que l'erreur lexicale. Chartrand, Blaser et Gagnon (2006: 276) mentionnent d'ailleurs que la dimension affective du rapport à l'écrit « se manifeste par deux indicateurs : le temps et la fréquence des pratiques de lecture et d'écriture à l'école et hors de l'école ainsi que les efforts que le sujet consacre à l'écrit pour l'école. » Nous voyions mal comment, d'un point de vue méthodologique, nous pourrions mesurer la dimension affective du rapport à l'erreur lexicale des enseignants sur de telles bases ; certaines informations

relatives à cette dimension pourraient par ailleurs ressortir de l'analyse des dimensions axiologique et praxéologique du rapport à l'erreur lexicale.

CONCLUSION

Ce chapitre nous a donc permis de définir plus clairement les dimensions que nous désirons éclairer lorsque nous affirmons vouloir décrire le rapport à l'erreur lexicale d'enseignants de français du secondaire. Évidemment, les différentes composantes de notre rapport à l'erreur lexicale restent difficilement « désintriquables », mais, tout comme la notion de rapport à l'écrit(ure), le rapport à l'erreur lexicale devrait avoir un potentiel euristique suffisant pour permettre d'intéressantes observations didactiques. Nous verrons dans notre méthodologie de quelle façon nous entendons dégager les représentations/conceptions, les attitudes et les pratiques qui constituent le rapport à l'erreur lexicale d'enseignants du secondaire.

PREMIER OBJECTIF - MÉTHODOLOGIE

VII. MÉTHODOLOGIE – PREMIER OBJECTIF

Avant d'aborder notre méthodologie, rappelons d'entrée de jeu notre premier objectif de recherche.

1. Analyser les erreurs lexicales d'élèves francophones de troisième secondaire en production écrite.

Notre cadre conceptuel nous a permis de définir ce que nous entendons par *erreur lexicale* et de nous familiariser avec les différents aspects à considérer dans l'analyse d'erreurs. Nous présenterons maintenant la méthodologie associée à notre premier objectif. Nous détaillerons notre méthode de collecte de données, avant de décrire notre démarche d'analyse des erreurs. Nous parlerons finalement de l'analyse de données.

1. CONSTITUTION DU CORPUS D'ANALYSE

La première étape de notre analyse d'erreurs a bien entendu été la constitution d'un corpus de textes d'où nous avons tiré les erreurs à analyser. La sélection d'un corpus homogène et représentatif de la population à laquelle nous nous intéressons était bien sûr cruciale pour assurer la validité des résultats tirés de nos observations. Voici donc les caractéristiques du corpus que nous avons constitué et de la population visée par notre étude.

1.1 CONSIDÉRATIONS LIÉES AU CORPUS

Les textes que nous avons analysés pour réaliser notre premier objectif de recherche ont été produits en classe par des élèves de troisième secondaire. Le choix de cibler cette année scolaire en particulier peut être justifié par plusieurs raisons. Tout d'abord, nous avons envisagé récolter des rédactions tout au long du secondaire, mais, l'acquisition du vocabulaire étant encore très rapide à l'adolescence (Beck et McKeown, 1991 ; Stahl et Nagy, 2006), nous croyons que les problèmes lexicaux d'élèves de première secondaire peuvent différer

considérablement de ceux d'élèves de cinquième secondaire. Par ailleurs, nous n'avons pas voulu récolter des données qui auraient demandé une analyse comparative entre différentes populations d'élèves du secondaire, notre corpus n'étant pas assez important (300 rédactions) pour que de telles comparaisons soient pertinentes. Nous avons donc arrêté notre choix sur les élèves rendus à mi-parcours de leurs études secondaires. Ce choix nous semble justifiable dans la mesure où nous voulions, à la lumière de notre analyse, proposer des pistes d'interventions didactiques. De telles interventions sont encore possibles en troisième secondaire, alors qu'elles s'avèreraient moins réalisables en cinquième secondaire, où presque toute l'attention des élèves est portée vers l'épreuve unique de fin d'études secondaires; nous savons par ailleurs que les cours du collégial sont plutôt axés sur la littérature et que les leçons visant la maîtrise du code à proprement parler s'y font rares et constituent donc un terrain moins propice aux interventions en lien avec l'erreur lexicale.

Malgré le fait que le type de texte et les consignes d'écriture puissent selon certaines études influencer la fréquence et la nature des erreurs lexicales commises (Reid, 1990 ; Agustín Llach, 2005b), nous n'avons pas voulu nous imposer des contraintes précises quant au type de textes à analyser. Comme nous désirions tracer un portrait global des difficultés lexicales des élèves en rédaction, nous n'avions pas d'objections à ce que les textes de notre corpus ne soient pas uniformes quant à leur type ou au thème abordé. Par contre, le choix de nous concentrer sur les élèves de troisième secondaire nous a conduit à recueillir des textes de types narratif ou descriptif/explicatif, puisque ce sont les types à l'étude au deuxième cycle du secondaire.

Il nous semblait important que les textes analysés aient été produits en classe afin de pouvoir savoir, par l'intermédiaire de l'enseignant, quels étaient les ouvrages de référence permis lors de la rédaction ainsi que les conditions de production (durée, consignes, type d'évaluation, etc.). Cela avait aussi pour but d'éviter de retenir dans notre corpus des textes qui auraient pu être plagiés sur Internet ou corrigés par des scripteurs plus experts que ceux dont nous voulons observer les

problèmes lexicaux. La longueur des textes devait varier entre 350 et 500 mots, ce qui correspond à la longueur moyenne des textes demandés au deuxième cycle du secondaire.

1.2 CONSIDÉRATIONS LIÉES À LA POPULATION VISÉE

Les élèves à l'origine des textes que nous avons analysés sont des élèves ayant été scolarisés en français depuis la première année du primaire et parlant couramment le français à la maison; ces informations ont été recueillies à l'aide d'un questionnaire distribué aux élèves fournissant leur rédaction (voir Formulaire de consentement – élèves (1^{er} objectif) présent dans l'annexe 2).

Nous avons choisi de limiter notre échantillon à des élèves francophones pour assurer une certaine homogénéité à notre corpus et ne pas produire des données nécessitant une analyse comparative selon les profils linguistiques. En effet, certains travaux ont démontré que les erreurs lexicales produites par des apprenants de L1 diffèrent de celles commises par des apprenants de L2, principalement en ce qui concerne leur fréquence et leur distribution (voir IV-4); un portrait produit en tenant compte des deux populations ne serait donc représentatif d'aucune. Par ailleurs, comme nous nous intéressons aussi à l'explication des erreurs observées, il aurait été hasardeux de mélanger des textes de L1 et de L2 dans le même corpus, puisque les sources peuvent varier considérablement selon le profil linguistique (Ringbom, 1987).

Il nous était difficile d'établir le niveau de français écrit des élèves participant à l'étude, notre seul indice étant le résultat obtenu pour la rédaction que nous avons recueillie. Nous savions par contre l'âge de chaque élève ayant fourni une rédaction, ce qui nous a permis de ne retenir que les productions des élèves âgés de 14 ou 15 ans, qui est l'âge normal en troisième secondaire; ainsi, aucun des élèves de notre étude n'avait vécu de redoublement. Nous avons conservé l'ensemble des copies recueillies qui correspondaient au profil linguistique et à l'âge visés sans tenir compte des notes obtenues par les élèves. Nous croyons que cette façon de procéder nous a permis de dresser un portrait global des

problèmes lexicaux des élèves, ce qui ne serait pas le cas si nous avions écarté, par exemple, les copies s'étant vu attribuer un échec. Un seul groupe, très faible selon les dires de l'enseignant, a été écarté ; la majorité des élèves y avaient obtenu une note avoisinant les 50 % pour la rédaction recueillie.

Nous avons tenté d'obtenir des rédactions provenant d'écoles secondaires publiques situées dans un environnement socioéconomique moyen selon les indices de défavorisation des écoles publiés chaque année par le ministère de l'Éducation. Nous avons à l'origine établi que les écoles retenues devraient se situer entre les rangs déciles 4 et 7 pour les critères ISR (*indice du seuil de faible revenu*) et IMSE (*indice de milieu socioéconomique*). Malheureusement, certaines difficultés liées à la cueillette des données nous ont obligé à assouplir nos contraintes. Les quatre écoles nous ayant fourni des rédactions se situent entre les rangs déciles 5 et 7 pour l'ISR et entre les rangs 2 à 5 pour l'IMSE.

1.3 LA COLLECTE DES DONNÉES ET LES CONSIDÉRATIONS ÉTHIQUES

Un court descriptif de notre projet de thèse a été envoyé à des directions d'écoles secondaires ainsi qu'à des conseillères pédagogiques pour solliciter les enseignants à participer à notre étude. Nous avons ensuite transmis aux enseignants qui ont répondu à notre appel des formulaires de consentement à faire remplir par leurs élèves ainsi qu'un questionnaire à remplir eux-mêmes (le questionnaire, les formulaires de consentement ainsi que le certificat d'éthique y correspondant sont disponibles dans l'annexe II). Afin de ne pas attirer l'attention des élèves et des enseignants sur la dimension lexicale lors des tâches de rédaction et de correction, le descriptif de projet transmis aux écoles et présent sur les formulaires de consentement mentionnait seulement que l'étude s'intéressait aux problèmes de français écrit des élèves, sans spécifier que nous nous concentrons sur ceux relatifs au vocabulaire.

Une fois les formulaires remplis et les productions écrites faites (et dans certains cas corrigées par l'enseignant), nous nous sommes rendu dans les différentes écoles pour emprunter les rédactions et les photocopier. Évidemment, seules les

productions des élèves ayant fourni leur consentement et correspondant au profil linguistique et à l'âge visés ont été photocopiées. Les informations du formulaire (sexe, âge, autres langues parlées) ont été transcrites sur les copies de rédactions correspondantes, mais sans le nom de l'élève, de façon à préserver la confidentialité des participants.

Les caractéristiques du corpus de textes que nous avons analysé seront présentées en détail juste avant la présentation des résultats (VIII-1).

2. PRÉSENTATION DE LA DÉMARCHE D'ANALYSE

Une fois les rédactions recueillies et anonymisées, nous avons pu nous attaquer à l'analyse des productions, étape qui nous a permis de constituer notre corpus d'erreurs. Les étapes de notre démarche d'analyse ainsi que les outils utilisés sont décrits dans les sections suivantes.

Comme nous l'avons vu dans notre cadre conceptuel, l'analyse d'erreurs peut s'intéresser à différentes dimensions des problèmes rencontrés, qu'il importe de bien distinguer si l'analyse se veut rigoureuse. Par conséquent, l'utilisation d'une grille d'analyse d'erreurs multidimensionnelle est préconisée par de nombreux auteurs (James, 1998 ; Granger, 2003). L'étiquetage des erreurs lexicales recueillies dans notre corpus de productions d'élèves a donc distingué trois niveaux d'analyse. Le premier fournit une description linguistique structurale des erreurs, le second permet une évaluation de la gravité des problèmes alors que le troisième s'intéresse aux sources possibles de l'erreur; les outils d'analyse correspondant à ces trois niveaux seront décrits plus bas.

La démarche d'analyse d'erreurs que nous préconisons respecte les grandes étapes propres aux travaux des différents courants de recherches abordés dans notre cadre conceptuel (voir III-3.1). Voici, en détail, les étapes de notre démarche d'analyse des erreurs lexicales.

1. Le repérage des erreurs : Cette étape consistait à relever dans chaque rédaction tous les problèmes de nature lexicale. Chaque texte était lu en entier une première fois avant l'identification des erreurs, dans le but de bien comprendre le sens général du texte pour être en mesure de relever

efficacement les erreurs de nature sémantique, dont le domaine dépasse régulièrement la phrase et nécessite de faire des liens entre les phrases, voire entre les paragraphes. C'est à partir de nos intuitions de locuteur expert que nous avons fait le repérage, bien sûr en nous appuyant sur les ouvrages de référence chaque fois que nous avons des doutes (voir 2.1).

2. Saisie des problèmes relevés dans le logiciel de codage : Nous avons choisi d'utiliser le logiciel *Excel* de la suite *Office pour Mac* pour l'analyse des données de notre premier objectif de recherche, ce logiciel permettant une saisie facile des données et présentant une grande flexibilité en ce qui a trait à leur tri, tout en proposant des outils d'analyse intéressants, notamment les tableaux croisés dynamiques.

Chacun des problèmes recueillis dans la production a donc été saisi dans le logiciel de codage des erreurs lexicales pour pouvoir être décrit en détail. L'item linguistique problématique était d'abord copié dans le logiciel avec un contexte suffisant pour permettre de faire ressortir son statut erroné⁹⁶. Dans les cas où le contexte à considérer était trop large (erreur lexicale relevant plutôt de la cohérence textuelle) ou pour les problèmes se justifiant par des arguments extratextuels (cf. connaissances sur le monde), nous avons fourni une explication de sorte que l'erreur soit compréhensible sans qu'il faille retourner à la rédaction. Pour chaque erreur, nous avons aussi indiqué le code permettant d'identifier la rédaction d'où elle provenait, le numéro attribué à l'élève l'ayant produite (de 1 à 300) et la longueur du texte d'où elle est extraite (en nombre de mots).

3. Correction de l'erreur : Cette étape consistait à proposer une solution pour rendre correcte la phrase dans laquelle apparaît l'erreur. En d'autres mots, la correction a pour but de rétablir la forme correcte dans le cas d'erreurs formelles ou morphosyntaxiques et de restituer le sens visé par le scripteur dans le cas d'erreurs sémantiques. Comme les corrections ont été proposées à partir d'inférences de l'analyste et sans interroger le scripteur, elles ne constituent, comme nous l'avons vu dans le cadre conceptuel, que des reconstructions plausibles dont le caractère hypothétique augmente sans doute dans le cas de problèmes impliquant une restitution du sens; dans quelques cas, il a même été impossible de proposer une correction (voir point *iii* plus bas). Lors de la correction, les principes suivants ont été respectés :

- a. Les corrections proposées doivent conserver intact le texte du scripteur le plus possible, c'est-à-dire impliquer le minimum d'interventions pour rendre l'énoncé sémantiquement, formellement, syntaxiquement et pragmatiquement correct. C'est ce que Milićević et Hamel (2007) appellent une *correction minimale*. C'est l'adéquation sémantique qui

⁹⁶ Nous avons à l'origine voulu faire correspondre ce contexte au contexte minimal nécessaire pour identifier l'erreur (*domaine de l'erreur*, voir IV-3.2.3), mais l'opérationnalisation de ce concept s'est avérée plus difficile que prévu et nous avons pris la décision de déterminer de façon intuitive le contexte nécessaire pour mettre en évidence chaque erreur.

aura préséance sur la correction formelle ou syntaxique; ainsi, il sera important de s'assurer que le terme erroné est approprié sémantiquement avant de proposer une correction relevant des autres dimensions.⁹⁷

- b. Dans le cas où plus d'une correction est possible, toujours se demander laquelle est la plus probable. S'il est impossible de trancher, fournir plus d'une correction et ajouter une note à l'erreur pour y revenir en fin de codage et prendre une décision.

À la suite de la correction, un premier code a été attribué à l'erreur quant à notre degré de certitude relativement à la correction proposée. Voici les trois niveaux de certitude :

- i. *Élevée* : indique que l'analyste est à peu près convaincu que la correction proposée est fidèle au message que voulait livrer le scripteur.
- ii. *Modérée* : indique que l'analyste n'est pas certain qu'il a réussi, par sa correction, à rétablir le message que voulait livrer le scripteur.
- iii. *Aucune correction proposée* : pour les cas recevant ce code, aucune correction n'a été proposée, puisque le contexte ne permettait pas à l'analyste de rétablir le sens visé.

4. Confrontation du repérage de l'analyste avec la correction de l'enseignant : Le fait d'avoir dans notre corpus un nombre important de rédactions déjà corrigées par les enseignants (224/300) s'est révélé intéressant, puisque cela a en quelque sorte servi de second repérage d'erreurs par un locuteur expert et, dans certains cas, nous a permis d'obtenir une seconde proposition de correction par le biais des annotations de l'enseignant.

Pour tenir compte de la position des enseignants par rapport aux problèmes lexicaux, un deuxième code a été attribué à chaque erreur pour indiquer si elle a été relevée ou non par l'enseignant, ce qui nous a permis de vérifier quels sont les types de problèmes généralement corrigés et de les mettre en relation avec le degré de gravité des erreurs (voir X-1.2). Dans le cas où l'erreur était relevée par l'enseignant, nous avons indiqué si elle a été pénalisée, c'est-à-dire comptabilisée dans la note finale. Lorsqu'une erreur a fait l'objet d'un commentaire de la part de l'enseignant (proposition de correction, code d'erreurs, commentaire développé, etc.), ce dernier a été consigné dans un espace prévu à cet effet dans notre grille de codage. Ces commentaires ne nous ont pas été utiles directement pour répondre à notre premier objectif de recherche, mais nous ont aidé à alimenter les questions

⁹⁷ Il est à noter que cette façon de procéder et celle que nous avons privilégiée au moment de l'analyse. Dans le cas d'interventions didactiques visant des problèmes ayant comme source une confusion entre des synonymes, Martin (1984: 132) suggère plutôt de proposer aux étudiants une correction préservant le contexte grammatical dans lequel apparaît l'erreur, ceci dans le but de bien mettre en évidence que l'énoncé qu'ils ont produit aurait été acceptable avec un terme sémantiquement proche, mais que l'item qu'ils ont spécifiquement choisi est incompatible avec le contexte grammatical.

posées aux enseignants dans le cadre de notre deuxième objectif de recherche (voir chapitre IX) et à faire certaines observations relatives au traitement de l'erreur lexicale par les enseignants.

5. Description de l'erreur : Cette étape correspondait à la description linguistique de l'erreur. Décrire l'erreur consiste à caractériser la différence entre ce qui a été produit (l'erreur) et ce qui aurait dû l'être (la correction proposée). Voici les critères selon lesquels les problèmes ont été codés :

a. **Classe de mots :** Cette étiquette spécifie la classe de mots de l'UL cible, parmi les étiquettes suivantes : *nom, adjectif, verbe, adverbe, préposition, conjonction (subordonnant/coordonnant), pronom, déterminant*. Il existe quand même certaines exceptions à cette règle. Dans le cas d'une erreur de préposition régie, la classe de mots associée à l'erreur n'est pas celle de la préposition, mais celle de l'UL qui la contrôle. Dans le cas d'erreurs de collocation, la classe de mots est celle de la base de la collocation, puisque l'erreur découle d'une méconnaissance des propriétés collocationnelles de cette UL. Dans le cas de pléonasmes, nous avons considéré la classe de mots du terme du pléonisme dont le sens englobe celui de l'autre terme (par exemple CHUCHOTER dans **chuchoter faiblement*, donc *verbe*).

b. **Type d'UL en jeu :** Cette rubrique nous indique si l'UL cible est un lexème ou une locution, de façon à voir si certains types de problèmes sont plus susceptibles de concerner un type particulier d'UL.

c. **Description structurale de l'erreur :** À l'aide de la typologie descriptive des problèmes lexicaux (voir 3.1), chaque erreur a ensuite reçu un code pour indiquer la classe à laquelle elle appartient. Tous les cas de double codage (voir 3.1.2) ont été clairement identifiés de sorte à pouvoir être revérifiés à la fin du codage.

d. **Gravité de l'erreur :** À la suite de l'analyse structurale de l'erreur, nous avons posé sur chaque erreur un jugement de gravité. Toute erreur s'est donc vu attribuer un score de 1 à 6 à l'aide du Tableau 13 présenté au point 3.2.

6. Diagnostic de l'erreur : Cette étape consistait à fournir une explication sur la source de l'erreur. Ce diagnostic a été fait à l'aide d'une liste de sources d'erreurs (voir 3.3), qui a été étoffée en cours d'analyse. À côté du diagnostic proposé, un espace nous permettait de fournir une courte explication (par exemple pour mentionner l'autre UL en jeu dans le cas d'une analogie formelle ou sémantique). Dans un cas où il y a des diagnostics concurrents, nous les avons mentionnés tous. Lorsque des sources différentes semblent cumulatives (cf. *compound error*), nous l'avons aussi mentionné. Cette étape de l'analyse est bien entendu la plus subjective et les diagnostics posés ne constituent que des hypothèses. Pour cette raison, nous n'avons tenté d'établir la source de l'erreur pour tous les

problèmes; si aucun diagnostic ne nous venait spontanément à l'esprit, nous avons tout simplement laissé cet espace vide.

Le Tableau 10 résume les informations fournies pour chaque problème lexical relevé. Un extrait de notre grille de codage remplie est aussi disponible à l'annexe 3 pour permettre de visualiser de quelle façon l'information y figure.

Tableau 10. Informations fournies lors de l'étiquetage des erreurs

PROBLÈME ET CORRECTION	<i>Erreur et contexte</i>	
	<i>Correction de l'erreur</i>	
	<i>Degré de certitude de la correction</i>	Élevé Modéré Aucune correction proposée
POSITION DE L'ENSEIGNANT	<i>Problème relevé</i>	Oui / Non
	<i>Problème pénalisé</i>	Oui / Non
	<i>Commentaire</i>	
DESCRIPTION DE L'ERREUR	<i>Classe syntaxique</i>	Nom Verbe Adjectif Adverbe Préposition Conjonction Pronom Déterminant
	<i>Type d'unité lexicale</i>	Lexème Locution
	<i>Description structurale de l'erreur</i>	Type d'erreur (codes de la typologie descriptive)
		Commentaires sur le classement
<i>Gravité (score de 1 à 6)</i>	À l'aide de la grille d'évaluation de la gravité	
DIAGNOSTIC	<i>Source de l'erreur</i>	Codes de la typologie explicative
	<i>Commentaire</i>	

2.1 OPÉRATIONNALISATION DE LA NOTION DE LOCUTEUR EXPERT ET CHOIX DES OUVRAGES DE RÉFÉRENCE

Nous avons choisi de construire notre définition de l'erreur autour du concept de locuteur expert (voir IV-1.2), locuteur dont nous rappelons ci-dessous les caractéristiques.

- Il s'agit d'un locuteur natif.
- Il a un bon niveau d'éducation et une très bonne maîtrise de la langue, c'est-à-dire qu'il maîtrise les règles de grammaire (norme prescriptive) et qu'il possède un vocabulaire précis et étendu.

- Il s'agit d'un locuteur flexible qui sait bien s'adapter à différents contextes d'énonciation (norme fonctionnelle), notamment par l'emploi de différents registres de langue. Il se permet aussi de transgresser parfois certaines règles pour produire des effets de style (humour, jeux de mots, figures de style, etc.).
- Il s'agit d'un locuteur idéalisé, c'est-à-dire qu'il n'est jamais confronté à des problèmes d'accès au vocabulaire et qu'il ne commet pas d'erreurs de performance.

Recentrer notre définition de l'erreur autour de la notion de locuteur expert nous permet de contourner en quelque sorte la notion de norme, que nous jugeons trop contraignante pour l'erreur linguistique en général, et plus spécifiquement pour l'erreur lexicale. Bien évidemment, il nous est impossible d'esquiver complètement la norme et il est important d'expliquer ici de quelle façon nous l'avons prise en compte dans notre analyse des erreurs lexicales.

Nous avons vu que le dictionnaire demeure la meilleure illustration de la norme (Guilbert, 1972), notamment pour l'aspect lexical. Comme nous avons défini notre locuteur expert comme ayant un vocabulaire précis et entendu (et qui n'est par ailleurs jamais confronté à des problèmes d'accès au vocabulaire), nous devons nécessairement nous servir du dictionnaire comme balise dans notre analyse des erreurs lexicales. S'est donc posée la question du choix des ouvrages de référence à retenir. Les dictionnaires étant des outils descriptifs nécessairement incomplets (voir III-3), nous avons jugé préférable d'en retenir plus d'un aux fins de notre travail, d'autant plus vrai qu'il fallait s'assurer que les ouvrages retenus rendent compte de la réalité linguistique québécoise. Voici donc la liste des dictionnaires que nous avons choisis pour notre étude.

1. ***Le Petit Robert*** : Ce dictionnaire est sans aucun doute l'ouvrage de référence par excellence dans toute la francophonie. *Le Petit Robert* est particulièrement complet en ce qui concerne la polysémie et les locutions et s'avère très riche pour les liens de synonymie/antonymie; les collocations les plus fréquentes y apparaissent aussi dans les exemples. Cet ouvrage à forte tendance descriptive étant publié en France, il rend surtout compte de la norme française hexagonale, malgré la présence de certains usages propres aux autres français européens, puis de certains usages africains et « canadiens ».

2. **Le Multidictionnaire** : Probablement le dictionnaire québécois le plus utilisé, *Le Multidictionnaire* n'a pas pour seul objectif la description lexicographique, mais vise à « fournir une réponse sûre à toutes les questions linguistiques qui peuvent se poser » (de Villers, 2003: XI). Cet ouvrage contient donc, en plus des définitions, des tableaux sur différentes notions grammaticales, des mises en garde relatives aux difficultés les plus fréquentes rencontrées par les scripteurs, des notes sur la norme typographique, etc. Deux grands principes guident le choix de la nomenclature et de la rédaction des articles : « l'inventaire des difficultés prévisibles, d'une part, le respect de la norme sociale la plus généralement admise, d'autre part » (p. XI). Ce premier principe a pour corolaire une nomenclature moins grande que celle du *Petit Robert* et une description lexicographique moins complète quant à la polysémie des vocables; le nombre d'exemples est aussi plus limité que dans *Le Petit Robert*. Le deuxième principe fait que *Le Multidictionnaire* est un ouvrage à forte tendance prescriptive, puisqu'il rend compte des prescriptions normatives émises par les différentes instances linguistiques, particulièrement celles de l'Office québécois de la langue française. L'un des intérêts principaux du *Multidictionnaire* pour notre travail tient dans le fait que cet ouvrage met bien en lumière les particularités du français québécois, particulièrement du français québécois standard. Il serait donc pratiquement impossible d'envisager un travail de description d'erreurs d'élèves en contexte scolaire québécois sans inclure parmi nos ouvrages *Le Multidictionnaire*.

3. **Antidote RX** : Cet outil informatique révolutionnaire, conçu au Québec par la firme Druide informatique, est beaucoup plus qu'un dictionnaire : il contient aussi plusieurs outils d'aide à la rédaction, dont des guides grammaticaux, un correcteur extrêmement puissant ainsi que des dizaines de filtres qui permettent d'analyser les textes sous différents angles. Toutefois, c'est à la dimension lexicographique d'*Antidote RX* que nous ferons appel dans le cadre de cette thèse. Cet outil referme en effet une dizaine de dictionnaires qui fournissent une description lexicographique rarement égalée. Outre le classique dictionnaire de définitions – très précis quant à la polysémie des vocables –, *Antidote* dispose aussi de dictionnaires des synonymes et des antonymes, d'un dictionnaire d'expressions et d'un dictionnaire des cooccurrences extrêmement étoffé⁹⁸ qui s'est avéré fort utile pour la description de problèmes de collocation. L'outil fournit aussi un dictionnaire de citations, qui donne accès à plusieurs exemples d'utilisation de la plupart des UL décrites, exemples tirés de romans et d'essais, mais aussi de journaux. Beaucoup moins normatif que *Le*

⁹⁸ C'est d'ailleurs l'ajout de ce dictionnaire des cooccurrences à *Antidote RX* qui fait que nous ne croyons pas pertinent de retenir parmi nos ouvrages de référence le *Dictionnaire des cooccurrences* de Jacques Beauchesne (2002), qui fut pourtant un pionnier du genre. En effet, la nomenclature d'*Antidote RX* est beaucoup plus vaste et le nombre de cooccurrences y est, de façon générale, plus élevé pour chaque entrée. Des informations sur la fréquence relative des différentes cooccurrences y sont aussi indiquées.

Multidictionnaire, Antidote RX décrit aussi les UL propres au français québécois, mais s'avère plus exhaustif en ce sens, parce que les emplois familiers y sont beaucoup plus présents et que les expressions y sont décrites systématiquement.

Avec ces trois ouvrages de référence, nous disposons d'un portrait satisfaisant de la norme lexicale du français québécois, du moins dans un contexte de production écrite en milieu scolaire. Nous avons considéré comme correct tout emploi lexical apparaissant dans l'un ou l'autre de ces dictionnaires, pourvu bien sûr que cet emploi soit approprié dans le contexte dans lequel il était utilisé (cf. norme fonctionnelle). De la même façon, tout emploi condamné explicitement par l'ensemble des ouvrages a été considéré comme une erreur. Dans le cas où les ouvrages avaient des positions différentes relativement à certains usages, nous avons adopté celle du dictionnaire le plus permissif⁹⁹.

Dans le cas où l'emploi lexical qui retient notre attention ne figurait pas dans nos ouvrages de référence – surtout dans le cas de locutions, de collocations ou d'« extensions sémantiques » –, nous nous sommes fié à notre intuition; c'est dans ces cas précis qu'est intervenue la notion de locuteur expert. Bien entendu, nous avons validé nos intuitions à l'aide d'Internet, notamment par des observations sur *Google*, qui nous donne accès à un corpus d'une ampleur impressionnante. Énormément de travaux de description lexicographique se font à l'aide de corpus (Stubbs, 2001); le corpus linguistique offert par Internet, bien qu'il échappe à toute sélection quant aux contenus qui le constituent, présente tout de même l'avantage de donner un portrait des usages d'une communauté linguistique. Bien entendu, la personne qui consulte un tel moteur de recherche dans un but de validation linguistique doit le faire avec une certaine réserve. Une dizaine d'occurrences d'un emploi lexical douteux ne témoigne pas d'une intégration au système lexical de la langue et ne signifie en rien que cet emploi doit être considéré comme correct. Par contre, plusieurs centaines d'occurrences d'une UL employée dans un sens non attesté dans les ouvrages ou régissant une structure

⁹⁹ Si un usage n'apparaissait que dans l'un des ouvrages et y était explicitement condamné, il a été considéré comme une erreur.

syntactique différente de celle figurant dans les dictionnaires, dans la mesure où ces occurrences proviennent de sites qui semblent dignes de considération, peuvent en quelque sorte légitimer ces emplois ; elles témoignent en effet d'une intégration dans l'usage.

Le tableau ci-dessous récapitule la façon par laquelle les ouvrages de référence ont servi d'assises à notre concept de locuteur expert en nous permettant de statuer sur le caractère erroné ou non de certains emplois lexicaux ayant retenu notre attention lors de l'étape du repérage. Nous sommes tout à fait conscient que les cas où nous avons exercé notre jugement de locuteur donnent place à plus de subjectivité dans notre description des problèmes, mais cette subjectivité nous semble moins importante que celle d'une approche suivant laquelle nous aurions considéré comme des erreurs tous les emplois qui ne figurent pas dans les dictionnaires. Par ailleurs, il ne faut pas oublier que la subjectivité est inhérente à toute entreprise d'analyse d'erreurs et qu'elle se manifeste dès l'étape de repérage. En effet, les items lexicaux qui ont retenu notre attention et nous ont poussé à consulter les ouvrages de référence ont été relevés parce qu'ils ne correspondaient pas à des emplois normaux selon notre propre conception de la langue; ils semblaient donc problématiques à travers la lunette de notre propre subjectivité. C'est la raison pour laquelle nous jugeons important qu'une personne qui s'engage dans l'analyse de tels problèmes soit un locuteur expert, dont la connaissance de la langue est assez approfondie pour qu'il puisse avoir de fortes intuitions sur la norme réelle.

Tableau 11. Utilisation des ouvrages pour statuer sur le caractère erroné d'un emploi lexical

CAS	EMPLOI CORRECT	ERREUR	JUGEMENT DU LOCUTEUR EXPERT
Usage attesté par au moins un ouvrage (même si condamné explicitement par les autres)	✓		
Usage condamné explicitement par les trois ouvrages		✓	
Usage condamné explicitement par un ouvrage et absent des autres		✓	
Usage absent des ouvrages de référence			✓

3. PRÉSENTATION DES OUTILS D'ANALYSE

Les trois niveaux de notre démarche d'analyse correspondent à autant d'outils qui permettront l'étiquetage de nos erreurs. Nous les présenterons tour à tour, en commençant par le plus complexe : notre typologie descriptive des erreurs lexicales.

3.1 LA TYPOLOGIE DESCRIPTIVE DES ERREURS LEXICALES

Les différents travaux sur l'erreur lexicale analysés dans notre cadre conceptuel nous permettent de proposer ici une typologie descriptive que nous croyons complète et susceptible de dépasser certaines limites inhérentes aux autres classements.

Certaines des typologies d'erreurs lexicales consultées dans notre revue de littérature ont été construites de façon empirique, par l'observation des erreurs lexicales tirées d'un corpus (voir par exemple Zughoul, 1991 ; Patris et Vansnick, 1992). Cette façon de faire présente l'avantage de déboucher sur une typologie d'erreurs étroitement liée aux problèmes relevés dans le corpus, ce qui peut paraître tout à fait justifié dans une perspective didactique. Pourquoi, alors, proposons-nous plutôt d'utiliser une typologie structurale dont les classes

d'erreurs sont déterminées à l'avance? Plusieurs raisons motivent ce choix méthodologique.

Tout d'abord, nous avons déjà travaillé sur l'erreur lexicale (cf. Anctil, 2005) et la typologie proposée alors s'est révélée assez opérationnelle pour que nous jugions approprié de fonder notre analyse de thèse sur le même type de grille, en y apportant bien sûr des améliorations. Par ailleurs, la confrontation de notre typologie avec celles utilisées par d'autres auteurs a fait ressortir assez de points communs – tant relativement aux classes d'erreurs y figurant qu'à la structure de la grille – pour nous convaincre de la pertinence de notre classement.

Ensuite, et plus important encore, notre conception de l'erreur lexicale repose sur un cadre théorique précis et exhaustif qui, selon nous, permet de prévoir tous les types de problèmes lexicaux possibles. La grille que nous proposons ici est donc susceptible d'être plus complète qu'une grille qui ne contiendrait que les types d'erreurs effectivement observés dans un corpus. En effet, il se peut tout à fait que le corpus analysé ne contienne pas certains types d'erreurs, ce qui ferait qu'une grille construite de façon empirique pourrait se révéler incomplète et donc difficilement réutilisable avec un autre corpus. Comme nous souhaitons que notre typologie ait un potentiel didactique élevé, donc qu'elle puisse servir à décrire les problèmes lexicaux d'élèves de différents horizons (L1, L2, primaire, secondaire, etc.) et dans différentes situations d'écriture, il est important qu'elle couvre l'ensemble des cas possibles.

Bien entendu, l'information que nous aurait fournie une grille construite empiriquement, à savoir les types d'erreurs réellement présents dans les rédactions d'élèves, sera aussi accessible lorsque nous analyserons quantitativement les problèmes lexicaux codés à l'aide de notre typologie descriptive.

3.1.1 SURVOL DES CLASSES D'ERREURS DE LA TYPOLOGIE DESCRIPTIVE

Voici la typologie descriptive que nous avons utilisée pour décrire, d'un point de vue structural, les erreurs relevées dans notre corpus. Une description détaillée de chacune des classes suit la grille.

Tableau 12. Typologie descriptive des erreurs lexicales

		CLASSES D'ERREURS		CODES
FORME			Erreur d'orthographe lexicale	F-Orth
			Barbarisme (signifiant déformé)	F-Barb
			Erreur de flexion	F-Flex
			Forme analytique	F-Déf
			UL inexistante	F-Inex
SENS			Utilisation inappropriée d'un quasi-synonyme	S-Syn
			Utilisation inappropriée d'une UL de sens proche	S-Proc
	Improprété	Sens fictif		S-Fict
		Anglicisme sémantique		S-Ang
			Non-respect des contraintes sémantiques pesant sur les actants d'une UL	S-Act
		Pléonasme	S-Pléo	
COMBINATOIRE RESTREINTE	GRAMMATICALE	SYNTAXIQUE (RÉGIME)	Absence d'un complément obligatoire de l'UL	CS-AbsC
			Emploi transitif indirect d'un verbe transitif direct	CS-Trans
			Utilisation d'une préposition/conjonction erronée	CS-ChPr
			Préposition régie absente	CS-AbsPr
			Emploi transitif d'un verbe intransitif	CS-Intrans
			Construction erronée dans un complément	CS-Const
		MORPHO-SYNTAXIQUE	Non-respect de la classe de mots de l'UL	CM-Class
			Non-respect du genre du nom	CM-Genr
			Non-respect du caractère invariable de l'UL	CM-Inv
			Utilisation d'un verbe auxiliaire erroné	CM-Aux
			Classe résiduelle	CM-Rés
		LEXICALE	Choix d'un collocatif inapproprié (collocation inexistante)	
	Choix d'un collocatif valable, mais mauvaise fonction lexicale		CL-FL	
	Non-respect des contraintes d'emploi du collocatif		CL-Cont	
	Erreur de dérivation sémantique		CL-Dér	
	PRAGMATIQUE	Emploi d'une UL appartenant à un registre de langue inapproprié		CP-Regis
		Emploi d'une UL véhiculant un contenu connotationnel incompatible avec le contexte		CP-Conn
		Emploi d'une UL appartenant à une variété dialectale incompatible avec le contexte		CP-Dial

Cette grille contient l'ensemble des erreurs lexicales susceptibles d'être rencontrées dans une production écrite. Voici maintenant un survol des différentes classes d'erreurs qui y figurent.

3.1.1.1 LES ERREURS FORMELLES

Nous considérons qu'il y a une erreur relative au signifiant d'une UL chaque fois que le scripteur ne semble pas connaître ou maîtriser la forme qui véhicule le sens qu'il tente d'exprimer. Il existe cinq types d'erreurs de forme.

- **Erreur d'orthographe lexicale** : Il s'agit de la traditionnelle « faute d'orthographe ». Dans ce cas-ci, l'orthographe utilisée ne correspond pas à la norme, mais l'erreur ne se manifeste qu'à l'écrit et n'altère pas la prononciation du mot (**encors* pour *encore*). Nous avons remarqué (Anctil, 2005) que les problèmes d'orthographe, du moins à l'université, concernent souvent l'emploi du trait d'union (**bien être*) et de la majuscule (**la marine Allemande*) ainsi que la division des mots-formes (*assis *au tour du feu*). Les erreurs d'orthographe lexicale peuvent bien sûr affecter les locutions (**entre autre*).
- **Signifiant déformé** : Dans ce cas, l'orthographe erronée altère la prononciation du mot, ce qui laisse supposer que la personne ayant produit l'erreur a une méconnaissance plus profonde de la forme associée au sens visé (**assuidité* ou **rénumérer*). Dans le cas de lexèmes, nous parlerons de *barbarisme*, alors que nous nommerons *locution déformée* une expression dont la forme aura été altérée (*c'est la goutte qui fait déborder le *verre*). Nous incluons aussi dans cette classe les cas de confusion entre des paronymes (*invasion* pour 'évasion') et les cas de verbes essentiellement pronominaux employés de façon non pronominale.
- **Erreur de flexion** : Cette classe regroupe tous les problèmes de conjugaison de verbes irréguliers (**vous disez, *ils sontaient*) ainsi que les formes fléchies de noms ou d'adjectifs constituant des exceptions dans leur variation en genre ou en nombre (**une grandeur standarde, *des yeux bleux*). Il est important de bien distinguer les problèmes de cette classe des problèmes relatifs au caractère invariable de certaines UL; ces problèmes ont plutôt été classés parmi les transgressions de propriétés de combinatoire grammaticale.
- **« Forme analytique »** (Milićević et Hamel, 2007) : Une forme analytique correspond en quelque sorte à une décomposition sémantique se rapprochant d'une définition plus ou moins complète. Cette formulation maladroite et peu idiomatique est remplacée dans la correction par une seule UL; c'est ce qui distingue les problèmes de « forme analytique » de la « périphrase », qui serait plutôt classée parmi les problèmes de style. L'utilisation d'une forme analytique peut résulter de la méconnaissance complète de la forme correspondant au sens visé (l'élève ne sait pas qu'il existe un mot pour exprimer ce sens), d'un problème d'accès à la forme au moment de l'écriture ou de la volonté d'éviter la répétition d'un mot déjà utilisé. Dans les deux derniers cas, le scripteur produit délibérément la

forme analytique en ayant recours à une stratégie d'écriture : la paraphrase.

- **UL inexistante** : Il s'agit du cas extrême d'un problème formel : la forme produite est inexistante ou du moins, l'association entre la forme utilisée et le sens visé est inexistante¹⁰⁰. Dans certains cas, cette forme pourra avoir été produite par une mauvaise dérivation morphologique; certaines dérivations conduiront à des formes inexistantes (**impersévérant*) alors que d'autres correspondront, de façon fortuite, à des UL existantes (**ennuyeux* au sens de 'qui s'ennuie facilement'). Pour des raisons de cohérence dans notre analyse, nous avons classé les erreurs relevant du second cas de figure (forme correspondant à UL existante) parmi les erreurs sémantiques (cf. *impropriété*, voir plus bas); l'étape d'explication de l'erreur nous a par contre permis de faire mention de la dérivation morphologique comme source possible. Finalement, il pourrait aussi arriver que le signifiant utilisé ne présente aucune parenté formelle permettant de croire qu'il découle de la déformation d'une UL existante. Par contre, ce dernier cas de figure est peu susceptible d'apparaître dans un corpus de langue première, mais est fréquent en L2, où le recours à des formes de la langue première (donc souvent inexistantes dans la langue cible) constitue une stratégie pour pallier un manque lexical.

3.1.1.2 LES ERREURS SÉMANTIQUES

Comme nous l'avons vu dans le cadre conceptuel, l'approche holistique dans la description des erreurs sémantiques permet de rendre compte plus aisément de ce type de problèmes qu'une approche par composantes sémantiques. C'est la raison pour laquelle nous avons adopté une méthode de description fortement inspirée de Milićević et Hamel (2007), c'est-à-dire qui vise à caractériser la « distance sémantique » entre l'UL utilisée et le sens visé. Voici donc les types d'erreurs sémantiques que nous avons retenus.

- **Utilisation inappropriée d'un quasi-synonyme** : Un lien de (quasi)synonymie existe entre l'UL erronée et la lexie visée. Pour que deux termes soient considérés comme des (quasi-)synonymes, il faut qu'ils soient substituables dans certains contextes sans altérer le sens de la phrase de façon importante.
- **Utilisation inappropriée d'une UL de sens proche** : Des UL de sens proche partagent des composantes de sens, peuvent appartenir à un même champ sémantique, mais ne sont jamais substituables en contexte.

¹⁰⁰ C'est ce que nous appelions une *lexie fictive* dans le cadre de notre maîtrise. (Anctil, 2005)

- **Impropiété** : Ce cas correspond au « sens fictif » de Milićević et Hamel (2007) : l'association entre la forme utilisée et le sens visé est inexistante dans la langue. Comme nous l'avons souligné plus haut, les cas de mauvaise dérivation morphologique menant à une forme existante ont été placés dans cette classe.
- **Anglicisme sémantique** : L'anglicisme sémantique constitue un cas particulier d'impropiété. Ici, le scripteur a recours à une forme appartenant à la langue anglaise (ou une version francisée d'une UL de l'anglais) pour exprimer un sens qui n'est pas associé à cette forme en français. Comme nous l'avons mentionné dans le cadre conceptuel (V-3.2.2), nous ne croyons pas que le fait de qualifier un problème sémantique d'anglicisme relève de l'explication, puisque la plupart des anglicismes sémantiques sont tout à fait intégrés au système linguistique et sont utilisés par les locuteurs comme des UL de leur langue, c'est-à-dire sans passer par l'anglais.
- **Non-respect de contraintes sémantiques pesant sur les actants** : Il s'agit de la seule classe d'erreurs sémantiques qui repose sur une analyse plus componentielle du sens des UL en jeu. Une erreur appartient à cette classe lorsqu'il y a incompatibilité entre l'UL utilisée et un de ses actants sémantiques présent dans l'énoncé.
- **Pléonasmе** : Un pléonasmе résulte de la cooccurrence de deux termes dont le sens de l'un est inclus dans celui de l'autre. Il s'agit en quelque sorte d'une répétition sémantique.

3.1.1.3 LES ERREURS RELEVANT DE LA COMBINATOIRE RESTREINTE

Nous avons vu dans la section sur le lexique que la combinatoire restreinte d'une UL peut regrouper une grande variété de propriétés, que nous divisons ici en trois : la combinatoire grammaticale, la combinatoire lexicale et la combinatoire pragmatique.

LA COMBINATOIRE GRAMMATICALE regroupe les propriétés liées au comportement morphosyntaxique de l'UL. Voici d'abord la liste des erreurs liées au régime, puis de celles relatives aux autres propriétés morphosyntaxiques.

RÉGIME (structure syntaxique contrôlée) : Les problèmes de régime concernent la structure syntaxique régie par une UL; il en existe différents types.

- **Absence d'un complément obligatoire de l'UL** : Certains verbes, noms, adjectifs ou adverbes exigent la présence d'un complément. L'absence de ce dernier conduit à une erreur.

- **Emploi transitif indirect d'un verbe transitif direct** : Dans ce cas, le complément direct d'un verbe est introduit par une préposition superflue (**se rappeler de quelque chose*).
- **Utilisation d'une préposition/conjonction erronée** : Ici, la structure du complément (GPrép) est appropriée, mais c'est le choix d'une mauvaise préposition qui entraîne l'erreur (**avoir de l'intérêt envers qqch*). Entrent aussi dans cette classe les cas où une conjonction régie est remplacée par une préposition (*Son second séjour à Londres fut une expérience meilleure *de ce qu'il avait vécu la première fois.*).
- **Préposition régie absente** : À l'inverse, le complément indirect d'un verbe est ici introduit sans préposition, comme s'il s'agissait d'un verbe transitif direct (**échouer un examen*). Cette erreur pourrait concerner une UL appartenant à une autre classe de mots que les verbes.
- **Emploi transitif d'un verbe intransitif** : Dans ce cas, un verbe intransitif est utilisé comme s'il contrôlait un complément (**débuter son travail*).
- **Construction erronée dans les compléments** : Les constructions qui peuvent servir de complément à certaines UL sont limitées et il est important de respecter ces contraintes pour ne pas commettre d'erreur (**Le plaisir que j'aie réussi mon examen m'a grandement encouragé.*)¹⁰¹.

AUTRES PROPRIÉTÉS MORPHOSYNTAXIQUES DE L'UL

- **Non-respect de la classe de mots de l'UL** : Une erreur de ce type survient lorsqu'une UL est utilisée pour remplir une fonction syntaxique incompatible avec la classe de mots à laquelle elle appartient (**un logo triangle*).
- **Non-respect du genre du nom** : C'est par les accords dans le groupe du nom que nous avons détecté ce genre d'erreurs; il se pourrait donc que certaines soient demeurées indétectables, si le nom dont le scripteur ignore le genre apparaissait dans un contexte qui ne demandait pas d'accord dans le GN.
- **Non-respect du caractère invariable de l'UL** (ex. MATHÉMATIQUES_N, ARMOIRIES, CULTURE au sens d'un 'ensemble des connaissances acquises dans un ou plusieurs domaines par un individu') : Nous incluons aussi dans cette catégorie les formes non existantes des verbes défectifs.
- **Utilisation d'un verbe auxiliaire erroné** : seulement dans les cas où le choix de l'auxiliaire ne peut être prédit par une règle, par exemple dans le cas de verbes dont l'auxiliaire varie en fonction du sens.
- **Non-respect d'une autre propriété grammaticale de l'UL** : comme il existe une infinité de propriétés liées à la combinatoire grammaticale, nous avons cru pertinent d'ajouter une classe résiduelle, pour les erreurs qui ne correspondent pas aux cas présentés ci-dessus.

¹⁰¹ Exemple tiré de Boivin, Pinsonneault et Philippe (2003).

COMBINATOIRE LEXICALE : Les résultats de notre mémoire de maîtrise (Anctil, 2005) nous avaient conduit à proposer différentes sous-classes parmi les problèmes de collocation, divisions qui nous semblent toujours pertinentes aujourd'hui.

- **Choix d'un collocatif inapproprié (collocation inexistante)** : Dans ce cas-ci, la collocation utilisée est plutôt inusitée et semble maladroite ou carrément déviante par rapport à l'usage : le collocatif choisi ne fait pas partie de la combinatoire lexicale de la base de la collocation (ex. **Les élèves devraient **appliquer** un plus grand respect envers les enseignants.*).
- **Choix d'un collocatif valable, mais dont le sens ne correspond pas au sens visé** : Dans un tel cas, la collocation existe bel et bien, mais ne véhicule pas le sens visé. En d'autres termes, le collocatif n'est pas une valeur de la fonction lexicale correspondant au sens que voulait exprimer le locuteur. Par exemple, l'élève qui écrit *#Les parents devraient **avoir** plus de discipline à la maison avec les jeunes.* alors qu'il veut exprimer le sens 'imposer de la discipline'.
- **Non-respect des contraintes d'emploi du collocatif** : Il arrive que le collocatif servant à exprimer un sens auprès de la base présente certaines contraintes d'emploi, contraintes qui doivent être respectées pour que la collocation soit correcte. Par exemple, il semble que le collocatif *porter*, qui fonctionne avec ATTENTION, exige que le nom soit précédé d'un déterminant possessif (*Les scientifiques portent **leur** attention sur cette nouvelle molécule.*); dans le cas où le déterminant est indéfini, le nom doit absolument être modifié pour que la collocation *porter une attention* soit correcte (*Ils portent **une** attention **particulière** à ce type d'élèves.*).
- **Erreur de dérivation sémantique** : Même si ce type de problème ne relève pas de la collocation à proprement parler¹⁰², nous avons choisi de le placer ici parce que le phénomène de dérivation sémantique peut aussi être modélisé à l'aide des fonctions lexicales. Une seule erreur de ce type avait été relevée dans notre corpus de maîtrise, ce qui avait conduit à l'ajout de la présente classe. Un exemple de ce type de problème serait l'emploi de POSTE pour parler d'une *caserne* [de pompiers].

COMBINATOIRE PRAGMATIQUE :

- **Emploi d'une UL appartenant à un registre de langue inapproprié** : Les ouvrages de référence se sont révélés très utiles pour le classement de ce type d'erreurs. Dans les cas où leur position était divergente, nous avons adopté celle du dictionnaire le plus permissif (en l'occurrence celui

¹⁰² En effet, le dérivé sémantique peut tout à fait apparaître sans le terme à partir duquel il est dérivé; il ne s'agit pas toujours de cooccurrence. Par exemple, AUTEUR est un dérivé sémantique de ROMAN (plus exactement son S₁, c'est-à-dire le nom typique de son premier actant sémantique), mais n'apparaît pas nécessairement en cooccurrence avec cette UL.

qui affirmait, par l'absence de marque d'usage, que l'UL en jeu appartient au registre standard).

- **Emploi d'une UL véhiculant un contenu connotationnel incompatible avec le contexte :** Le contenu connotationnel des UL n'étant pas toujours indiqué explicitement dans les ouvrages de référence, il nous a fallu faire appel à notre intuition de locuteur pour placer les erreurs dans cette classe.
- **Emploi d'une UL appartenant à une variété dialectale incompatible avec le contexte :** Les erreurs de ce type sont à peu près absentes de notre corpus, puisque les textes recueillis sont écrits par des élèves francophones québécois dans un contexte de français québécois standard.

STYLE : Comme nous l'avons mentionné dans le cadre conceptuel, certains problèmes liés au style ont à voir avec la dimension lexicale, par exemple la répétition, l'imprécision ou la périphrase. Par contre, ces problèmes donnent souvent lieu à des jugements très variables quant à leur correction, puisqu'ils ne sont pas clairement « condamnables » en vertu d'une norme. Par conséquent, ils se prêtent mal à une analyse quantitative comme celle que nous privilégions dans cette thèse. Par ailleurs, ces problèmes ne sont pas liés à la transgression d'une propriété d'une UL, ce qui fait qu'ils ne sont pas couverts par notre définition de l'erreur lexicale. Pour ces raisons, nous avons décidé de ne pas en tenir compte dans l'analyse de notre corpus.

Une remarque mérite par contre d'être faite à propos de la périphrase, qui est une formulation lourde produite pour exprimer une idée qui aurait pu l'être de façon plus économique et élégante; bien que la formulation utilisée ne soit pas erronée en soi, le résultat est peu idiomatique et n'aurait jamais été produit par un locuteur expert. Ce type de problème est à distinguer de la « forme analytique », que nous avons classée parmi les problèmes formels. Lorsqu'un scripteur produit une forme analytique, il utilise une formulation se rapprochant d'une définition par décomposition sémantique, alors qu'il aurait clairement pu utiliser une UL précise, ce qui constitue donc réellement une erreur lexicale. Par contre, dans le cas de la périphrase, le scripteur a recours à une formulation lourde, mais qu'on ne peut pas vraiment qualifier de forme analytique (ou remplacer par un seul

terme); dans un tel cas, le problème est plutôt d'ordre stylistique et ne sera pas pris en compte dans notre analyse.

3.1.2 LE DOUBLE CODAGE

Dans tout travail s'intéressant au classement d'erreurs à l'aide d'une typologie le moins précise se pose la question du double codage. En effet, certains problèmes peuvent parfois être décrits de différentes façons.

À priori, le fait que certains scénarios soient concurrents dans le cas de certaines erreurs ne semble pas poser problème aux chercheurs. Selon Agustín Llach (2005a), il est normal que les différents types d'erreurs des typologies ne soient pas mutuellement exclusifs, mais il est important qu'ils soient le moins ambigus possible. Par rapport à leur typologie d'erreurs lexicales, Milićević et Hamel (2007) ont aussi été confrontées au double codage :

On notera qu'il existe un recoupement entre les erreurs ayant trait au signifiant de L et celles ayant trait à la combinatoire morphologique de L ; ainsi la forme erronée d'une L peut être due au fait que L a été associée à la mauvaise classe flexionnelle. De tels recoupements sont normaux dans une typologie aussi détaillée que la nôtre. (p.158)

Zimmermann (1986: 33) est encore plus catégorique lorsqu'il parle, en s'appuyant sur les travaux de Legenhausen (1975) sur les erreurs grammaticales, de la « nécessité » des doubles classifications. L'auteur donne l'exemple d'un scripteur ayant utilisé le verbe GET RID [ɔʁ] à la place du verbe ERADICATE; non seulement s'agit-il ici d'une erreur sémantique, GET RID étant trop large pour le contexte de la phrase, mais en plus, le verbe utilisé appartient au langage familier. Nous nous sommes donc donné la possibilité d'attribuer deux étiquettes à une même erreur lors notre codage, en gardant en tête que les problèmes ayant reçu un double codage devraient faire l'objet d'un traitement particulier au moment de l'analyse des données, afin que chacun ne soit pas comptabilisé plus d'une fois, ce qui fausserait nos conclusions. Nous verrons dans notre chapitre de résultats (VIII-2.6.2) pourquoi nous sommes finalement revenu sur notre décision lors de l'analyse des données et que nous avons finalement évité le double codage des erreurs.

3.2 PROPOSITION D'UNE ÉCHELLE DE GRAVITÉ DES ERREURS

Nous avons évoqué dans notre cadre conceptuel la question de la gravité des erreurs, qui sera pertinente dans le cadre de notre thèse pour deux raisons. Premièrement, nous voulons voir s'il existe certains liens entre les types d'erreurs lexicales et la perception de la gravité de ces erreurs, ce qui nous permettrait notamment de vérifier s'il est vrai que les problèmes sémantiques sont généralement perçus comme le type le plus grave d'erreurs lexicales. Nous aimerions aussi voir si les erreurs que nous percevons comme étant plus graves sont celles qui sont pénalisées par les enseignants lors de la correction de productions écrites.

Comme nous l'avons vu plus haut, une opposition dichotomique (*erreur vs maladresse*) comme celle que nous avons privilégiée dans notre mémoire de maîtrise (Anctil, 2005) n'est pas satisfaisante pour rendre compte du positionnement des problèmes sur un continuum allant du « unanimement jugé non acceptable » au « parfaitement acceptable ». Nous avons à l'origine pensé avoir recours à une échelle de Likert, qui nous permettrait de situer avec plus de finesse les problèmes rencontrés sur un continuum. Par contre, nous ne croyons pas que ce type d'échelle puisse rendre compte du fait que plusieurs aspects sont impliqués dans un jugement de gravité (voir IV-3.2). Par conséquent, les jugements de gravité émis avec ce type d'outil semblent très subjectifs.

Pour diminuer un peu cette part de subjectivité inhérente au jugement de gravité, nous proposons d'opérationnaliser le concept de gravité en trois critères, qui nous semblent les plus pertinents pour l'erreur lexicale parmi ceux proposés dans les travaux sur la gravité des erreurs. Privilégier une évaluation à trois critères plutôt qu'une appréciation globale permettra de mieux comprendre ce qui motive les jugements posés. Voici les critères que nous avons retenus :

1. L'acceptabilité linguistique (cf. clarté de la norme)
2. L'impact sur la compréhension
3. Perception de l'intégration de l'erreur à l'usage

Chacun des trois critères est évalué séparément avant que le « calcul de la gravité » soit effectué. Pour chaque dimension évaluée, l'analyste attribue un score de 0 à 2, selon les indications qui apparaissent dans le tableau ci-dessous, ce qui fait que chaque problème peut potentiellement obtenir un score entre 1 et 6. Si nous n'envisageons pas la possibilité qu'un item obtienne un score de 0, c'est que cela impliquerait qu'il soit attesté par les ouvrages, qu'il ne pose aucun problème de compréhension et qu'il semble tout à fait intégré à l'usage en français. Comme nous nous intéressons seulement aux erreurs dans le cadre de cette thèse, les items linguistiques présentant ces trois caractéristiques n'ont pas été retenus pour notre analyse.

Le premier critère, *l'acceptabilité linguistique*, nous permet d'analyser le problème en regard des normes descriptive et prescriptive telles qu'elles se présentent dans les ouvrages de référence retenus (voir 2.1); ce critère n'est pas subjectif dans la mesure où il s'appuie sur des sources extérieures à l'analyste. Les deux autres critères relèvent quant à eux du jugement du correcteur et laissent donc plus de place à la subjectivité. Le deuxième critère, *impact sur la compréhension* (cf. compréhension), a été retenu parce qu'il apparaît dans la plupart des travaux s'intéressant à la gravité des erreurs et qu'on lui attribue généralement un rôle prépondérant dans les jugements posés par les locuteurs. Finalement, la *perception de l'intégration de l'erreur à l'usage* nous permet de rendre compte de la norme évaluative dont nous avons fait mention à la section III-2.1; en effet, nous croyons qu'il arrive souvent que nos jugements se fondent principalement sur nos propres intuitions de locuteurs. Ce critère nous permet aussi de constater que ces intuitions peuvent aller à l'encontre de ce que l'on trouve dans les ouvrages de référence (critère 1).

Tableau 13. Échelle de gravité des erreurs

Acceptabilité linguistique selon les ouvrages		est attesté par les ouvrages	0
		n'apparaît pas dans les ouvrages	1
		est en contradiction avec les ouvrages de référence (<i>téléphoner qqn</i>) ou y est explicitement condamné (<i>*téléphone</i> au sens 'appel téléphonique')	2
Jugement du correcteur	Impact sur la compréhension	n'a aucun impact sur la compréhension	0
		peut être interprété, mais pas avec assurance	1
		compromet la compréhension	2
	Perception de l'intégration de l'erreur à l'usage	la formulation semble tout à fait naturelle en français (intégration au code, ex. anglicisme comme <i>opportunité</i>)	0
		la formulation semble maladroite et marquée par rapport à l'usage normal	1
		la formulation semble s'écarter fortement de l'usage normal	2

Évidemment, l'utilisation de l'échelle de gravité que nous proposons ici présente certaines limites. Tout d'abord, nous n'avons retenu que les trois critères qui nous semblaient les plus pertinents. Parmi les critères présentés par James (1998) et Johansson (1978), certains ne nous paraissaient pas pertinents dans le cas d'erreurs lexicales (par exemple la fréquence ou la généralité), d'autres étaient difficilement opérationnalisables sous forme d'échelle (par exemple la saillance) et certains nous semblaient un peu redondants par rapport à ceux que nous avons retenus. Par exemple, nous croyons que notre troisième critère englobe, en statuant sur le caractère naturel d'un emploi linguistique, une appréciation de son adéquation pragmatique. C'est pour ces raisons que nous n'avons retenu que trois critères pour l'évaluation de la gravité des erreurs; par contre, nous sommes conscient qu'il se peut que d'autres facteurs jouent dans l'évaluation de la gravité d'un problème lexical et que notre grille n'en donne qu'un portrait partiel.

Le deuxième reproche que l'on pourrait faire à l'endroit de la grille d'évaluation de la gravité des problèmes lexicaux que nous proposons et d'accorder un poids identique aux trois critères dans le jugement de gravité. En effet, il est probable qu'un des critères, l'impact sur la compréhension du lecteur par exemple, ait une répercussion plus importante sur le jugement de gravité que les autres. Ce poids

relatif des différents critères est occulté par une grille comme celle que nous proposons.

Finalement, la méthode d'évaluation de la gravité des erreurs proposée ici n'est pas absolue et les jugements posés pourront varier en fonction de l'attitude plus ou moins normative de la personne qui posera les jugements, principalement en ce qui concerne les critères 2 et 3. Ce problème n'est pas unique à notre méthode d'évaluation, mais découle plutôt de la subjectivité inhérente à tout jugement de gravité. Il nous sera bien entendu impossible de soumettre l'ensemble des problèmes recueillis à un nombre élevé de juges afin de considérer une moyenne dans les jugements de gravité dans notre analyse. Néanmoins, nous pensons qu'en soumettant une partie des problèmes à d'autres juges pour un contre-codage (voir 4), nous serons en mesure de constater l'ampleur de la variation dans les jugements.

Malgré les limites soulevées, nous croyons que l'échelle de gravité que nous proposons, par sa simplicité et sa convivialité, permettra de porter des jugements nuancés sans trop alourdir le processus de codage des erreurs lexicales. Appliquée de façon cohérente, elle permettra de faire ressortir les écarts dans la gravité des différents problèmes analysés.

3.3 LA TYPOLOGIE EXPLICATIVE DES ERREURS LEXICALES

Nous avons constaté dans le cadre conceptuel que les travaux sur l'analyse d'erreurs en général, et l'analyse d'erreurs lexicales en particulier, nous laissent avec bien peu de pistes pour expliquer les erreurs lexicales commises dans un contexte de L1. La solution méthodologique la plus pertinente pour nous était donc de partir des quelques scénarios d'explication qui nous semblaient pertinents pour constituer notre typologie explicative. Cette typologie s'apparentait au départ à une simple liste de sources d'erreurs, liste que nous avons enrichie au fur et à mesure de notre analyse; il s'agissait donc d'un classement de type ouvert. Voici donc la liste de départ que nous avons utilisée pour l'explication des erreurs du corpus.

Tableau 14. Typologie explicative des erreurs lexicales

Méconnaissance d'une UL ou d'une propriété d'une UL	
Problème d'accès à un mot en mémoire	
Application d'une règle de dérivation inappropriée (utilisation de préfixe/suffixe, changement de classe de mots) ¹⁰³	
Calque de l'anglais (sémantique ou lexicosyntaxique)	
Intégration de l'erreur à la langue (fait d'avoir entendu ou lu la forme produite à plusieurs reprises) = erreur induite	
Extension sémantique abusive : par métaphore, métonymie ou causativité	
Ellipse sémantique : L'erreur découle du fait que le scripteur omet certaines informations implicites, ce qui conduit à des cooccurrences inhabituelles (ex. <i>*Le renforcement positif démontre (= 'favorise') une amélioration rapide des résultats des élèves. → Des études démontrent / Il a été démontré que le renforcement positif favorise une amélioration...</i>).	
Amalgame de deux locutions	
Analogie	formelle (confusion entre des paronymes ou insécurité formelle) : les problèmes de ce type sont liés au concept de <i>synforms</i> (V-4.4)
	sémantique (confusion entre des termes sémantiquement proches)
	de construction : transposition de la structure syntaxique régie d'une acception différente du même vocable, d'un synonyme ou d'un terme sémantiquement lié à une unité lexicale qui n'a pas le même régime
	choix d'un collocatif appartenant à une autre acception du même vocable ou à une unité lexicale sémantiquement proche (greffe collocationnelle)

Nous soulignons une fois de plus le caractère hypothétique des scénarios d'explication présentés dans cette grille, puisque nos diagnostics d'erreurs seront posés sans faire appel aux scripteurs ayant commis ces dernières.

4. LE CONTRE-CODAGE

Nous avons soumis une partie de notre corpus à un contre-codage pour certains des critères d'analyse des erreurs : ceux mettant en jeu une plus grande subjectivité, c'est-à-dire ceux relatifs à l'évaluation de la gravité de l'erreur et à son explication. Si nous avons décidé de ne pas faire contre-coder la dimension plus descriptive de l'erreur, c'est parce que nous jugions que l'utilisation de notre typologie descriptive des erreurs lexicales, vu sa complexité, demande une compréhension profonde de nombreux concepts liés à l'étude du lexique et que de

¹⁰³ Comme le souligne Warren (1982), la règle de dérivation peut dans certains cas être appliquée à la mauvaise acception d'un vocable (ex. *Il a interprété les résultats de son étude de façon remarquable. *Quel interprète! – élever des enfants ~ *élevage des enfants*). Elle présente aussi des cas, plus rares, où il y a confusion entre deux termes dérivés d'une même base ou de problème d'identification de la base (voir V-4.1).

former des contre-codeurs pour qu'ils atteignent un niveau d'aisance suffisant aurait demandé trop de temps. Par ailleurs, le contre-codage que nous avons effectué sur une grille similaire dans le cadre de notre recherche de maîtrise (Anctil, 2005) avait produit un accord interjuges satisfaisant, ce qui nous permet de supposer que cela aurait été le cas ici aussi; d'ailleurs, les nombreuses séances de codage avec notre directrice de recherche, habituée à travailler avec notre grille, nous ont montré que nous associions les erreurs aux mêmes classes dans la très grande majorité des cas. Nous aurions bien sûr pu contre-coder nous-même une partie du corpus, mais l'analyse des données a permis de démontrer que notre codage de départ était très homogène. Le contre-codage a donc porté spécifiquement sur l'évaluation de la gravité des erreurs et sur leur explication.

Le contre-codage a été assuré par deux collègues étudiantes au doctorat en didactique; l'une, didacticienne du lexique, a une formation en linguistique et connaît bien la Théorie Sens-Texte, tandis que l'autre est plutôt spécialisée en didactique de la grammaire et possède une expérience d'enseignement du français au secondaire. Après leur avoir présenté notre *Guide de contre-codage* (voir annexe 4) et offert une formation d'environ 1 h 30, nous avons soumis à ces deux contre-codeuses 10 % des erreurs de notre corpus (sans les erreurs d'orthographe, soit 130 erreurs), choisies de façon aléatoire. Les erreurs étaient présentées dans un court contexte mettant en évidence le problème, identifiées en gras et accompagnées d'une proposition de correction. Les contre-codeuses ont ensuite eu deux semaines pour effectuer le codage.

En ce qui concerne l'évaluation de la gravité des erreurs, nous avons choisi de faire contre-coder deux des trois critères, soit *l'impact de l'erreur sur la compréhension* et *la perception de l'intégration de l'erreur à l'usage*; *l'acceptabilité linguistique selon les ouvrages* n'a pas fait l'objet d'un contre-codage, puisque le score attribué pour ce critère dépendait directement de ce que nous trouvions dans les dictionnaires et qu'il aurait été inutile que les contre-codeuses refassent le travail de recherche dans les ouvrages que nous avons nous-même fait de façon méticuleuse. Pour ce qui est de l'explication des erreurs, nous avons

simplement demandé aux codeuses de fournir une hypothèse (ou plus d'une) chaque fois qu'elles croyaient pouvoir expliquer le problème rencontré. Ainsi, comme nous, elles n'avaient à proposer un scénario d'explication que lorsqu'un tel scénario leur apparaissait spontanément. Dans le but de ne pas orienter leur réflexion sur les sources possibles des erreurs lexicales, nous ne leur avons fourni aucune liste de sources possibles.

Mentionnons pour conclure sur le contre-codage que ce que nous avons transmis aux codeuses était un fichier *Excel* présentant directement les erreurs à contre-coder plutôt que des rédactions en entier, ceci dans le but d'alléger leur tâche. Par conséquent, l'étape de repérage des erreurs était absente du contre-codage, ce qui fait que nous n'avons pas obtenu d'avis extérieurs concernant cette partie de la démarche d'analyse¹⁰⁴. Les résultats relatifs au contre-codage seront présentés aux sections VIII-3.1 et VIII-4.1.

5. ANALYSE DES DONNÉES

La correction des 300 productions écrites et le codage des quelque 2700 erreurs récoltées se sont étalés sur une période de quatre mois, de novembre 2009 à mars 2010. Chaque erreur relevée dans le corpus a été codée selon l'ensemble des critères présentés dans ce chapitre, sauf les erreurs orthographiques pour lesquelles nous n'avons pas proposé de scénarios d'explication.

Ce volume important de données a ensuite pu être analysé à l'aide du logiciel *Excel*, qui nous a permis d'effectuer un nombre important de tris et de générer plusieurs tableaux croisés dynamiques pour mettre en relation différentes variables.

C'est par ces manipulations sur notre corpus d'erreurs que nous avons pu répondre à notre premier objectif de recherche, soit la description des erreurs lexicales commises par des élèves de 3^e secondaire; ces résultats seront présentés au chapitre VIII.

¹⁰⁴ Nous verrons tout de même, à la section X-1.1, quelles sont les erreurs lexicales qui ont été relevées par les enseignants dans les rédactions de leurs élèves.

6. LES LIMITES ASSOCIÉES À NOTRE PREMIER OBJECTIF

Tout choix méthodologique se répercute sur la fiabilité et la validité des données recueillies ainsi que sur l'analyse qu'on en fait et influence nécessairement le potentiel de généralisation des observations qui en découlent; notre recherche n'y échappe pas. Voici les principales limites liées à notre premier objectif de recherche.

Les choix effectués quant aux critères de sélection des élèves retenus pour participer à l'étude l'ont été dans le but de produire un échantillon homogène représentant les élèves francophones de troisième secondaire issus de milieux socioéconomiques moyens; cette homogénéité, compte tenu du nombre de textes récoltés (300), nous paraissait une condition *sine qua non* à la constitution d'un corpus représentatif. Les résultats de la présente étude se limitent donc à cette population et ne peuvent être généralisés à des élèves pour qui le français est une langue seconde ou qui proviennent de milieux socioéconomiques différents; il faut aussi être prudent si nous voulons appliquer nos conclusions aux élèves des autres niveaux du secondaire, et encore plus des autres ordres d'enseignement. Il est aussi important de rappeler que la sélection des élèves a été faite sur la base seulement de l'âge, du profil linguistique et de la situation socioéconomique de l'école fréquentée : aucun contrôle n'a été fait sur le degré de maîtrise en français écrit des élèves. Notre corpus de textes comprend donc tant des productions d'élèves forts que d'élèves plus faibles et notre analyse ne permet pas d'établir de liens entre les types d'erreurs lexicales commises et le degré de maîtrise en français écrit. Encore une fois, cette limite ne nous semble pas particulièrement problématique dans la mesure où nous voulons fournir un portrait global des erreurs lexicales des élèves francophones de troisième secondaire.

Les rédactions constituant notre corpus présentent certaines différences quant aux types de textes produits (narratif et explicatif/informatif) et à la consigne d'écriture proposée aux élèves. Si ces paramètres influencent peut-être les erreurs lexicales commises – ce que notre recherche n'a pas pour but de vérifier –, nous pensons qu'une certaine part d'aléatoire dans la constitution du corpus

présente tout de même l'avantage de permettre une description « générale » des erreurs lexicales des élèves, c'est-à-dire qui transcende les types de textes ou les thèmes abordés. Par ailleurs, le fait que nous ayons recours à des textes qui sont produits dans un cadre scolaire régulier et qui sont extérieurs à la recherche (cf. données invoquées, voir IX-1.1.5) procure à notre corpus un caractère naturel susceptible de mener à un portrait réel des difficultés des élèves. Relativement au corpus, il est aussi important de souligner que nous n'avons pas réussi à obtenir les versions déjà corrigées des productions dans tous les cas – 76 copies sur les 300 n'étaient pas corrigées; par conséquent, nous n'avons pas pu comparer systématiquement notre correction à celle des enseignants.

Relativement à l'analyse des erreurs lexicales, certaines limites doivent aussi être considérées. Tout d'abord, rappelons que toute entreprise d'analyse d'erreurs contient une certaine part de subjectivité qu'il est impossible d'éliminer complètement; cette subjectivité est particulièrement présente au moment du repérage des erreurs et lors de l'évaluation de leur gravité. Le contre-codage d'une partie du corpus d'erreurs nous a permis de limiter cette part de subjectivité en confrontant notre analyse à celles de deux contre-codeuses, mais, comme nous l'avons mentionné plus haut, ce contre-codage occulte certaines dimensions de l'analyse, notamment l'étape de repérage et de description linguistique des erreurs ; par ailleurs, il ne portait que sur 10 % du corpus.

Ensuite, nous soulignons de nouveau le caractère hypothétique des scénarios d'explication des erreurs; consulter les élèves à l'origine des erreurs plutôt que de poser des hypothèses à partir des seuls textes aurait pu éclairer davantage la question de la source des erreurs, mais aurait considérablement alourdi le processus d'analyse. Nous jugeons que le volume d'erreurs analysées compense en quelque sorte cette limite méthodologique.

Vu l'ampleur du corpus de textes à analyser, nous avons aussi dû limiter nos efforts aux problèmes de nature lexicale; ainsi, notre analyse ne permet pas d'établir la proportion d'erreurs lexicales par rapport aux erreurs relevant des autres dimensions du code écrit (syntaxe, accords grammaticaux, cohérence,

etc.)¹⁰⁵, ce qui pourrait bien sûr être intéressant pour mesurer l'importance relative de l'erreur lexicale. Dans le même ordre d'idées, il aurait été intéressant de pouvoir comparer le nombre d'items lexicaux erronés au nombre de structures correspondantes bien utilisées; par exemple, si la rédaction d'un élève présente trois erreurs de prépositions régies, il aurait été intéressant de confronter ce nombre au nombre d'occurrences de prépositions régies bien utilisées dans son texte. Bien sûr, ceci aurait exigé un travail colossal et nous nous concentrons dans cette thèse seulement sur les occurrences erronées.

C'est en gardant en tête ces quelques limites que nous avons analysé nos données et tiré les conclusions de notre étude. Passons maintenant à la méthodologie associée à notre deuxième objectif de recherche.

¹⁰⁵ Nous avons songé à nous appuyer pour cela sur les erreurs relevées par les enseignants, mais comme le corpus contenait des textes corrigés par différentes personnes et un certain nombre de textes non corrigés, cette façon de procéder nous semblait peu concluante.

PREMIER OBJECTIF - RÉSULTATS

VIII. LES ERREURS LEXICALES EN PRODUCTION ÉCRITE EN 3^E SECONDAIRE

Le présent chapitre sera consacré à la présentation des résultats relatifs à notre premier objectif de recherche, soit l'analyse des erreurs lexicales tirées d'un corpus de textes produits par des élèves de troisième secondaire. Nous présenterons tout d'abord le corpus de textes analysé, avant de nous intéresser tour à tour aux trois niveaux de notre démarche d'analyse : la description linguistique des erreurs (section 2), l'évaluation de leur gravité (section 3) et l'explication de leurs sources (section 4).

1. LE CORPUS DE TEXTES

Notre corpus d'analyse comprend 300 rédactions d'élèves francophones de troisième secondaire réalisées en classe à l'automne 2007 ou au printemps 2008 dans le cadre d'une évaluation sommative de fin d'étape ou de fin d'année; 62 % des élèves ayant fourni une rédaction sont des filles et 38 %, des garçons. Les productions proviennent de quatre écoles de taille variable – moins de 300 élèves à plus de 3000 – de différentes régions : une école de l'ouest de l'île de Montréal (48 copies), une de la Rive-Sud de Montréal (75 copies), une située sur la Couronne Nord (73 copies) et une en Estrie (104 copies).

Le thème abordé dans les rédactions variait d'un établissement à l'autre. Environ la moitié des élèves (148) devaient écrire un conte merveilleux (texte narratif) à partir de certains éléments proposés dans la consigne de rédaction : personnage, lieu, événement déclencheur. Près d'un sixième des rédactions (48) consistaient en la description d'une publicité et la présentation des techniques publicitaires qui y sont utilisées (texte descriptif/explicatif). Finalement, le tiers des productions écrites du corpus (104) portaient sur le thème du travail d'été; les élèves étaient invités soit à exposer les avantages et les inconvénients d'un emploi

d'été (texte explicatif), soit à parler de leur première expérience en tant que jeune occupant un tel emploi (texte narratif). Environ les deux tiers des textes sont donc à dominante narrative, alors que le tiers restant est à tendance descriptive/explicative.

Les rédactions, dont le temps total de production allait de 3 h 15 à 3 h 45, ont été réalisées soit en une seule fois soit en plusieurs périodes de 75 minutes; dans tous les cas, le recours au dictionnaire, à la grammaire et à un guide de conjugaison était permis. Un travail de préparation à l'écriture a été réalisé en classe pour l'ensemble des groupes : lecture de contes merveilleux ou d'articles traitant du travail d'été, leçons sur la structure du texte narratif ou du texte descriptif/explicatif, visionnement d'un reportage sur la publicité, examen de compréhension écrite sur des textes liés au thème de rédaction. Nous pouvons supposer que ce travail préalable a fourni aux élèves un certain vocabulaire de base pour traiter le thème imposé.

La longueur moyenne des rédactions recueillies est de 385 mots, alors que la longueur minimale demandée était de 350 mots; la plus courte rédaction contenait 190 mots et la plus longue, 716¹⁰⁶. Comme la plupart des productions écrites avaient déjà été évaluées au moment de notre collecte de données, nous avons pu obtenir 224 rédactions corrigées par les enseignants, ce qui représente 75 % du corpus. Ces copies nous ont permis d'observer de quelle façon les enseignants traitent l'erreur lexicale dans leur correction (voir X-1); si ces informations ne sont pas cruciales à notre analyse d'erreurs, elles nous ont permis certaines observations qui ajouteront un éclairage supplémentaire aux réponses fournies par les enseignants lors des entretiens de notre deuxième objectif, puisqu'elles constituent une fenêtre ouverte sur les pratiques réelles de correction.

¹⁰⁶ Hormis la courte rédaction de 190 mots et trois rédactions de plus de 700 mots, la longueur des textes variait entre 240 et 607 mots.

2. LA DESCRIPTION DES ERREURS À L'AIDE DE LA TYPOLOGIE DES ERREURS LEXICALES

Comme nous l'avons expliqué dans notre méthodologie, notre méthode d'analyse des erreurs lexicales comporte trois niveaux : 1) une description linguistique à l'aide de la typologie d'erreurs proposée, 2) une évaluation de la gravité de chaque erreur à l'aide d'une échelle de gravité et 3) une explication de la source de l'erreur dans le cas où une hypothèse nous apparaît. La présente section portera précisément sur le premier niveau de notre analyse : la description des erreurs.

2.1 LES ERREURS LEXICALES AU SECONDAIRE : PORTRAIT GLOBAL

La correction des 300 rédactions a révélé un total de 2720 erreurs lexicales. La moyenne est donc de 9,1 erreurs par copie ou, plus précisément, de 8,4 erreurs par 350 mots¹⁰⁷. Par contre, cette moyenne est somme toute assez peu représentative, puisqu'un écart type élevé (=6) nous indique que le nombre d'erreurs lexicales varie énormément d'un élève à l'autre, tout comme le nombre d'erreurs en général d'ailleurs¹⁰⁸. Le tableau ci-dessous présente la répartition des erreurs des différents types.

¹⁰⁷ Si nous décidons ici d'établir la moyenne pondérée pour 350 mots, c'est qu'il s'agit de la longueur de texte minimale spécifiée dans les consignes d'écriture.

¹⁰⁸ Comme la correction des rédactions recueillies a été faite par sept enseignants différents et l'évaluation soumise à différents barèmes (notes en pourcentages, en lettres, en nombre d'erreurs), nous n'avons pu établir le nombre total d'erreurs par copie de façon systématique, ce qui nous aurait permis d'établir la proportion d'erreurs lexicales par rapport au nombre d'erreurs linguistiques en général. Par contre, nous avons constaté, à l'aide des rédactions déjà corrigées par les enseignants sur lesquelles apparaissait le nombre total d'erreurs relevées, que le nombre d'erreurs lexicales semblait être proportionnel au nombre total d'erreurs, particulièrement si l'on inclut les erreurs orthographiques; en d'autres mots, un élève qui commet un nombre important d'erreurs de tous types commet aussi davantage d'erreurs lexicales.

Tableau 15. Répartition des erreurs selon les classes de la typologie

Grandes classes d'erreurs	Types d'erreurs ¹⁰⁹	Nombre d'erreurs	Pourcentage de chaque type d'erreurs dans chaque grande classe	Nombre d'erreurs dans chaque grande classe	Pourcentage du total d'erreurs
Forme	F-Orth	1426	78,5	1817	66,8
	F-Barb	209	11,5		
	F-Flex	150	8,3		
	F-Inex	21	1,2		
	F-Déf	11	0,6		
Sens	S-Proc	107	31,3	342	12,6
	S-Fict	85	24,9		
	S-Syn	42	12,3		
	S-Act	39	11,4		
	S-Ang	36	10,5		
	S-Pléo	33	9,6		
Combinatoire pragmatique	CP-Regis	192	98,5	195	7,2
	CP-Conn	2	1,0		
	CP-Dial	1	0,5		
Combinatoire syntaxique	CS-ChPr	49	38,3	128	4,7
	CS-Const	30	23,4		
	CS-AbsC	23	18,0		
	CS-Trans	15	11,7		
	CS-AbsPr	6	4,7		
	CS-Intrans	5	3,9		
Combinatoire lexicale	CL-Inex	100	78,7	127	4,7
	CL-Cont	15	11,8		
	CL-FL	8	6,3		
	CL-Dér	4	3,1		
Combinatoire morphologique	CM-Inv	51	45,9	111	4,1
	CM-Genr	47	42,3		
	CM-Class	7	6,3		
	CM-Aux	3	2,7		
	CM-Rés	3	2,7		
Total		2720			

Ce qui frappe à première vue à la lecture de ce tableau est le nombre élevé d'erreurs liées à la forme. Les 1817 erreurs formelles représentent en effet les deux tiers (66,8 %) du corpus. On constate aussi que la très grande majorité de ces problèmes (78,5 %) sont des erreurs d'orthographe d'usage (F-Orth). Comme ce type d'erreurs est souvent traité à part par les enseignants lors de la

¹⁰⁹ Les codes apparaissant dans cette colonne correspondent à ceux de la typologie présentée notre chapitre de méthodologie (VII-3.1.1, Tableau 12).

correction¹¹⁰ et n'est pas considéré comme relevant de problèmes lexicaux dans la plupart des travaux sur l'erreur lexicale, nous avons décidé de les exclure temporairement de nos résultats afin de conserver les erreurs plus « purement » lexicales. Bien entendu, une analyse en profondeur des erreurs orthographiques pourrait être faite, mais renverrait à un autre cadre théorique (par exemple Catach *et coll.*, 1986) et ne constitue pas l'objet de la présente thèse; nous ferons tout de même certaines observations sur les erreurs d'orthographe à la section 2.5.

Les erreurs formelles liées à la flexion (F-Flex) posent quant à elles un problème pour l'analyse des données, car la plupart ont à voir avec la conjugaison des verbes (141/150), principalement au passé simple, temps privilégié (mais mal maîtrisé!) par les élèves ayant produit des textes narratifs. Comme l'orthographe, la conjugaison est traitée à part par les enseignants dans la correction. Par ailleurs, elle est aussi soumise à des règles strictes, qui ne sont pas propres à chaque verbe; les différentes formes fléchies correspondant à un verbe sont généralement prédictibles à partir de règles morphologiques et ne constituent pas, par conséquent, des connaissances proprement lexicales. Bien entendu, certains verbes sont irréguliers et la connaissance des différentes formes non prédictibles peut être assimilée à une connaissance lexicale, mais il demeure plutôt difficile de départager les verbes réguliers des verbes irréguliers. Par exemple, même un verbe aussi irrégulier qu'ALLER produit des formes tout à fait prédictibles à certains temps comme l'imparfait ou le passé simple, ce qui nous obligerait à considérer que l'élève qui écrit **j'allerai* commet une erreur lexicale – la forme *irai* étant non prédictible –, mais pas celui qui écrit *nous *vons* – la forme *allons* correspondant à la règle morphologique de formation de l'indicatif présent des verbes du premier groupe. À cause de cette difficulté, nous jugeons préférable d'exclure aussi les problèmes de flexion de notre analyse, de façon à ne pas gonfler le nombre d'erreurs de notre corpus avec des cas qui ne relèvent pas de connaissances purement lexicales, c'est-à-dire des connaissances concernant des

¹¹⁰ Toutes les grilles de correction de productions écrites que nous avons rencontrées présentent une rubrique « orthographe » ou « usage ».

propriétés idiosyncrasiques d'une UL. Soulignons toutefois que cette décision nous conduit à ne pas considérer dans notre analyse neuf erreurs qui n'affectent pas des verbes, mais qui sont pour la plupart aussi soumises à des règles de flexion du français : formation du pluriel de noms (**journals, *travails, *passes-temps, *attrapes-nigaud*) ou de locutions nominales (**coup de pieds, *coups de poings, *armes à feux*), choix de la forme de l'adjectif NOUVEAU en fonction de la position syntaxique (*cet entrain *nouvel*) et flexion d'un déterminant (*sans *aucuns doutes*).

L'exclusion des problèmes orthographiques et flexionnels ramène le nombre total d'erreurs recueillies à 1144, ce qui représente 3,5 erreurs par 350 mots (écart type = 2,5). Ce corpus plus modeste permet de mieux observer la répartition des grandes classes d'erreurs lexicales, que montre la figure ci-dessous.

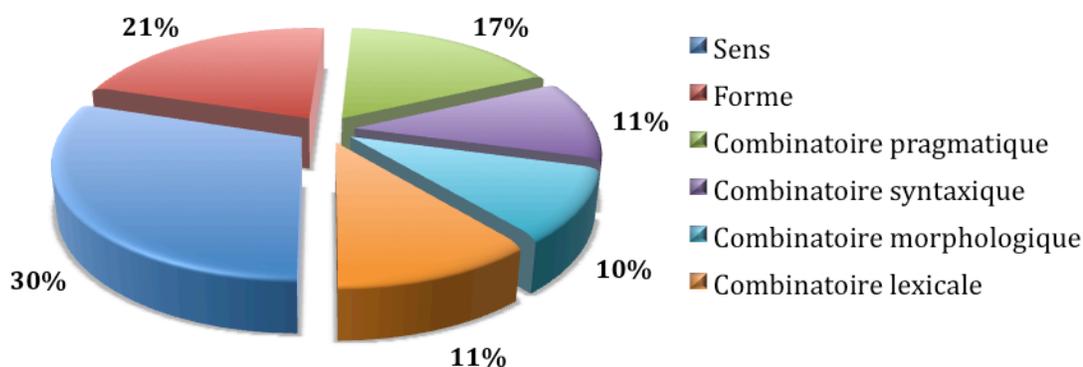


Figure 7. Répartition des 1144 erreurs lexicales (sans orthographe et flexion)

Les erreurs sémantiques occupent maintenant le premier rang (30 %), suivies par les problèmes formels (21 %), puis par les erreurs lexicopragmatiques (17 %); les problèmes lexicosyntaxiques, lexicomorphologiques et de combinatoire lexicale sont présents dans des proportions presque identiques (10 à 11 %). Ce qui est frappant, c'est que près de la moitié des erreurs (49 %) sont liées à la combinatoire des UL, d'où l'importance de ne pas considérer que la forme et le sens dans une définition du concept d'*erreur lexicale*. Observons maintenant la répartition des différents types d'erreurs à l'intérieur de chaque grande classe à l'aide d'exemples d'erreurs de chaque type.

2.2 TOUR D'HORIZON DES DIFFÉRENTES CLASSES D'ERREURS DU CORPUS

La présente partie du chapitre aborde tour à tour chaque grande classe d'erreurs de notre typologie pour présenter et discuter des exemples de chaque sous-classe d'erreurs. Nous commencerons par les erreurs les plus fréquentes, soit les erreurs sémantiques, et terminerons avec les erreurs les moins fréquentes : celles liées à la combinatoire lexicale. Chacune des sections commence par une présentation de la répartition des différentes sous-classes d'erreurs avant des les détailler en ordre décroissant de fréquence.

2.2.1 LES ERREURS SÉMANTIQUES

La figure suivante présente la répartition des 342 problèmes sémantiques.

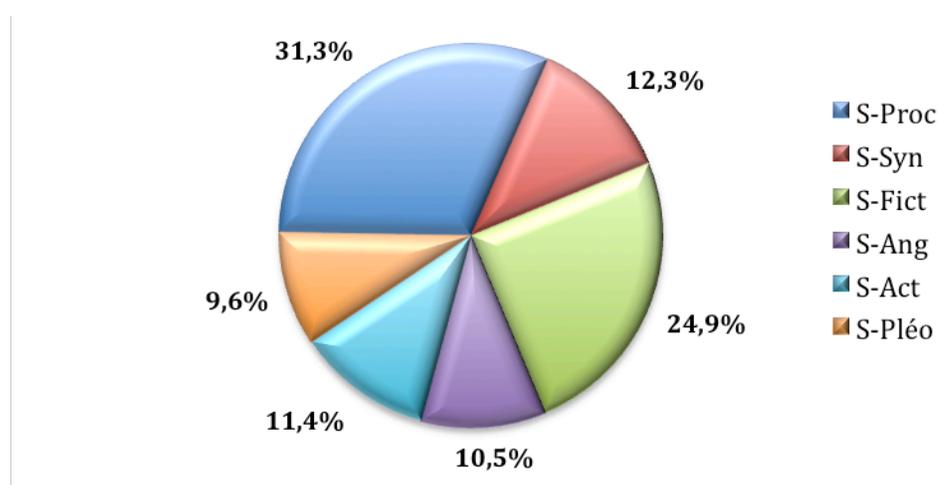


Figure 8. Répartition des 342 erreurs sémantiques

La confusion entre des termes liés sémantiquement constitue 43,6 % des erreurs sémantiques : 31,3 % pour les erreurs de sens proche (S-Proc = 107 cas) et 12,3 % pour les synonymes (S-Syn = 42 cas). Les impropriétés représentent 35,4 % des erreurs de sens de notre corpus, si l'on considère à la fois les impropriétés « pures » (S-Fict = 85 cas) et les anglicismes (S-Ang = 36 cas). La non-maitrise des contraintes pesant sur les actants sémantiques (S-Act = 39 cas) et les pléonasmes (S-Pléo = 33 cas) complètent le tableau des erreurs sémantiques.

2.2.1.1 LES CAS D'UTILISATION D'UNE UL AYANT UN SENS PROCHE OU SYNONYME DU SENS VISÉ

Comme nous l'avons mentionné, l'utilisation erronée d'une UL ayant un sens proche de celui de l'UL visée constitue le type d'erreurs sémantiques le plus fréquent. Voici quelques exemples d'erreurs¹¹¹ appartenant à cette classe.

- (1) *Il passait tous ses **passes-temps** [temps libres] sur celui-ci (= son projet). [31]*
- (2) *Tout d'un coup, les fous fût **stabilisé** [immobilisé] et Bobby traversa la forêt sans danger. [73]*
- (3) *Aux pieds du volcan Asama **trônait** [se trouvait] un village. [187]*

Dans la première phrase, le lien sémantique entre PASSE-TEMPS et TEMPS LIBRES est évident, puisque le sens de la seconde UL fait partie de celui de la première, un passe-temps étant une 'activité agréable que l'on pratique dans ses temps libres'. Dans la phrase (2), le lien est moins direct, mais il existe tout de même une parenté sémantique entre IMMOBILISER et STABILISER (= 'rendre stable' → STABLE = 'qui a tendance à ne pas bouger à cause d'une bonne répartition des points d'appui'). En (3), c'est plutôt une incompatibilité entre TRÔNER et la situation dénotée qui permet de constater que l'élève n'a pas une compréhension complète du sens de ce verbe : TRÔNER signifie 'être comme sur un trône, être bien en évidence', ce qui n'est pas le cas d'un village situé au pied d'un volcan. Pourtant, il existe bel et bien un lien de sens entre TRÔNER et SE TROUVER, puisque quelque chose qui trône sur quelque chose d'autre s'y trouve forcément, par exemple un village trônant au sommet d'une falaise. Dans ces trois erreurs, nous constatons un lien sémantique entre l'UL utilisée et l'UL visée, mais sans pouvoir considérer les deux items en jeu comme des synonymes¹¹², contrairement aux cas présentés dans les phrases suivantes.

- (4) *La mauvaise sorcière (...) ne laissait personne **rentrer** [entrer] dans la maison. [56]*

¹¹¹ Les phrases tirées des rédactions que nous reproduisons dans la thèse sont recopiées telles quelles, c'est-à-dire que nous n'y avons corrigé aucune erreur, ceci dans le but de fournir des extraits représentatifs du niveau de maîtrise linguistique des élèves. L'UL erronée apparaît en gras et la correction proposée, entre crochets. Le nombre entre crochets à la suite de l'exemple correspond quant à lui au numéro de la rédaction analysée. Comme toutes les phrases présentées contiennent des erreurs, nous avons omis l'astérisque marquant l'agrammaticalité.

¹¹² Le cas de la phrase (3) est moins clair, puisqu'il serait possible de considérer TRÔNER comme un hyponyme, donc un synonyme plus riche, de SE TROUVER, mais les dictionnaires ne répertorient pas ces deux verbes comme des synonymes.

- (5) *Tous se retournèrent pour s'apercevoir que Billy venait de se faire **kidnapper** [enlever] par une espèce de géant poilu. [206]*
- (6) *[L'héroïne du récit part à la chasse d'un dragon pour délivrer son village.] Après une longue **promenade** [marche] d'environ une heure, elle dut passer par la Vallée verte. [112]*

Dans ces trois phrases, l'erreur naît d'une confusion entre deux termes clairement synonymes qui ne se distinguent que par quelques nuances de sens. En (4), RENTRER est utilisé pour signifier 'entrer' plutôt qu' 'entrer de nouveau', alors que le héros n'a jamais pénétré dans la maison de la sorcière; cette confusion semble fréquente, car on la retrouve treize fois dans notre corpus. Dans le cas de KIDNAPPER, c'est la notion de 'rançon' associée à ce verbe (KIDNAPPER = 'enlever qqn dans le but d'obtenir une rançon') qui pose problème, puisqu'il y a fort à parier que le yéti enlève Billy plutôt pour le dévorer que pour en tirer quelconque avantage pécuniaire. Enfin, l'élève qui a écrit la phrase (6) ne semble pas conscient du caractère récréatif de la promenade, caractère plutôt incompatible avec une chasse au dragon! Dans ces trois cas, l'élève utilise un terme sémantiquement plus riche que l'UL visée; ce cas de figure, le plus représenté dans le corpus¹¹³, témoigne peut-être d'un effort d'utilisation d'un vocabulaire sophistiqué, mais pas totalement maîtrisé. Par contre, nous avons aussi rencontré des cas, plus rares, où l'UL utilisée est au contraire un hyperonyme de l'UL visée (exemple (7)) ou simplement un cohyponyme¹¹⁴ (exemple (8)).

- (7) *Triton décida, donc de prendre sa grande **fourche** [trident] d'or et il avait transformé Émina en crevette. [176]*
- (8) *Celui-ci prit un **bas**¹¹⁵ [chaussette], crayonna rapidement à l'aide d'un feutre et transforma le bas en marionnette. [273]*

Il nous semble aussi intéressant de pointer quelques erreurs mettant en jeu des déictiques, comme dans les phrases suivantes.

- (9) *Il y a plusieurs centaines d'années, La Roumanie (...) était infesté de petits groupes de vampire. **Jusqu'à maintenant** [Jusque là, Jusqu'alors], elles ne constituaient pas vraiment une menace pour la population. [161]*

¹¹³ D'autres exemples de ce type de confusion seraient PARADE au sens de 'défilé' [147], FILAMENT au sens de 'fil' [27], FRACTURER au sens de 'briser' [80] ou ISSUE pour 'ouverture' [107].

¹¹⁴ Par *cohyponyme*, nous entendons un terme qui se situe au même niveau qu'un autre par rapport à un hyperonyme dans une hiérarchie sémantique (voir II-4.3). Par exemple, nous pourrions dire que CHALOUPE et CANOT sont des cohyponymes par rapport à EMBARCATION.

¹¹⁵ Bien entendu, BAS au sens de 'chaussette' constitue aussi un québécoisisme de registre familier.

(10) *Le seul hic, c'était que le voyage était prévu pour l'été **prochain** [suivant]. [263]*

(11) *La nuit tombée, elle s'endormit avec l'espoir de pouvoir sortir d'**ici** [(de) là]. [199]*

Dans les deux premiers cas, l'emploi des déictiques MAINTENANT et PROCHAIN est problématique, puisque ces deux UL n'ont de sens qu'en relation avec le moment présent : MAINTENANT signifie 'au moment présent, au moment où je prononce *maintenant*' et PROCHAIN signifie 'qui suit chronologiquement la date où je prononce *prochain*' (ex. *mardi prochain, l'an prochain*). Comme les textes dans lesquels apparaissent les phrases (9) et (10) sont ancrés dans le passé, il est impossible d'y utiliser ces deux UL. Le problème est similaire en (11), sauf que l'incompatibilité du déictique ICI avec le contexte d'énonciation est spatiale plutôt que temporelle. ICI signifie 'à l'endroit où se trouve la personne qui prononce *ici*'; comme c'est à ce moment le narrateur (omniscient) qui s'adresse au lecteur, il ne peut utiliser le déictique ICI, puisqu'il ne se trouve pas lui-même à l'endroit auquel il fait référence. Évidemment, si c'est l'héroïne qui s'était posé la question « *Comment vais-je sortir d'ici?* », la phrase aurait été tout à fait correcte.

Comme la linguistique textuelle l'a bien montré (cf. par exemple Adam (2008)), les déictiques constituent un problème redoutable pour les élèves en situation de production de textes et il n'est pas toujours aisé d'expliquer l'origine de ce type d'erreurs. Un travail précis sur ce type d'UL peut s'avérer intéressant avec les élèves, notamment pour voir de quelle façon les dictionnaires décrivent le sens de ces items lexicaux particuliers.

2.2.1.2 LES IMPROPRIÉTÉS

Les impropriétés représentent près du quart des erreurs sémantiques. Il s'agit de cas où l'UL utilisée ne véhicule pas du tout le sens visé, comme dans les phrases suivantes.

(12) *Tous les objets de la cuisine se déplaçaient, ce qui causait tout un **charabia** [désordre]. [18]*

(13) ***Dernièrement** [Finalement], un slogan apparaît également au bas de la page. [159]*

(14) *Quand on **fait appel** [fait référence] à Noël, bien du monde pensent au mot "cadeaux" ou "famille". [127]*

Dans de tels cas, le lien sémantique entre l'UL produite et l'item visé est très ténu ou inexistant; nous verrons par contre que ces confusions peuvent souvent être

expliquées par d'autres types de liens lorsque nous nous intéresserons aux sources des erreurs. Nous tenons par ailleurs à souligner la présence dans notre corpus de certaines impropriétés si fréquentes qu'elles deviendront sans doute normées d'ici quelques années. L'emploi des expressions SUITE À et AU NIVEAU DE sont de bons exemples de ces erreurs découlant de la transgression d'une règle « de correction » imposée par la norme prescriptive, que James (1998) nomme *solecisms* et qui, à notre avis, ne témoignent pas réellement de lacunes lexicales chez la personne qui les utilise, mais simplement de la méconnaissance de certaines prescriptions normatives.

Les anglicismes constituent une sous-classe d'impropriétés et sont surtout présents dans notre corpus sous la forme d'anglicismes sémantiques comme BATTERIE au sens de 'pile' [48], FUTUR au sens d' 'avenir' [297], DÉFINITIVEMENT au sens d' 'assurément' [275], OPPORTUNITÉ au sens d' 'occasion' [184] et ACADÉMIQUE au sens de 'scolaire' [234]. Nous retrouvons aussi certains calques, comme PRENDRE POUR ACQUIS [257], PÂTE À DENTS [136] et BIEN-ÊTRE SOCIAL [224], ainsi que des anglicismes formels tels que COACHER_v [49], STAR [86], CHIFFRE¹¹⁶ [de travail] [270] et FAN au sens d' 'admiratrice' [282]. Ces anglicismes sont tout à fait intégrés à la langue, ce qui nous laisse croire que les élèves les ayant produits ne sont pas passés par l'anglais, mais ont simplement utilisé un item lexical de leur vocabulaire sans être conscients de la position de la norme prescriptive relativement à ces usages ou simplement parce qu'ils n'avaient pas accès à un terme équivalent du français standard¹¹⁷.

2.2.1.3 LE NON-RESPECT DES CONTRAINTES PESANT SUR LES ACTANTS SÉMANTIQUES

Le non-respect des contraintes pesant sur les actants sémantiques occasionne un peu plus d'une erreur sémantique sur dix. Au-delà de la traditionnelle confusion entre les verbes APPORTER, EMPORTER, AMENER et EMMENER [56, 92, 110, 120] et l'emploi de

¹¹⁶ Déformation de SHIFT [de travail].

¹¹⁷ On remarque d'ailleurs qu'une élève semblait consciente du caractère un peu problématique de SEXY – ce terme étant considéré comme un anglicisme au Québec – dans un contexte de rédaction, puisqu'elle met ce mot-forme entre guillemets dans son texte.

[aller] PORTER au sens de ‘reconduire’ [238], nous trouvons plusieurs cas intéressants, dont voici quelques exemples.

- (15) *L'arbre s'affala [tomba] sur le ruisseau. [8]*
 (16) *Au sommet de la montagne se tenait [se trouvait] la boule de feu magique qui remettrait vie à son village. [78]*
 (17) *Toutpetit empoigna [saisit] la poudre de Papif et en souffla dans la pièce puis les insectes devinrent de joli papillons. [74]*

Dans ces trois cas, c’est la cooccurrence du verbe avec son sujet ou son complément qui permet de constater que l’élève ne maîtrise pas tout à fait les contraintes sémantiques qui pèsent sur les actants de ces verbes. Dans le cas d’AFFALER, le type d’actant sémantique pouvant remplir le rôle de sujet est clairement décrit dans *Antidote RX*, où on définit ce verbe comme ‘se laisser tomber lourdement, s’affaisser, **en parlant d’une personne**’. Pour les exemples (16) et (17) par contre, il n’est pas aussi aisé d’expliquer l’erreur, puisque les contraintes pesant sur les actants sémantiques des verbes en jeu ne sont pas énoncées explicitement dans les ouvrages. Devant une définition comme ‘se trouver dans tel endroit, être quelque part’ pour le verbe SE TENIR, difficile en effet de justifier que ce verbe ne peut avoir pour sujet une UL dénotant un être inanimé comme une plante, une lampe ou une boule de feu; c’est plutôt notre intuition de locuteur (appuyée par l’observation d’exemples d’utilisation), qui nous permet de trouver la phrase (16) maladroite. Même chose pour EMPOIGNER, défini ainsi par *Le Robert* : ‘prendre en serrant dans la main’. Comment expliquer qu’on ne puisse dire **empoigner de la poudre*, comme il aurait été étrange de dire **empoigner une feuille* ou *une ficelle*? Sans doute existe-t-il certaines contraintes sur les UL pouvant jouer le rôle de complément du verbe EMPOIGNER (rigidité? grosseur?), mais elles demeurent difficiles à cerner¹¹⁸, même si la plupart des locuteurs s’entendent pour dire qu’**empoigner de la poudre* est carrément bizarre. Les erreurs de typage des actants sont donc souvent difficiles à diagnostiquer, puisque le recours aux

¹¹⁸ Si on considère que la définition d’EMPOIGNER proposée par *Le Robert* est exacte, une décomposition sémantique du sens de SERRER, son genre prochain, nous permettrait peut-être d’éclairer le problème sémantique présent en (14).

ouvrages de référence n'aide pas toujours; on peut donc imaginer qu'un enseignant devant expliquer ce genre d'erreur puisse être un peu dépourvu.

2.2.1.4 LES PLÉONASMES

Le dernier type d'erreurs sémantiques est le pléonasme, qui témoigne en quelque sorte d'une méconnaissance du sens d'une UL, puisqu'un élément supplémentaire est utilisé dans le but d'exprimer une composante de sens déjà présente. Laissons de côté les cas typiques de doubles connecteurs (*comme par exemple* [70, 128, 231], *donc par conséquent* [66]) pour analyser d'autres cas intéressants.

(18) Ce **cyclope** n'avait qu'un seul oeil [cyclope] et était très brutal. [121]

(19) Il **adorait énormément** [adorait/aimait énormément] ce sport! [215]

(20) Le soleil, qui **éblouissait ses yeux** [l'éblouissait], éclairait également la route. [224]

Dans ces trois phrases, le sens du terme souligné englobe déjà celui de la seconde moitié du pléonasme, ce qui alourdit le texte sans apporter d'information nouvelle. Nous ne sommes pas certain, par contre, que ce type d'erreur révèle véritablement que l'élève l'ayant commise ne comprend pas le sens de l'UL « principale » du pléonasme. En fait, si l'élève ayant produit la phrase (18) a choisi le nom CYCLOPE, c'est sans doute parce qu'il savait que ce nom désigne un 'membre d'une race de géants possédant un œil unique au milieu du front', tout comme l'élève qui a utilisé ÉBLOUIR en (20) connaissait le sens de ce verbe. Difficile, donc, de comprendre pourquoi ces deux scripteurs ont ajouté un élément redondant à la suite de ces termes; serait-ce justement par désir de bien montrer qu'ils comprennent le sens de l'UL utilisée? S'agit-il simplement de maladresses de style? L'exemple (19) nous semble un peu différent, dans la mesure où le scripteur semblait ici vouloir intensifier le sens d'un verbe qui a déjà un sens superlatif par rapport au terme plus neutre AIMER. En effet, ce pléonasme pourrait en quelque sorte être analysé comme un problème de collocation : l'élève a tenté de produire un collocatif intensificateur pour ADORER, alors que ce verbe correspond déjà à une intensification d'AIMER¹¹⁹.

¹¹⁹ En terme de fonctions lexicales, nous dirions qu'ADORER est une valeur fusionnée de Magn(aimer).

Magn(aimer) = beaucoup < passionnément, à la folie, de tout son cœur // adorer.

IMPLICATIONS DIDACTIQUES

La forte proportion d'erreurs mettant en jeu une UL véhiculant un sens proche du sens visé nous porte à croire qu'il serait pertinent de travailler en classe sur des groupes d'UL liées par le sens. Comme il est plutôt difficile d'anticiper des confusions mettant en jeu des UL présentant un lien de sens ténu, il serait sans doute préférable de se concentrer en priorité sur des ensembles de termes synonymes. En effet, il est facile, à partir d'un dictionnaire de synonymes, voire d'un dictionnaire comme *Le Petit Robert*, de trouver un ensemble d'UL pour bâtir une activité d'analyse des nuances sémantiques qui les distinguent. Il ne s'agit pas de travailler sur un ensemble important d'UL, mais bien de n'en retenir que quelques-unes, qui risquent d'être utiles à l'élève, dont certaines qu'il connaît déjà et auxquelles il pourra rattacher les autres. Par exemple, on pourrait imaginer un travail sur REGARDER, VOIR, APERCEVOIR et SCRUTER; il peut en effet être intéressant d'inclure dans notre liste des hyperonymes qui permettent de prendre conscience du spectre de contextes plus large dans lesquels ils peuvent apparaître et de l'emploi plus contraint des hyponymes. Ce type d'activités pourrait favoriser l'acuité sémantique des élèves, élargir leur vocabulaire et leur démontrer qu'il est pertinent d'utiliser conjointement le dictionnaire des synonymes et le dictionnaire « régulier ». Ils comprendraient notamment qu'il faut tenir compte de l'acception qui nous intéresse avant de choisir un synonyme!

Il pourrait aussi être intéressant de faire avec les élèves un travail de réflexion sur les déictiques, qui posent souvent problème. Par exemple, comment exprime-t-on ICI, MAINTENANT OU DEMAIN lorsqu'on écrit un texte narratif au passé et que le narrateur est omniscient?

Les erreurs liées au typage des actants peuvent difficilement faire l'objet d'un travail en amont, parce qu'elles sont difficilement prévisibles. Par contre, devant ce type d'erreurs, il est possible de consulter le dictionnaire avec l'élève pour voir si les contraintes pesant sur les actants d'une UL y sont indiquées. Dans le cas contraire, on peut tenter d'expliquer ces contraintes en s'appuyant sur des exemples d'utilisation de l'UL.

Des listes d'anglicismes, d'impropriétés et de pléonasmes fréquents peuvent être distribuées aux élèves dans le but qu'ils évitent de les utiliser dans leurs productions écrites. Il est aussi possible de demander à ces derniers de consigner ces types de problèmes dans un aide-mémoire lexical qu'ils peuvent consulter lors de l'écriture. L'enseignant peut aussi, à partir des erreurs relevées lors de la correction de textes, bâtir des activités sur ces types d'erreurs et demander aux élèves de trouver les formulations correctes correspondantes. Dans le cas de pléonasmes, notamment, il est généralement facile de justifier à l'aide de la définition du terme fort de l'expression en quoi le second terme est superflu.

Dans le cas où la rédaction ne fait pas l'objet d'un travail subséquent de correction de la part de l'élève, nous jugeons que l'enseignant devrait systématiquement fournir

Les deux barres obliques indiquent ici que la valeur qui les suit inclut à la fois le sens de la base de la collocation et celui du collocatif; c'est pourquoi on utilise le terme de *valeur fusionnée*.

une reformulation appropriée pour toute erreur sémantique; en indiquant à l'élève l'UL correcte à utiliser dans un contexte précis, l'enseignant expose l'élève à une occurrence de cette UL, contribuant ainsi, ne serait-ce que minimalement, à l'accroissement de son vocabulaire. Si tant est que l'élève lise la correction proposée, bien entendu...

2.2.2 LES ERREURS FORMELLES

Parmi les erreurs formelles, auxquelles ont été soustraites les erreurs d'orthographe et de flexion, les barbarismes (F-Barb) représentent l'écrasante majorité (86,7 % = 209 cas), suivi par une vingtaine de formes inexistantes (F-Inex) en français (8,7 % = 21 cas), puis de onze formulations analytiques (F-Déf = 4,6 %). La figure qui suit illustre cette répartition des erreurs formelles.

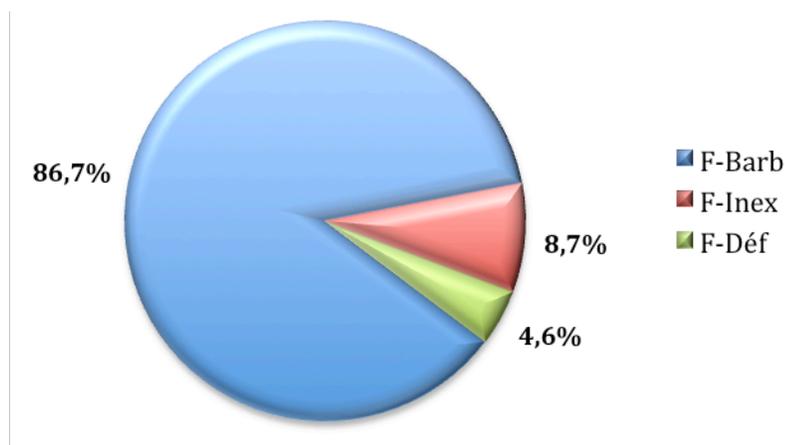


Figure 9. Répartition des 241 erreurs formelles (sans orthographe et flexion)

2.2.2.1 LES BARBARISMES

Les barbarismes consistent en une altération plus ou moins importante des formes associées à une UL. Plusieurs cas de figure ont été observés.

La confusion entre phonèmes est le phénomène le plus fréquent et conduit à produire de nombreux barbarismes comme **suptile* [84] pour *subtil*, **tournade* [170] pour *tornade*, **emponpoint* [190] pour *embonpoint*, **tribucha* [197] pour *trébucha*, **mignature* [89] pour *miniature* et **manufacture* [197] pour *manufacture*. On constate que la confusion intervient souvent entre une consonne sourde et la consonne sonore correspondante (ex. /p/ pour /b/) ou entre des voyelles proches dans le triangle vocalique (ex. /i/ pour /y/ ou /e/). L'omission d'un phonème est aussi un

phénomène assez fréquent (**acquiesça* [187], **ostination* [209], **abrute* [31]) et apparaît de façon récurrente auprès des organisateurs textuels **deuxièmement* [39, 47, 86, 126, 137, 157], **troisièmement* [47, 86] et **quatrièmement* [86]. Les chutes de liquides en fin de mot, très présentes à l'oral, se manifestent aussi à l'écrit dans des cas comme **monste* [111] pour *monstre* et **perde* [23] pour *perdre*. Plus rarement, on constate l'ajout d'un phonème (**ridiculeusement* [134] pour *ridiculement*, **tornitruant* [265] pour *tonitruant*, **minutres* [76] pour *minutes*, et **désambuler* [84] pour *déambuler*) ou l'inversion de phonèmes (**dédaigna* [111] pour *dégaina* [son épée]¹²⁰). Dans quelques cas, c'est plutôt la correspondance graphème-phonème qui ne semble pas maîtrisée : **déga* [11] pour *déjà* et **altèse* [15] pour *altesse*.

Parmi les barbarismes, nous avons aussi considéré les erreurs liées à l'emploi des verbes pronominaux. Dans les phrases suivantes, par exemple, le verbe aurait dû être employé à la forme pronominale.

(21) *Le moteur **arrêta** [s'arrêta] subitement.* [298]

(22) *Très heureuse que ses amies aient **embarqué dans** [se soient embarquées dans] son projet, Caroline commença immédiatement la recherche d'un répertoire.* [281]

(23) *L'horloger **aperçut** [s'aperçut/constata] que les gens vieillissaient très rapidement.* [88]

C'est trois cas sont très représentatifs des problèmes mettant en jeu des verbes pronominaux. La plupart du temps, ces verbes existent aussi à la forme non pronominale, mais sont alors des verbes transitifs directs qui, sémantiquement, peuvent être considérés comme des causatifs de leur correspondant pronominal. Par exemple, le verbe REFLÉTER signifie 'renvoyer la lumière, la couleur, l'image d'une chose de façon atténuée et plus ou moins vague', alors que SE REFLÉTER signifie 'être renvoyé de façon atténuée et plus ou moins vague, en parlant d'une image'; il serait donc possible de considérer que REFLÉTER est un causatif de SE REFLÉTER, puisque sa définition pourrait être 'causer que la lumière, la couleur, l'image d'une chose se reflète'. Cette nuance sémantique entre les deux formes de ce type de verbes, assez subtile il va sans dire, donne lieu à de nombreuses erreurs comme celles présentes dans les phrases (21) et (22). Le cas de la phrase (23) est tout à fait

¹²⁰ Dans ce cas-ci, la forme altérée correspond à celle d'une autre UL : il s'agit de deux verbes paronymes.

différent, puisque le verbe *APERCEVOIR*, au sens de ‘se rendre compte’, n’existe qu’à la forme pronominale. Nous avons aussi rencontré des emplois pronominaux qui, bien qu’ils aillent à l’encontre de ce que l’on retrouve dans les ouvrages, nous paraissaient tout à fait corrects à première vue : **se mériter* [126] au sens de ‘remporter’ (ex. *Le gagnant du concours *se méritera un weekend à New-York.*) et **se dégoter* [270] au sens de ‘découvrir’¹²¹.

Nous avons jusqu’à présent analysé les déformations du signifiant qui affectent des lexèmes, mais plus du tiers des barbarismes de notre corpus concernent des locutions. Voici quelques cas, parmi les plus frappants.

- (24) *S’il ne buvait pas de l’eau à temps, il allait tomber **sur la connaissance** [sans connaissance]. [18]*
- (25) *... l’un de vos chevaliers l’affrontera **au périple de sa vie** [au péril de sa vie]. [60]*
- (26) *Sinon, elle (= sa mère) se ferait un plaisir de lui **couper les finances** [couper les vivres]. [261]*
- (27) ***Peu d’instant** après [Peu de temps après, quelques instants plus tard], Eric sortit sa cantine de son sac pour ensuite pouvoir dîner. [177]*

Dans les trois cas présentés, l’un des éléments de la locution a été remplacé par un autre, ce qui va à l’encontre du caractère figé de ce type d’UL, même si certaines locutions sont plus malléables que d’autres. Ce figement est particulièrement évident en (26), puisque *FINANCES* et *VIVRES* sont pourtant presque synonymes dans ce contexte, mais que seul *VIVRES* peut apparaître dans l’expression; un autre exemple du même type est l’erreur de la phrase (27).

Il existe d’autres cas, comme les suivants, où le figement de l’expression est transgressé de manière plus subtile, ce qui produit malgré tout des phrases douteuses.

- (28) *Ils ont **fait une pierre deux coups** [fait d’une pierre deux coups], car ça sensibilise les conducteurs présents sur les routes, mais aussi les adolescents. [139]*
- (29) *Quand l’école recommença, j’**eus du plaisir fou** [eus un plaisir fou] à raconter à toutes mes amies ce que j’avais vécu. [230]*
- (30) *Je les (= les canards) nettoyé tout doucement avec **un coeur gros** [le coeur gros] tellement ils faisaient pitié. [230]*
- (31) *... elle venait littéralement de **me boucher un coin** [m’en boucher un coin]. [284]*

¹²¹ L’emploi de ce verbe est doublement problématique, puisqu’il appartient en plus au registre familier.

Au-delà du figement, nous avons vu que la locution se caractérise souvent par une forte cohésion interne, caractéristique parfois en jeu dans certaines erreurs lexicales de notre corpus. Par exemple, dans les phrases suivantes, l'élève a tenté d'intensifier le sens de la locution utilisée, ce qui est impossible, du moins avec ces expressions.

- (32) *Juste **par gros malheur** [par malheur], William a complètement oublié le rendez-vous. [49]*
- (33) ***Jamais, au plus grand jamais** [Jamais, au grand jamais], elle parviendront à s'entendre durant toute l'été. [265]*

Nous terminerons notre présentation des barbarismes en présentant deux erreurs récurrentes que nous avons choisi de placer dans cette classe. Tout d'abord, la confusion entre les locutions ENTENDRE DIRE et ENTENDRE PARLER, qui, bien qu'elles ne soient pas si éloignées sémantiquement, contrôlent des structures syntaxiques différentes et ne peuvent être utilisées indifféremment, comme dans la phrase suivante.

- (34) *J'ai entendu **parler** [dire] que vous faisiez un reportage sur la publicité. [147]*

L'ajout abusif de la préposition À devant les expressions marquant la périodicité, tournure très fréquente à l'oral, est aussi un problème que nous avons observé à plusieurs reprises : *à *chaques matins* [19], *à *chaque jour* [47, 241, 265], *à *chaque fois* [57, 64, 146, 194, 245, 269], *à *chaque nuits* [14], *à *chaque mercredi* [229], *à *chaque année* [268], *à *toutes les fois* [118], *à *tous les jours* [208, 227], *à *tous les Vendredis* [180]. Dans le même ordre d'idées, un élève a écrit *à *quelque part* [182].

2.2.2.2 LES CAS D'UTILISATION D'UNE UL FICTIVE

Le second type d'erreurs formelles de notre corpus est constitué des formes fictives, c'est-à-dire des UL inexistantes. Dans la plupart des cas rencontrés, il est aisé de comprendre le sens de l'UL inexistante, puisque les mots-formes « inventés » par les élèves partagent souvent une grande parenté formelle avec une UL existante. Voici des exemples de ces formes fictives sémantiquement « transparentes » : **apeurant* [57] au sens d' 'inquiétant', **sombré* [78] au sens de 'plongé dans le noir', **exagèrement* [134] au sens d' 'exagération', **incompréhensible* [171] au sens d' 'incompréhensible', **infinissable* [280] au sens d' 'interminable', **ennuyée* [282] au

sens de ‘qui s’ennuie’, **mornement* [265] au sens d’ ‘avec morosité’, **directives* [268] au sens de ‘directions’. Certains cas, aussi facilement compréhensibles, étaient tout simplement des termes empruntés directement à l’anglais, que l’on aurait pu classer parmi les anglicismes : **dunk* [86] au sens de ‘smash’, et **drum* [290] au sens de ‘tambour’. Notons aussi l’emploi de SAUVETEURE pour ‘sauveteuse, surveillante de piscine ou maitre-nageuse’, qui n’apparaît dans aucun des ouvrages de référence retenus, mais qui est pourtant très fréquent sur Internet... et semble être le terme privilégié par les principales intéressées! Au-delà de ces UL fictives parfaitement compréhensibles, quelques cas nous ont laissé plus perplexe, d’autant plus que le contexte nous permettait difficilement d’en inférer le sens. En voici deux exemples.

(35) *Sorter moi de là, c’est **oublignant** [???*] [30]

(36) *... l’animal avait été relâché. Donc, son coup **fit chûte** [???*] [203]

Soulignons finalement, parmi les lexies inexistantes, l’emploi de la locution FAIRE LE TITANIC [203], qui signifie sans doute ‘se tenir debout à la proue d’un bateau, suspendu au-dessus de l’eau les cheveux au vent et les bras étendus de chaque côté, dans le but de séduire la femme de ses rêves’. Nul besoin de spécifier que la compréhension de cette expression nécessite certains référents culturels!

2.2.2.3 LES CAS D’UTILISATION D’UNE FORME ANALYTIQUE

Le dernier type d’erreurs formelles correspond à l’utilisation d’une forme analytique plutôt que d’une UL précise, ce qui alourdit inutilement le texte et dénote selon nous une lacune dans le vocabulaire de l’élève. Nous rappelons qu’il ne s’agissait pas de relever l’ensemble des périphrases ou constructions maladroites, mais bien celles qui auraient aisément pu être remplacées par un seul terme en rendant le texte plus élégant, par exemple **bicyclette à trois roues* [265] au lieu de TRICYCLE. Voici quatre des onze cas relevés.

(37) *Alors il prit les saucisses empoisonnées d’un poison **qui fait dormir** [somnifère].* [104]

(38) *... elle n’était pas très sur de ce qu’elle devait faire pour transformer Alexandre en **homme sirene** [triton]* [174]

(39) *Plus tard, on voit un autre groupe de filles, **à une autre location** [ailleurs], courrir vers la même place que les autres.* [157]

(40) *... la tondeuse, qui se trouvait dans un entrepôt avec le **coupeur de gazon** [taille-bordure] et le souffleur.* [292]

Parmi ces problèmes, nous avons aussi observé des cas où l'élève utilise une négation pour exprimer le sens contraire d'une UL plutôt que son antonyme (**non connue* [157] pour 'inconnu' et **pas certaine* [175] pour 'incertaine').

IMPLICATIONS DIDACTIQUES

Sauf dans le cas de barbarismes très répandus, comme **aréoport*, il est plutôt difficile de prévenir les erreurs formelles. Par contre, il serait important que tout barbarisme soit rectifié par l'enseignant lors de la correction. Comme la forme constitue le moyen d'accès à l'information contenue dans le dictionnaire, nous jugeons qu'il serait inapproprié de demander à l'élève de chercher lui-même la réponse à un problème formel; l'élève risque de croire que le « mot » n'existe tout simplement pas s'il ne le trouve pas sous l'orthographe qu'il croit la bonne ou de se décourager s'il doit chercher à différents endroits dans le dictionnaire selon les orthographes possibles.

Dans le cas d'erreurs fréquentes, du type **à tous les jours* ou **deuxièmement*, il peut être pertinent de faire un retour en classe avec les élèves pour attirer leur attention sur la forme correcte; l'enseignant peut aussi se construire une liste de termes problématiques dont il fait part aux élèves.

Prévenir les erreurs relatives à l'emploi des locutions n'est pas non plus chose facile. Par contre, faire davantage de travail sur les expressions en classe contribuerait sans doute à augmenter le stock de locutions des élèves, à les encourager à les utiliser en rédaction et à leur montrer de quelle façon on les trouve dans le dictionnaire pour en vérifier le sens et la forme.

Des activités sur les préfixes et les suffixes peuvent aider les élèves dans leur compréhension du sens d'UL nouvelles rencontrées dans leurs lectures et contribuent à mieux cerner les liens sémantiques qui unissent les UL d'une même famille morphologique. Il est également important de faire prendre conscience aux élèves que la dérivation morphologique n'est pas très productive en français (du moins en synchronie) et qu'il est préférable de toujours vérifier l'existence d'un « mot nouveau » que l'on crée à l'aide d'un préfixe ou d'un suffixe. Utiliser lors de la rédaction des ouvrages présentant des familles de mots, le *Lexis* de Larousse par exemple, permettrait sans doute d'éviter des erreurs comme **exagèrément_N* ou **incomprénable*.

Dans les cas où l'élève produit une forme analytique au lieu d'utiliser une UL précise, il serait pertinent que l'enseignant lui suggère cette UL parce que l'élève ne sait peut-être même pas qu'un terme précis existe pour véhiculer le sens visé. Si l'élève a de la difficulté à nommer une chose concrète, on peut lui demander de trouver le terme approprié dans un dictionnaire encyclopédique de type *Le Petit Larousse*, qui contient plusieurs planches favorisant la nomination; les dictionnaires visuels sont encore plus efficaces en cette matière et il pourrait être intéressant d'en posséder un en classe.

2.2.3 LES ERREURS PRAGMATIQUES

Voici la façon dont se répartissent les erreurs liées à la combinatoire pragmatique.

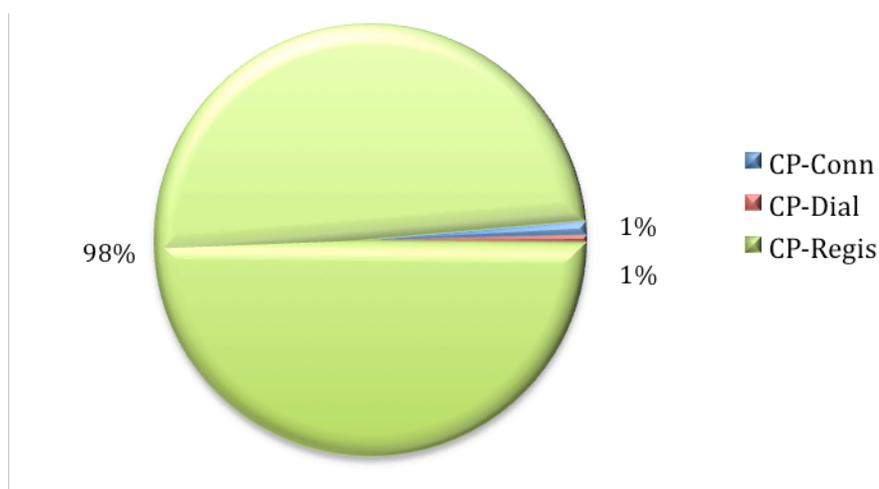


Figure 10. Répartition des 195 erreurs de combinatoire pragmatique

La quasi-totalité des erreurs liées à la combinatoire pragmatique (98,5 %) correspondent à des cas d'utilisation d'UL appartenant au registre familial (CP-Regis), ce qui est inapproprié dans une production écrite en contexte scolaire, du moins dans les textes analysés¹²².

2.2.3.1 LES CAS D'UTILISATION D'UN REGISTRE DE LANGUE INAPPROPRIÉ

Les erreurs de registre affectent toutes les classes de mots. En voici quelques exemples.

Noms : JOB [39, 43, 47, 230, 236, 245, 263, 266, 275, 279, 288], PUB¹²³ [53, 81, 126] au sens de 'publicité', GARS [29, 208, 271], CHUM [256], BLONDE [250], AFFAIRES [259] au sens d' 'effets personnels'¹²⁴, PLACE au sens d' 'endroit' [157, 158, 213, 245, 258, 264, 270].

¹²² Soulignons que, lorsque nous avons relevé les erreurs de registre dans les rédactions, nous avons tenu compte du contexte dans lequel les UL familiares apparaissaient. Ainsi, une UL familière utilisée par un personnage dans le contexte d'un dialogue n'était pas considérée comme une erreur, alors qu'elle l'aurait été en apparaissant dans un passage narratif.

¹²³ PUB et JOB sont parmi les UL familiares les plus fréquentes dans notre corpus. Cette surreprésentation est assurément liée aux thèmes de rédaction proposés aux élèves et au fait que ces UL familiares apparaissaient sans doute dans les textes de préparations consultés par les élèves, ce qui légitime en quelque sorte leur utilisation. D'ailleurs, nous avons remarqué que les enseignants ont peu relevé ou pénalisé ces « erreurs ».

¹²⁴ Nous avons relevé un deuxième emploi familier d'AFFAIRES, qui renvoie cette fois à un sens assez vague correspondant à 'choses' (ex. *Cela faisait beaucoup d'affaires [choses] à gérer en même temps.* [256]). Cet emploi, pourtant fréquent à l'oral, est assez mal décrit dans les ouvrages de référence.

Verbes : SE CHICANER [49], OBSTINER [42] au sens de ‘manifester son opposition’, LÂCHER [un emploi] [43, 44], QUITTER [42] au sens de ‘partir’ (emploi intransitif), DÉCAMPER [57] au sens de ‘s’enfuir’, PARTIR [à] au sens de ‘se mettre à’ [169, 173], VEDGER [251] au sens de ‘ne rien faire’.

Adverbes : TOUT D’UN COUP [48] au sens de ‘tout à coup’, POUR DE VRAI [132] au sens de ‘réellement’, PAREIL [48, 49] au sens de ‘quand même’ (ex. *Il n’était pas trop convaincu, mais James essaya pareil.*), DUR [20] au sens de ‘fort, avec intensité’ (ex. *La reine du diable dormait tellement dur...*).

Adjectifs : BÊTE [260] au sens d’ ‘antipathique’, TANNÉ [291] au sens de ‘qui en a assez’, BARAQUÉ [261], BOUSILLÉ [292].

Prépositions : À pour exprimer la possession [59, 61, 166] (ex. *la maman à Christopher*), CÔTÉ [300] au sens de ‘en matière de’ (ex. *Nous nous entendions bien côté goûts musicaux.*), QUESTION DE [251, 278] au sens de ‘afin de’ (ex. *... question de les sensibiliser davantage.*), APRÈS [251, 278] au sens de ‘à’, ‘sur’ ou ‘contre’ [21, 118] (ex. *... le colla après un poteau.*).

Déterminants : PLEIN DE [57, 61, 73, 111, 217, 245, 252, 254, 262, 266, 277] au sens de ‘beaucoup de, plusieurs’, UNE COUPLE DE [82] au sens de ‘quelques’ (ex. *Elle est âgée d’une couple d’années...*), GRAND [267] comme déterminant en contexte négatif (ex. *Je ne mis pas grand temps à regretter d’avoir voulu cet emploi.*).

Conjonctions : À CAUSE QUE [48, 134, 203, 266] au sens de ‘parce que’.

Si beaucoup des termes familiers relevés dans le corpus appartiennent au français québécois, on note tout de même la présence de plusieurs termes propres au français de France, comme BOULOT [12, 229, 288], BALÈZE [224], BOUSILLÉ [292] OU EN AVOIR MARRE [227]. Nous pouvons supposer que les élèves qui utilisent ces termes à « consonance européenne » ont l’impression d’utiliser un vocabulaire sophistiqué, alors qu’ils emploient pourtant un registre de langue familier. La même hypothèse s’applique aux cas où les élèves utilisent le masculin pour des termes familiers généralement féminins au Québec, notamment avec [un] JOB [39, 43, 230, 236, 245, 279, 288]; nous avons d’ailleurs rencontré un cas où l’enseignant pénalisait un élève pour avoir utilisé JOB au féminin [227] plutôt que pour le caractère familier de cette UL.

Dans certains cas, le verbe familier utilisé nous semble la façon la plus typique d’exprimer le sens visé en français québécois et serait sans doute le premier terme à nous être venu en tête.

- (41) *Etiv se tassa [bougea, se déplaça, se poussa] juste à temps pour ne pas être mangé. [95]*
- (42) *... (le) cyclope pila sur [marcha (sur), posa le pied (sur); écrasa] la main de l'un des pêcheurs. [197]*
- (43) *il poussa un énorme souffle qui (fesa) revolver [projeta] Bob. [69]*

En effet, dans les phrases (41) et (42), le terme familier semble correspondre plus exactement au sens visé que les termes standards que nous proposons comme correction. En (43), la correction PROJETER est correcte dans ce contexte syntaxique, mais cet exemple nous a permis de réaliser que le verbe familier REVOLER n'a pas d'équivalent dans la langue standard, qui nous oblige à utiliser une tournure passive pour exprimer ce sens (ex. *L'automobile explosa est des débris **revolèrent** [→ furent projetés] dans tous les sens.*)¹²⁵.

Si la plupart des termes familiers relevés sont des lexèmes, nous trouvons certaines locutions verbales familières comme NE PAS PRENDRE QQCHOSE [269] (au sens de 'ne pas accepter qqch'), NE PAS ÊTRE FORT SUR QQCHOSE [286] (au sens de 'ne pas apprécier qqch'), AVOIR DE LA MISÈRE [173], BOURRER LE CRÂNE DE QQN [86] et NE PAS ÊTRE UN CADEAU [239].

Un seul des problèmes de registres de langue de notre corpus ne correspond pas à l'emploi d'un terme familier, mais plutôt à l'utilisation d'une UL qui nous semble d'un niveau trop soutenu pour le style de texte dans lequel elle apparaît, plutôt malhabile. Il s'agit de l'adjectif OBLIGÉ [178] au sens de 'reconnaissant'.

- (44) *En faisant ceci, elle avait rendu un grand service aux sirènes. Celle-ci bien **obligé** [reconnaissante] allait lui demander! « Jeune fille qu'elle est ton plus grand souhait? » [178]*

Un dernier commentaire s'impose relativement aux problèmes de registres de langue. Les nombreux exemples présentés permettent de constater que toutes les UL familières utilisées ne sont pas perçues comme aussi « dérangeantes » les unes que les autres. En effet, nous croyons que la plupart des gens s'entendraient pour affirmer que l'emploi de termes comme [de la] VITRE [147, 254, 256] pour signifier '(du) verre', [des] SOUS pour exprimer '(de l') argent' ou PESER pour 'appuyer' est moins

¹²⁵ GICLER est en fait un synonyme de REVOLER, mais seulement dans le cas où le sujet du verbe dénote un liquide.

choquant que l'utilisation de [la] GANG [252, 284] pour signifier '(le) groupe', GLANDOUILLER [265] pour 'flâner' ou PARTIR SUR LE PARTY [210] pour exprimer 'faire la fête'. Pourtant, toutes ces UL reçoivent la marque d'usage *familier* dans les ouvrages, ce qui ne permet pas d'identifier où se situe chaque terme sur le continuum de « familiarité »; les enseignants doivent donc utiliser leur propre intuition afin de trancher sur ce type de problèmes, puisque l'on constate qu'ils pénalisent ou relèvent certains usages alors qu'ils en laissent passer d'autres sans les signaler, voire sans les remarquer. Peut-être d'ailleurs jugent-ils dans certains cas que l'emploi d'un terme familier est justifié parce qu'il fait image ou permet de rendre efficacement un message.

2.2.3.2 LES AUTRES ERREURS LIÉES À LA COMBINATOIRE PRAGMATIQUE

Comme l'indique la figure présentée au début de cette section, un très faible pourcentage des erreurs en lien avec la combinatoire pragmatique n'étaient pas liées aux registres de langue.

Tout d'abord, l'emploi de VIEUX_N [224, 240] pour 'vieillard' nous semble véhiculer une certaine connotation négative, ce que confirme *Le Petit Robert*, qui définit ainsi cette acception de VIEUX : '**un vieux, une vieille** : un vieil homme, une vieille femme (avec une valeur un peu méprisante ou condescendante)'

Le second problème de combinatoire pragmatique non lié au registre nous apparaît comme un dialectalisme.

(45) *Peu d'instant après, Eric sortit sa **cantine** [boite à lunch] de son sac pour ensuite pouvoir dîner. [177]*

L'emploi du terme CANTINE dans cette phrase nous a posé un problème de compréhension et nous avons eu recours au dictionnaire pour en vérifier le sens. Il n'y a que dans *Antidote RX* que nous avons trouvé l'acception CANTINE au sens de 'récipient à compartiments, en fer blanc ou en fer émaillé, utilisé pour transporter un repas'. Le dictionnaire mentionne que cette acception est propre à la Suisse et à l'Afrique, ce qui explique sans doute pourquoi ce sens ne nous est pas apparu facilement et nous laisse supposer que l'élève qui a produit la phrase (45) a été en contact avec un locuteur qui utilisait CANTINE de cette façon.

IMPLICATIONS DIDACTIQUES

Comme l'écrasante majorité des erreurs reliées à la combinatoire pragmatique sont attribuables à l'emploi de termes familiers, nous croyons qu'il est primordial d'aborder en classe la question des registres de langue. Il faut tout d'abord sensibiliser les élèves à l'existence de différents registres de langue et leur faire prendre conscience que beaucoup du vocabulaire que l'on emploie à l'oral n'est pas approprié dans un contexte de production écrite exigeant un niveau standard.

Il serait aussi important de bien établir la distinction, à l'aide d'exemples, entre registre familier et français québécois; en effet, le français de France a aussi son registre familier et tous les termes propres au français québécois ne sont pas nécessairement familiers.

Tout un travail pourrait aussi être fait sur la façon de trouver les informations relatives aux registres de langue et aux variations régionales dans les dictionnaires. On pourrait observer comment les dictionnaires fournissent ces informations, notamment à l'aide d'abréviations – comme *fam.* – qui sont souvent opaques pour les élèves. Il faut aussi faire prendre conscience à ces derniers que ce n'est pas parce qu'un mot « existe » dans le dictionnaire qu'il est nécessairement approprié dans tous les contextes.

Encore une fois, une liste de termes familiers utilisés fréquemment pourrait être consignée en classe et on pourrait demander aux élèves de trouver les UL standards correspondantes.

2.2.4 LES ERREURS LIÉES À LA COMBINATOIRE MORPHOSYNTAXIQUE

La combinatoire morphosyntaxique des mots explique 21 % de l'ensemble des erreurs analysées (11 % = 128 cas pour les propriétés lexicosyntaxiques et 10 % = 111 cas pour les propriétés lexicomorphologiques). Si nous choisissons de présenter dans la même section ces deux grandes classes d'erreurs lexicales, c'est qu'elles nous semblent intimement liées : elles constituent les deux classes de problèmes lexicaux à cheval sur la grammaire et le lexique et regroupent les sous-classes d'erreurs souvent comptabilisées par les enseignants parmi les problèmes de syntaxe et d'orthographe grammaticale. Attardons-nous d'abord sur les erreurs de combinatoire morphologique.

2.2.4.1 LES ERREURS LEXICOMORPHOLOGIQUES

La figure suivante illustre la répartition des erreurs liées à la combinatoire morphologique. Quarante-cinq et neuf dixièmes pour cent (= 51 cas) des erreurs lexicomorphologiques ont trait à l'invariabilité des unités lexicales (CM-Inv), alors que 42,3 % (= 47 cas) concernent le genre nominal (CM-Genr); les 11,7% d'erreurs restantes de cette classe correspondent à des cas de non-respect de la classe de mots (CM-Class), de mauvais choix d'auxiliaire (CM-Aux) ou à des cas particuliers (CM-Rés).

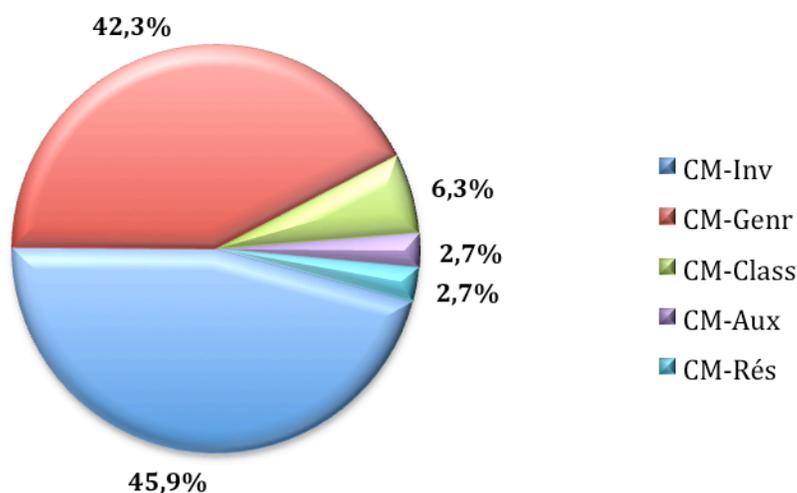


Figure 11. Répartition des 111 erreurs de combinatoire morphologique

A. LES CAS DE NON-RESPECT DE L'INVARIABILITÉ

Les erreurs liées à l'invariabilité de certaines UL concernent principalement des noms, comme dans les phrases suivantes.

- (46) ... des gens avec plus de 6 mois **d'expériences** [d'expérience] [46]
- (47) ... on peut ressentir les scènes **d'actions** [action] vécu du superhéro. [54]
- (48) Je suis présentement à New-York pour une fin de semaine de **plaisirs** [plaisir] avec mes camarades. [280]
- (49) ... mais peut après leur traversée de l'autre côté **des douanes** [de la douane], un Américain entra dans l'autobus armé d'un fusil. [255]
- (50) Si l'un parlait en **vacance** [vacances] en Floride, l'autre travaillait. [273]
- (51) Ils devaient penser à leurs transports, à trouver **de bons hébergements** [de l'hébergement convenable], à la nourriture... [213]

Les UL en jeu dans ce type d'erreur sont la plupart du temps des noms polysémiques dont seule une acception est invariable, comme dans les exemples

(46) à (49). Analysons l'erreur présente en (46). *EXPÉRIENCE*, dans le sens où l'élève voulait l'utiliser (*EXPÉRIENCE2*), signifie 'ensemble des connaissances acquises par une longue pratique' et est toujours employé au singulier; il s'agit en fait d'un nom non comptable. Or, cette acception est intimement liée à une autre du même vocable, qui cette fois peut varier en nombre : *EXPÉRIENCE1* au sens d' 'évènement vécu par une personne, susceptible de lui apporter un enseignement'. Il existe entre ces deux UL une relation de métonymie, *EXPÉRIENCE2* pouvant être considérée comme 'la somme des expériences₁ d'un individu'. Les cas (47) et (48) sont similaires, puisqu'un rapport métonymique similaire existe entre *ACTION1* ('ce que fait qqn'), variable, et *ACTION2* ('animation tenant aux faits et aux actes représentés ou racontés, aux actions₁ des personnages d'une œuvre narrative'), invariable, et entre *PLAISIR1* ('état affectif plus ou moins passager, lié à une sensation agréable ou à la satisfaction d'un besoin, d'un désir'), invariable, et *PLAISIR2* ('objet, action qui est la cause de l'état de plaisir₁'), variable. Un rapport métonymique existe aussi entre *DOUANE1* ('administration chargée de percevoir les droits sur les marchandises lors des importations et des exportations') et *DOUANE2* ('Lieu où la douane₁ est établie'), mais le *Multidictionnaire* mentionne qu'il est fautif d'utiliser la deuxième acception au pluriel, ce qui se reflète dans les exemples d'utilisation proposés par les autres ouvrages.

Dans d'autres cas, l'erreur met en jeu un nom dont toutes les acceptions sont invariables. Par exemple, tous les ouvrages mentionnent que *VACANCES* s'utilise seulement au pluriel¹²⁶ dans toutes ses acceptions liées à la notion de 'congé', ce qui explique l'erreur dans la phrase (50). Pour *HÉBERGEMENT*, l'information n'est pas aussi claire, mais l'absence complète d'exemples d'emploi au pluriel dans les ouvrages et notre propre intuition de locuteur nous permettent d'affirmer qu'il est étrange d'employer ce terme au pluriel et que la phrase (51) serait plus conforme à la norme si l'élève avait employé ce substantif au singulier.

Un dernier cas intéressant relatif à l'invariabilité nominale mérite d'être discuté, d'autant plus qu'il s'agit de l'une des erreurs que nous avons choisi de présenter

¹²⁶ On emploie le terme *plurale tantum* pour désigner ce type de noms.

aux enseignants lors des entrevues menées en lien avec notre deuxième objectif de recherche (voir IX-2.3.2).

(52) *Un premier job d'été peut s'avérer un très bon tournant pour tes autres **travails** [emplois] dans le futur. [39]*

Dans cette phrase, l'élève utilise TRAVAIL au sens d' 'emploi', ce qui est tout à fait correct. Seulement, il semble que TRAVAIL en ce sens n'existe qu'à la forme singulier. En effet, le pluriel de TRAVAIL est généralement *travaux*, mais une phrase comme **Il a eu plusieurs travaux au cours des dernières années.* ne semble pas fonctionner; c'est sans doute la raison pour laquelle l'élève a privilégié la forme *travails*¹²⁷. Encore une fois, c'est notre propre intuition de locuteur qui nous indique que cette acception de TRAVAIL est invariable; aucun des ouvrages consultés ne mentionne quoi que ce soit à cet effet, mais encore une fois, tous les exemples fournis sont au singulier. Soulignons toutefois que les dictionnaires répertorient deux acceptions de travaux toujours pluriel, qui signifient 'ensemble d'entreprises nécessitant l'activité physique de personnes et l'emploi de techniques particulières' (ex. *Les travaux d'entretien des routes*) et 'délibérations d'une assemblée' (ex. *Les travaux seront suspendus pour la période estivale.*).

Pour ce qui est des cas de non-respect de l'invariabilité affectant les autres classes de mots, nous retrouvons quelques adverbes qui reçoivent la marque du pluriel (**tards* [50], **à chevaux* [162] , **ensembles* [168]) ou du féminin (**une publicité forte attrayante* [148]) ainsi que quatre cas où l'adverbe TOUT a reçu la marque du féminin alors qu'il devait rester invariable, puisqu'il précédait un adjectif commençant par une voyelle [249, 265, 280, 282]. Les déterminants et les pronoms qui posent problème sont ceux qui peuvent renvoyer à plusieurs référents, même en contexte singulier : les déterminants singuliers (**chaques élèves rentra chez eux* [48], **chaques personnes vieillissaient* [88], **aucunes rides* [82], **Toutes personnes qui voulurent y entrée y restarent pour toujours.* [88]) et numéraux (**quatre* [38], **cinqs* [94], **milles* [287]) ainsi que les pronoms indéfinis (**pour qu'elle puisse avoir quelques choses à se mettre sous la*

¹²⁷ On trouve d'ailleurs une remarque sur ce problème précis dans *Le bon usage* (Grevisse et Goose, 2008: 679), où l'on mentionne deux cas à l'oral où c'est la forme *travails* qui est utilisée.

dent. [201]) et négatifs (**plus personnes souffrirent de l'incroyable chaleur.* [95]). La seule préposition touchée par ce type d'erreur est AU PIED DE, qui a été utilisée deux fois au pluriel (**un joli arc-en-ciel avec à ses pieds un énorme chaudron d'or.* [182], ***Aux pieds du volcan Asama** [187]) et le seul adjectif en jeu est un adjectif de couleur invariable (**ses beaux yeux marons* [193]). Nous n'avons recensé aucun problème lié aux verbes défectifs.

B. LES CAS DE NON-RESPECT DU GENRE NOMINAL

Les erreurs liées au genre nominal, comme nous pouvions nous y attendre, affectent en très grande majorité (42 cas sur 47) des noms commençant par une voyelle (ex. * [une] *endroit* [98], * [une] *autobus* [256], * [une] *éclair* [79, 268], * [une] *immeuble* [126], * [un] *allure* [23]) ou un *h* muet (ex. * [un] *heure* [66, 91], * [une] *habit* [120], * [une] *historique*_N [211]); par contre, nous remarquons qu'il n'est pas nécessaire que la liaison s'entende pour qu'un nom masculin soit utilisé au féminin, puisque nous avons observé des cas où un adjectif faisait écran et « bloquait » la liaison, sans pour autant empêcher l'erreur (ex. **une gigantesque éclair* [285], **un aussi grand étendue d'eau* [185]). Les deux erreurs les plus fréquentes – encore une fois à cause des thèmes de rédaction – sont ÉTÉ, utilisé au féminin par 14 élèves, et ARGENT, que sept élèves employaient systématiquement comme un nom féminin. Les quatre cas concernant des noms commençant par une consonne sont les suivants : **la palet* [149] au sens de 'rondelle de hockey', **une "toast"* [260] au sens de 'tranche de pain grillée'¹²⁸, **son troisième secondaire* [213] et **mon cinquième secondaire* [296] au sens de 'X^e année du secondaire' et **une pétale* [169].

La plupart du temps, l'erreur se révèle par l'emploi d'un déterminant qui trahit la méconnaissance de l'élève relativement au genre du nom, comme dans les quatre derniers exemples présentés. Dans certains cas, c'est plutôt l'accord fautif de l'adjectif dans le groupe du nom qui nous permet de constater l'erreur (ex. **l'argent amassée* [146], **de petite job* [245], **l'historique professionnelle* [211]). Plus rarement, il faut

¹²⁸ Notons tout de même qu'*Antidote*, contrairement aux deux autres ouvrages, ne mentionne pas de genre pour ce nom et spécifie que ce mot s'emploie souvent au féminin au Québec. Nous aurions tendance à admettre cet emploi.

regarder dans le groupe prédicat pour constater le problème (phrases (50) et (51)), voire même s'appuyer sur les pronoms de reprises (phrases (52) et (53)).

(53) ... leur âme sont **noir** [noire] comme une nuit sans lune. [15] ¹²⁹

(54) Il avait compris combien l'argent pouvait être **précieuse** [précieux]. [266]

(55) Gagner de l'argent, c'est bien, mais il faut apprendre à **la** [le] gérer. [270]

(56) Nous pesions sur le bouton de l'ascenseur, mais **elle** [il] ne fonctionnait plus. [289]

C. LES AUTRES ERREURS LEXICOMORPHOLOGIQUES

Dans les erreurs de non-respect de la classe de mots, nous observons l'emploi adverbial des adjectifs MEILLEUR (phrase (57)) et DIRECT (phrase (58)) et l'emploi adjectival du nom ENTHOUSIASME (phrase (59)).

(57) C'est comme [si] je jouerais comme Lebron James ou **meilleur** [mieux] encore. [86]

(58) ... un morceau de bois très pointu et le planta **direct** [directement] dans le cœur. [69]

(59) J'étais très **enthousiasme** [enthousiaste] à l'idée de m'envoler vers les Alpes-Maritimes. [209]

Nous avons aussi observé deux utilisations de la préposition À LA SUITE (phrases (60) et (61)) à la place de l'adverbe ENSUITE, ce qui ne fonctionne pas : même si ces deux UL ont pratiquement le même sens, la préposition exige la présence d'un complément.

(60) Il était favorable à lui demander à **la suite** [ensuite], car comme ça il avait une bonne raison de les aider. [14]

(61) **À la suite** [Ensuite] il enfourcha son fidèle destrier. [66]

Les trois cas d'utilisation d'un mauvais auxiliaire considérés comme des erreurs lexicales dans notre corpus concernent le verbe TOMBER employé de façon intransitive avec l'auxiliaire AVOIR plutôt qu'ÊTRE, soit pour signifier 'être entraîné de haut en bas par son propre poids' (ex. *tout à coup il avait [était] tombé dans un trou.* [24]), soit comme collocatif du nom NUIT (*La nuit avait [était] déjà tombée.* [4]).

¹²⁹ Voici un cas où il n'est pas évident que l'erreur est vraiment lexicale. En effet, vu l'absence de flexion au pluriel du nom, du déterminant et de l'adjectif, nous pourrions ici poser l'hypothèse que l'élève n'applique aucune règle d'accord et que l'absence de marque du féminin sur l'adjectif n'est par conséquent pas un indice de la méconnaissance du genre du nom. Ceci est d'autant plus probable que, d'un point de vue purement phonologique, aucune différence perceptible n'existe entre les formes fléchies de l'adjectif NOIR.

Trois cas classés parmi les erreurs lexicomorphologiques ne correspondent pas aux sous-classes de notre typologie et ont été classés dans une classe résiduelle; ces trois problèmes sont de nature différente.

(62) *Il s'annonça une grosse tempête inattendue [s'annonça]. [48]*

(63) *William, notre joueur **d'hockey** [de hockey] dans l'équipe des Tigres... [49]*

(64) *Il rencontra un chat qui lui disa qu'elle habitait la maison après le **géant** lac [géant]. [59]*

Dans le premier cas, il nous semble très bizarre d'utiliser le verbe S'ANNONCER de façon impersonnelle, ce qui nous fait dire que certains verbes présentent peut-être des contraintes quant à la possibilité d'apparaître dans des constructions impersonnelles, contraintes qui ne sont jamais, à notre connaissance, indiquées dans les ouvrages de référence. Par contre, il serait intéressant de soumettre cette erreur à d'autres locuteurs pour voir s'ils portent le même jugement que nous sur cette phrase.

La phrase (63) présente une erreur relative à l'élision de la préposition DE devant le nom HOCKEY, qui découle du fait que ce nom commence par un *h* aspiré, ce qui bloque l'élision. Nous aurions pu choisir de classer ce problème parmi les problèmes formels, mais comme il ne se manifeste que dans des contextes syntaxiques précis, nous avons choisi de le considérer ici. Il s'agit selon nous d'une erreur lexicale, puisque même si la règle phonotactique régissant l'élision de DE – la préposition DE devient *d'* devant un mot commençant par une voyelle ou un *h* muet – est tout à fait régulière, le fait de savoir que HOCKEY commence par un *h* aspiré est une connaissance lexicale et fait partie de la combinatoire de cette UL.

L'erreur de la phrase (64), finalement, est en lien avec la difficile question du positionnement de l'adjectif en français¹³⁰. Comme il n'existe pas en cette matière de règle stricte permettant de déterminer si un adjectif est « préposé » ou « postposé » par rapport au nom qu'il complète, il s'agit d'une connaissance qu'il

¹³⁰ Comme le soulignent Abeillé et Godard (1999: 10), « [la] position relative de l'adjectif et du nom dans le GN constitue l'un des problèmes les plus étudiés de la grammaire du français ». Pour un survol de la question, voir notamment Delomier (1980).

faut acquérir pour chaque adjectif¹³¹, bien que cette connaissance dépende presque toujours d'un apprentissage implicite qui repose en grande partie sur la lecture, ou sur toute exposition de l'adjectif en question dans des contextes d'utilisation. S'il demeure difficile d'expliquer pourquoi GÉANT est un adjectif plutôt postposé, il ne fait aucun doute que tout locuteur du français trouverait bizarre la phrase (64).

IMPLICATIONS DIDACTIQUES

Les interventions relatives aux erreurs en lien avec l'invariabilité sont plutôt limitées, puisque les ouvrages de référence sont généralement peu aidants en cette matière. En effet, hormis quelques UL décrites comme toujours plurielles (ex. FUNÉRAILLES, FIANÇAILLES, OBSÈQUES, MŒURS) et d'autres comme inusitées au singulier (ex. VIVRES, PRÉPARATIFS, FÉLICITATIONS), les dictionnaires donnent peu d'informations sur le caractère invariable d'acceptions spécifiques de certains vocables. Il faut tout de même apprendre aux élèves à retracer l'information relative à l'invariabilité dans les articles de dictionnaire, même si toutes les erreurs ne pourront être évitées à l'aide des ouvrages.

Pour le genre nominal, par contre, il est très important de savoir récupérer l'information dans les articles de dictionnaire. Une fois de plus, il est important de démystifier avec les élèves les abréviations utilisées par les lexicographes pour mentionner la classe de mots et le genre nominal; ces informations en apparence évidentes peuvent passer inaperçues pour les élèves qui n'y ont pas été sensibilisés.

Finalement, on peut une fois de plus fournir aux élèves une liste de noms au genre « difficile », principalement des noms commençant par une voyelle. D'ailleurs, le simple fait de savoir qu'il faut se méfier des noms commençant par une voyelle pourrait amener les élèves à douter stratégiquement du genre de ces derniers.

¹³¹ Les dictionnaires construits selon les principes de la *LEC* notent d'ailleurs cette information dans l'entrée des adjectifs chaque fois qu'elle est pertinente.

2.2.4.2 LES ERREURS LEXICOSYNTAXIQUES

Du côté lexicosyntaxique, l'utilisation d'une mauvaise préposition régie (CS-ChPr) représente 38,3 % (= 49 cas) des erreurs, tandis que la construction erronée d'un complément (CS-Const) ou l'absence d'un complément obligatoire (CS-AbsC) totalisent respectivement 23,4 % (= 30 cas) et 18 % (= 23 cas) des problèmes relevés. Vingt et trois dixièmes pour cent (= 26 cas) sont finalement liés au non-respect du caractère transitif ou intransitif des verbes (CS-Trans/CS-AbsPr/CS-Intrans).

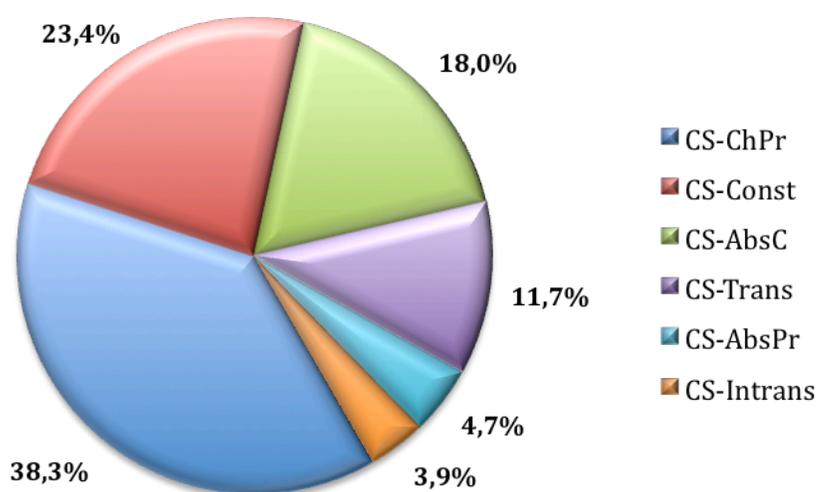


Figure 12. Répartition des 128 erreurs de combinatoire syntaxique

A. L'UTILISATION D'UNE MAUVAISE PRÉPOSITION RÉGIE

Le choix d'une mauvaise préposition auprès d'une UL contrôlant une structure syntaxique spécifique est le cas le plus fréquent d'erreurs liées à la combinatoire syntaxique. Ces erreurs touchent principalement les verbes (28 erreurs sur 49), comme dans les phrases suivantes :

- (65) *Le chef s'empressa à [d'] aller chercher cette formule dans une gigantesque bibliothèque. [18]*
- (66) *En plus, ils réussissent souvent de [à] nous convaincre. [55]*
- (67) *... l'horloger se dirigea à une vitesse éclair à [vers] la caverne. [88]*
- (68) *Pour clouer le bec du [au] public, Alex Kovalev lance un slogan en français. [151]*

Comme dans les exemples (65), (66) et (67), beaucoup des problèmes observés mettent en jeu les prépositions À et DE, prépositions par excellence pour introduire les compléments indirects des verbes, mais aussi difficilement prévisibles parce que vides d'un point de vue sémantique. D'ailleurs, si l'on observe la structure régie par le verbe S'EMPRESSER de la phrase (65) – obligatoirement construit avec la préposition DE – et on la compare à celle de son synonyme SE DÉPÊCHER – qui admet tant DE que À –, on saisit tout l'arbitraire relatif au choix de ces prépositions. Malgré cette difficulté de justifier pourquoi telle préposition plutôt que telle autre fonctionne auprès d'un verbe, les locuteurs natifs ont de très fortes intuitions en cette matière, et c'est pourquoi certaines erreurs nous ont paru plutôt surprenantes de la part d'élèves ayant le français comme langue première, par exemple celle de la phrase (66); il faut tout de même mentionner que dans ces cas plus « choquants », nous avons la plupart du temps constaté qu'un terme ou plus (*souvent* dans ce cas-ci) faisait écran entre la préposition erronée et l'UL la régissant.

La phrase (68) démontre que les problèmes de prépositions régies touchent aussi les locutions. La locution CLOUER LE BEC régit la préposition À et fonctionne comme un bloc. Or, l'élève semble ici avoir analysé *public* comme le complément du nom *bec* plutôt que comme le complément indirect de l'expression *clouer le bec*; cette analyse erronée de la structure de l'expression¹³², pourtant valable d'un point de vue purement syntaxique, est une bonne démonstration du caractère capricieux des locutions.

Les quatre exemples présentés plus haut constituent des cas assez clairs du point de vue de la norme, puisque les dictionnaires offrent des exemples d'utilisation de ces verbes présentant les prépositions qu'ils contrôlent. Dans le cas du verbe SE DIRIGER, présent dans la phrase (67), les ouvrages indiquent même explicitement qu'il se construit avec la préposition VERS. Par contre, il est fréquent que les

¹³² Nous pouvons établir que cette structure est erronée en transformant la phrase (65) à l'aide d'un pronom CI (correspondant à *au public*) et avec un déterminant possessif (correspondant à *du public*, complément du nom *bec*). La première transformation serait correcte (*Pour lui clouer le bec...*), mais pas la seconde (**Pour clouer son bec...*).

prépositions contrôlées par un verbe n'apparaissent pas de façon aussi évidente dans les entrées des dictionnaires, voire qu'elles en soient absentes. Difficile alors d'expliquer pourquoi certains emplois nous semblent douteux, comme celui de la phrase suivante.

(69) *Le gardien reprit les trésors volés et repartit à [vers, retourna à] la pyramide. [199]*

Dans ce contexte, l'emploi de la préposition À auprès du verbe REPARTIR nous semble problématique, même s'il nous est possible d'imaginer des contextes où cette cooccurrence serait davantage acceptable (ex. *Il repartit à la campagne/à la maison/au campement*). Ce problème rappelle beaucoup celui de la phrase (67), mais il est moins aisé ici de démontrer qu'il s'agit d'une erreur.

Il existe aussi des cas où l'emploi que l'on fait d'un verbe influence le choix de la préposition à privilégier. Par exemple, dans la phrase (70), le verbe DÉCIDER est employé de façon pronominale, ce qui appelle la préposition À, alors que DE aurait été de mise si le verbe n'avait pas été employé pronominalement.

(70) *Je me décides d' [à] appeler pour travailler dans des restaurants. [288]*

Les adjectifs constituent la deuxième classe de mots la plus touchée par les erreurs de prépositions régies.

(71) *Il était triste par le [du] refu de sa belle mais il continuait de rêver d'elle. [198]*

(72) *Anxieux à [de] partir, une force le poussait quand même. [241]*

(73) *Sarah se lève motivée d' [à] aller jouer de la musique. [274]*

(74) *Cet homme était très apprécié par [du] le village. [4]*

(75) *J'ai entendu parlé de votre reportage concernant la publicité dirigée aux [vers les] jeunes. [142]*

Le même problème existe pour les adjectifs relativement à la non-systématicité des ouvrages de référence quant aux prépositions à utiliser : si les dictionnaires nous ont permis de trouver les « bonnes » prépositions à utiliser dans les phrases (71) et (72), aucun ne nous a fourni d'argument pour étayer notre intuition à l'effet que DE ne fonctionne pas avec MOTIVÉ. Le problème est encore plus criant pour les participes passés employés comme adjectifs (phrases (74) et (75)), qui n'ont souvent pas d'entrée propre et ne sont que brièvement mentionnés dans l'entrée du verbe auquel ils correspondent.

Les quelques noms touchés par les erreurs de prépositions régies nous conduisent au même constat : il n'est pas toujours évident d'établir la structure qu'ils régissent à l'aide des ouvrages. Si l'élève à l'origine de la phrase (76) aurait sans doute pu éviter de produire une phrase aussi boiteuse en se référant au dictionnaire, l'auteur de la phrase (77) n'aurait pas pu comprendre que TRAITEMENT, lorsqu'il désigne un produit et a pour complément le mal à soigner, se construit préférablement avec la préposition CONTRE¹³³.

(76) *L'influence **envers** les enfants à **propos** de la publicité de gomme et de friandises [L'influence de la publicité de gomme et de friandises sur les enfants] est déjà assez forte. [134]*

(77) *S'est le meilleur traitement **d'** [contre l'] acné. [51]*

B. ERREURS DANS LA CONSTRUCTION D'UN COMPLÉMENT

Près des deux tiers des erreurs de cette classe (19 cas sur 30) sont en lien avec le mode verbal imposé par une UL dans la phrase subordonnée qu'elle régit. À part deux erreurs liées à des adjectifs appelant le subjonctif dans la subordonnée à verbe conjugué les complétant (IMPOSSIBLE [18] et HEUREUX [104]) et une en lien avec un verbe demandant le même mode (ATTENDRE [173]), l'ensemble des erreurs concernent des conjonctions de subordination.

(78) *L'expérience la amusé **jusqu'à ce que** soudain, un coin du plafond **est tombé** [tombe] avec fracas. [287]*

(79) *Nini se dépêcha d'aller donner la bouteille à la voyante **pour qu'elle** la distribut au plus vieux pour que toute les vieux **reprirent** [reprennent] une apparence plus jeune. [92]*

(80) ***Malgré que c'était** [ce fût] son but principal, elle fut très surprise de trouver¹³⁴... [259]*

La plupart du temps, comme dans les exemples ci-haut, la subordonnée introduite par la conjonction doit être au subjonctif, alors que les élèves privilégient l'indicatif; les erreurs correspondant à ce patron mettent en jeu les conjonctions

¹³³ Il nous semble en effet que TRAITEMENT, lorsqu'il désigne un médicament, se construit avec la préposition CONTRE, mais pas lorsqu'il désigne un ensemble de moyens utilisés pour guérir quelqu'un, un processus. Ainsi, nous aurions tendance à dire *Un nouveau traitement contre l'* (et non **de l') acné a été découvert.*, mais *Le traitement de l'* (et non **contre l')* *acné provoque souvent des effets secondaires.*

¹³⁴ Bien que certains grammairiens admettent l'emploi de l'indicatif après les conjonctions BIEN QUE et MALGRÉ QUE, c'est le subjonctif qui est privilégié dans les ouvrages de référence. C'est la raison pour laquelle nous avons pris le parti de considérer la construction avec l'indicatif comme une erreur.

SANS QUE [18], POUR QUE [21, 76, 92], JUSQU'À CE QUE [42, 109, 168, 287], AVANT QUE [46], MALGRÉ QUE [259] et BIEN QUE [300]. Nous avons aussi rencontré deux fois l'erreur classique où le conditionnel est utilisé après la conjonction SI [247, 275] alors qu'elle exprime une condition et demande l'imparfait. Finalement, l'utilisation du subjonctif dans la subordonnée introduite par APRÈS QUE, bien qu'elle ne corresponde pas à l'usage prescrit dans les ouvrages de référence, qui préconisent l'emploi de l'indicatif, semble être la norme dans les textes des élèves, tout comme dans l'usage contemporain d'ailleurs; nous avons relevé trois cas [21, 239, 242], dont voici un exemple.

(81) *L'école se termina après que les examens **soient faits** [eurent été faits]. [242]*

Au-delà des problèmes de mode verbal, d'autres erreurs dans la construction d'un complément affectent des verbes transitifs directs que les élèves utilisent comme des verbes ditransitifs à l'aide d'une subordonnée infinitive.

(82) *... les techniques publicitaires utilisées pour **attirer** [inciter, pousser] les gens à **acheter**. [81]*

(83) *Ce truc professionnel est pour [...] nous **divertire** [inciter, pousser] à **se procurer le produit annoncé**. [82]*

(84) *Cette publicité est le meilleur choix pour vous, car elle va **attirer** [inciter, pousser, amener] des gens à **lire votre journal**. [130]*

Dans les trois cas, le premier complément est tout à fait correct, mais la subordonnée infinitive ne trouve pas sa place dans la structure syntaxique régie par ces verbes, même si le verbe ATTIRER pourrait admettre un second complément (*X attire Y dans/vers Z*).

Dans les phrases (85) et (86), l'élève ajoute un complément à un adjectif qui ne peut en avoir alors que l'exemple (87) présente un complément incompatible avec l'adjectif utilisé. En effet, APTE est généralement suivi d'une subordonnée infinitive (ex. *apte à pratiquer cet emploi*). Bien que les dictionnaires recensent des exemples où APTE est suivi d'un nom (ex. *apte au travail, à la conduite automobile*), il nous semble que ce nom doit avoir un caractère générique et nous paraît étrange de l'introduire à l'aide d'un déterminant démonstratif en faisant référence à un emploi particulier.

- (85) ... une grosse boule blanche mesurant environ deux mètres 65 **velu d'un poil blanc** [recouverte d'un poil blanc] magnifique. [167]
- (86) Ce qui semblait être un synonyme de libération (= les vacances d'été) laissa rapidement les deux amies **pantois de toute idée d'activité** [pantois]. [265]
- (87) Je me demandais si j'étais **apte** à [faite pour] cet emploi [apte à occuper cet emploi]. [230]

Nous avons en dernier lieu observé quelques cas plus marginaux, mettant en jeu des constructions particulières, comme dans les phrases suivantes.

- (88) Tu auras de quoi **à** [---] être contente. [254]
- (89) La feuille avait beau ressembler à un dessin abstrait **que** [,] Gregory l'envoya tout de même. [219]
- (90) Ce projet me tient **bien** [si, tellement] occupé **que** je n'ai pas le temps de m'ennuyer. [240]

Le pronom DE QUOI, selon *Le Petit Robert*, désigne 'qqch. qui fournit un moyen ou une raison de...' et se construit avec une subordonnée infinitive introduite directement, ce qui explique pourquoi le **à** de la phrase (88) est superflu. Dans le cas de la locution AVOIR BEAU, la construction avec l'infinitive est juste, mais la deuxième phrase, obligatoire avec cette expression, doit nécessairement être juxtaposée à la première à l'aide d'une virgule et ne peut être introduite par le subordonnant QUE : le patron syntaxique régi par cette locution est *X a beau Y, Z*, où Y est une infinitive et Z une phrase à verbe conjugué. Finalement, en (90), l'élève utilise l'adverbe BIEN comme s'il s'agissait d'un adverbe corrélatif, alors que ce n'est pas le cas.

C. ABSENCE D'UN COMPLÉMENT OBLIGATOIRE

Plusieurs des cas observés dans cette classe d'erreurs pourraient être considérés plutôt comme des problèmes syntaxiques que lexicaux, puisqu'ils ne semblent pas révéler une incompréhension de la structure régie par l'UL en jeu – une propriété idiosyncrasique –, mais traduisent plutôt un problème relatif à la construction de phrases. Par exemple, dans les phrases ci-dessous, le complément absent est très facilement « inférable » à partir du contexte, et il y a tout lieu de croire que c'est cette transparence sémantique qui a poussé les élèves à ne pas répéter un élément déjà présent dans le texte (phrases (91) et (92)) ou fortement sous-entendu (phrases (93) et (94)). Nous croyons en effet que les élèves à l'origine de

ces phrases savent que le verbe EMPÊCHER contrôle un troisième argument (*X empêche Y de Z*) et que le prédicat sémantique DIRE implique ce qui est dit (*X dit Y à Z*), tout comme ils savent que les quasi-prédicats MOITIÉ et FOND désignent nécessairement la moitié et le fond de quelque chose; ce qu'ils ne comprennent pas, en revanche, c'est que pour que la phrase qui contient ces termes soit syntaxiquement correcte, leurs compléments doivent y apparaître explicitement.

- (91) *Il désida de faire une fugue. Sa mère ne put l'empêcher [l'en empêcher]. [167]*
 (92) *Je voudrais lui dire qu'il est beau (...) Non, au lieu de lui dire [le lui dire] je vais lui demander son numéro de téléphone. [272]*
 (93) *Ça fait maintenant deux heures que je nettoie et j'ai à peine la moitié ___ [du travail] d'accomplie. [278]*
 (94) *Je me rends à mon premier cours de l'année et je m'assoit dans le fond ___ [dans le fond de la classe]. [221]*

Certaines erreurs, par contre, semblent trahir une méconnaissance plus profonde de la structure syntaxique régie par le verbe. En voici quatre exemples.

- (95) *Quelques enfants commençaient à quitter [quitter le camp de jour]. [267]*
 (96) *Une nuit, Mathieu allait s'aventurer [se promener, s'aventurer dans la ville], mais son nez avait détecté une odeur très désagréable. [117]*
 (97) *Ils furent rassurés lorsqu'ils virent trois jolies sirènes assises sur un rocher et qu'ils remarquèrent que les jolis chants doux parvenaient [provenaient, leur parvenaient] d'elles. [173]*
 (98) *Cette publicité nous fait croire que les stars s'arrachent pour avoir un produit comme Sunsilik [s'arrachent le produit Sunsilik]. [83]*

L'erreur présente en (95) est très fréquente dans la langue de tous les jours (du moins au Québec), mais le verbe QUITTER est un verbe transitif¹³⁵ qui, selon les ouvrages de référence, ne peut être utilisé en emploi absolu, de la même façon que le deuxième argument du verbe S'AVENTURER, employé dans le sens qui nous intéresse ici, doit être énoncé. Dans la phrase (97), PARVENIR signifie 'se propager à travers l'espace jusqu'à un lieu ou une personne donnés' et il semble que le deuxième argument de ce verbe (*X parvient à Y*) doive obligatoirement être exprimé pour que la phrase soit syntaxique correcte¹³⁶. Le verbe S'ARRACHER, dans

¹³⁵ Cet emploi intransitif de QUITTER a aussi été présenté plus haut parmi les erreurs de registres de langue. Voici un bon exemple de cas où un double-codage aurait pu s'avérer pertinent lors de la description des erreurs.

¹³⁶ Nous décrivons ici l'erreur comme un problème de régime, mais la proximité formelle et sémantique avec le verbe PROVENIR, qui serait tout à fait correct dans ce contexte syntaxique, est frappante.

la phrase (98), est tout à fait approprié d'un point de vue sémantique, mais l'élève qui l'a utilisé l'a fait comme s'il s'agissait d'un verbe pronominal réfléchi (comme SE BATTRE), ce qui n'est évidemment pas le cas : S'ARRACHER demande un complément direct, alors que l'élève actualise cet argument obligatoire à l'aide d'une subordonnée infinitive complément de phrase.

Les derniers cas typiques d'absence d'un complément concernent l'utilisation des adverbes corrélatifs TELLEMENT [2, 68, 197, 222, 251, 289] et SI [228] dans des phrases où seule la première partie de la structure corrélatrice apparaît. En voici deux exemples.

(99) *Il était sûrement dans son lit à dormir, parce qu'il était **tellement** [très] occuper durant la nuit. [2]*

(100) *J'étais **si** soulagée de l'avoir retrouvé. [!, que] Je le pris dans mes bras et le réconfortai. [228]*

Bien entendu, nous pourrions aussi considérer que ces erreurs sont plutôt des problèmes syntaxiques dans la mesure où si ces adverbes étaient plutôt apparus dans des phrases exclamatives, il aurait été correct de ne pas trouver la deuxième partie de la corrélatrice; dans ce cas, nous les aurions plutôt considérés comme des marqueurs d'exclamation, particulièrement dans la phrase (100).

D. LES ERREURS LEXICOSYNTAXIQUES SPÉCIFIQUES AUX VERBES

Toutes les classes d'erreurs lexicosyntaxiques discutées jusqu'à présent affectent de façon plus ou moins importante les verbes. Nous présenterons maintenant trois classes d'erreurs qui touchent exclusivement cette classe de mots, soit celles liées à la notion de transitivité : CS-Trans, CS-AsbPr et CS-Intrans.

La classe CS-Trans regroupe les 15 cas où un verbe transitif direct a été utilisé de façon transitive indirecte, c'est-à-dire qu'il y a eu ajout d'une préposition superflue pour introduire le complément du verbe, comme dans les exemples suivants.

(101) *Les images démontrant l'eau éclabousser **dans** [---] son visage sont époustouflantes. [128]*

(102) *Déjà, l'amertume avait gagné **sur** [---] Amélie. [265]*

(103) *Elle se rappellera à jamais **de** [---] cette belle expérience. [274]*

Le verbe ÉCLABOUSSER et le verbe GAGNER au sens de ‘s’emparer de’ sont tous deux des verbes transitifs directs, alors que les élèves leur ont accolé des compléments indirects dans les phrases (101) et (102). Le cas du verbe SE RAPPELER est similaire, à la différence près qu’il est très fréquent que l’on rencontre ce verbe employé de façon transitive indirecte – cette erreur apparaît d’ailleurs six fois dans notre corpus [100, 119, 139, 174, 274, 294] –, bien que cet emploi soit considéré incorrect par les grammairiens, ce qui se reflète dans les dictionnaires¹³⁷.

Dans certains cas, comme les trois suivants, il est plus difficile d’établir si l’erreur découle vraiment d’une lacune dans les connaissances lexicales de l’élève où si c’est plutôt le contexte syntaxique qui a interféré avec ses connaissances.

- (104) *Elle désirait plus que tout au monde d' [---] aller à Londres. [214]*
 (105) *Ses efforts allaient **lui récompenser** [le récompenser → être récompensés] car, il aperçu une énorme montagne plus loin devant lui. [78]*
 (106) *Je vais vous expliquer pourquoi nous étions choqués et pourquoi je **m'en rapelle** [me la (= la publicité) rapelle] même si je ne l'ai vu qu'une fois. [139]*

Dans la phrase (104), par exemple, un modificateur dans le GV (*plus que tout au monde*) fait écran entre le verbe DÉSIRES et son complément. Sans ce « distracteur », l’élève n’aurait peut-être pas commis l’erreur de régime; pour s’en assurer, il aurait fallu disposer d’une autre occurrence du même verbe dans un contexte où il est immédiatement suivi de la subordonnée infinitive qui fait office de complément direct. Dans le cas de la phrase (105), on peut se demander si le fait de pronominaliser le complément a influencé l’élève ou si ce dernier croit vraiment que la structure de RÉCOMPENSER est *X récompense *à Y*; encore une fois, une autre occurrence d’utilisation de ce verbe nous permettrait d’y voir plus clair. L’erreur de la phrase (106), quant à elle, rejoint celle de la phrase (103) présentée plus haut, si ce n’est que le complément indirect erroné est ici pronominalisé; cette erreur est par contre si répandue que la forme correcte nous semble presque plus bizarre que l’erreur. On peut donc aisément comprendre que l’élève ait effectué cette pronominalisation de cette façon!

¹³⁷ *Le Petit Robert* mentionne tout de même que la forme transitive indirecte est correcte avec un pronom complément représentant un être humain (ex. *Il se rappelle de toi.*).

À l'inverse des cas où un verbe transitif direct est utilisé de façon transitive indirecte, nous avons relevé six cas où un verbe transitif indirect voit son complément introduit sans préposition, comme dans les phrases suivantes.

(107) *Avant de **pénétrer le couloir** [dans le couloir], il lanca un processeur. [10]*

(108) *... la police avait mis une loi qui interdirait **les** [aux] paysans de s'aventurer la nuit. [117]*

(109) *Le jour vint **succéder** [succéder à] la nuit. [200]*

Le verbe PÉNÉTRER correspond à différents régimes selon ses acceptions; dans la phrase (107), il signifie 'entrer, en parlant d'êtres vivants' et se construit avec les prépositions DANS ou A L'INTERIEUR DE. Le verbe INTERDIRE, quant à lui, peut correspondre à plusieurs structures, selon qu'il est complété par un nom (*Son régime lui interdit le blé.*) ou une subordonnée infinitive (*La loi interdit aux citoyens de boire sur la voie publique.*); par contre, son deuxième argument est toujours introduit par la préposition A (*X interdit Y à Z* ou *X interdit à Z de Y*), ce qui explique l'erreur de l'élève ayant écrit la phrase (108), qui a utilisé le régime **X interdit Z de Y*. La phrase (109), quant à elle, est erronée parce que le verbe SUCCEDER est toujours transitif indirect.

Ici encore, nous avons rencontré des cas où la pronominalisation vient rendre moins évidente la source lexicale d'erreurs liées au régime, dont les deux suivants.

(110) *Pourquoi j'ai choisi cette publicité est une très bonne question et je vais vous **la** [y] répondre. [130]¹³⁸*

(111) *Il l'aimait aussi cette fille, mais il avait décidé de ne plus la voir, car il **l'avait avoué** [lui avait avoué] qu'il était vampire. [172]*

Dans ces deux phrases, il est difficile de savoir si les élèves auraient vraiment produit les GV **répondre la question* et **avouer la fille qu'il était vampire* s'ils n'avaient pas pronominalisé le complément indirect. Il est permis de penser que oui, mais sans interroger directement ces élèves, il est impossible de le confirmer.

¹³⁸ Bien évidemment, nous nous intéressons ici précisément au caractère indirect du complément à pronominaliser. Le pronom Y permettrait de remplacer le complément indirect à la question, mais ne pourrait pas apparaître en cooccurrence avec le pronom VOUS (**Je vais vous y répondre.*), qui renvoie à une autre construction du verbe RÉPONDRE. Dans la construction *X répond à Y*, Y peut désigner soit la question à laquelle on répond ou la personne à qui l'on répond.

Le dernier cas de figure relatif aux erreurs de régime verbal est l'ajout d'un complément à un verbe intransitif, comme dans les deux phrases suivantes.

- (112) ... *les trois gars s'enfuirent du* [*se sauvèrent du, s'enfuirent loin du*] *loup-garou.* [29]
 (113) *Avant de débiter* [*commencer*] *votre recherche d'emploi...* [236]

Le verbe S'ENFUIR, contrairement à son synonyme SE SAUVER [de], est un verbe intransitif et ne doit donc pas avoir de complément; en le modifiant par un adverbe, comme nous le suggérons dans notre correction, il aurait été possible de conserver l'élément *du loup-garou*, qui devient alors complément de l'adverbe *loin*. L'erreur avec DÉBUTER est identique, mais tellement répandue qu'elle fait l'objet de remarques dans les ouvrages de référence, qui mettent en garde contre la structure transitive du verbe DÉBUTER.

Deux cas sont un peu plus complexes.

- (114) *L'écléragage est sombre pour mieux ressortir* [*faire ressortir*] *les appareils qui les entourent.* [126]
 (115) ... *il coupa un énorme cécoya et le fit tomber le géant* [*à l'aide duquel il fit tomber le géant*]. [66]

Dans ces phrases, les élèves utilisent un verbe intransitif de façon transitive, comme s'ils avaient créé une nouvelle acception de ce verbe ayant un lien de causativité avec l'acception de base. Dans la phrase (114), par exemple, ce sont les appareils qui ressortent, c'est-à-dire qui 'apparaissent clairement, par un effet de contraste', mais l'élève utilise RESSORTIR au sens de 'causer que qqchse apparaisse clairement, par un effet de contraste'; pourtant, ce sens causatif du verbe RESSORTIR n'existe pas, ce qui fait que l'on peut considérer erronée la structure syntaxique produite par l'élève. Le cas est similaire pour TOMBER, puisque l'élève utilise ce verbe pour exprimer 'causer que qqchse tombe', le sens de la phrase (115) étant 'l'arbre, utilisé comme arme par le héros du récit, *tombe le géant (= cause que le géant tombe)'. Encore une fois, une telle acception du verbe TOMBER n'existe pas, du moins dans le registre standard (les ouvrages font mention d'emplois régionaux ou vieillis correspondant à ce sens transitif direct). Par contre, la structure complexe du message visé par l'élève dans la phrase (115) nous permet de douter

qu'il aurait réellement produit la structure *X tombe Y dans un contexte syntaxique plus simple.

IMPLICATIONS DIDACTIQUES

Nous croyons que les notions d'actant sémantique et de structure prédicative sont centrales dans la compréhension du fonctionnement de la syntaxe : pour savoir bien analyser une phrase, il faut comprendre quels éléments régissent les différents constituants qui la composent. À notre avis, prendre comme point de départ le régime du verbe principal de la phrase constitue l'approche la plus logique et la plus efficace lorsque vient le temps d'identifier ces constituants et de départager ce qui est essentiel de ce qui ne l'est pas.

Un travail en profondeur sur la notion de régime verbal devrait selon nous occuper une place de choix dans les cours de français. Les concepts de transitivité, d'intransitivité, de complément direct et de complément indirect sont bien sûr abordés dès le primaire, mais nous constatons que même à l'université, ils demeurent flous et mal maîtrisés, sans doute parce qu'ils sont rarement expliqués en prenant comme point de départ la sémantique. Les élèves connaissent bien le sens d'un nombre important de verbes et nous croyons qu'il faudrait tirer parti de leurs connaissances lorsqu'on enseigne ces concepts. Prendre conscience que les verbes, par leur sens, constituent le moteur syntaxique des phrases constitue selon nous un premier pas vers une meilleure compréhension de la syntaxe.

Bien sûr, les élèves ne connaissent pas les structures associées à tous les verbes et il est important de leur offrir des moyens d'accéder à ce type d'informations dans les dictionnaires. Pour ce faire, un travail sur la compréhension des différentes abréviations associées à la notion de régime est primordial : il faut non seulement que l'élève comprenne le concept de « verbe transitif direct », mais aussi qu'il sache que l'abréviation *v. tr.* y correspond dans certains dictionnaires (par exemple *Le Multidictionnaire*). Ceci dit, nous doutons que les notions mentionnées plus haut soient assez maîtrisées au secondaire pour que des indications du type *v. intr.* soient vraiment utiles aux élèves qui consultent les dictionnaires et il nous paraît primordial d'encourager les élèves à lire aussi les exemples d'utilisation du verbe fournis dans les ouvrages; faute de comprendre l'étiquette qui qualifie le type de régime en jeu, l'élève pourra ainsi, par imitation, produire la structure syntaxique correcte. Ceci implique bien entendu que l'ouvrage consulté fournisse assez d'exemples pour montrer les différentes structures associées au verbe en question, ce qui n'est pas toujours le cas. Combien de fois avons-nous cherché à déterminer quelle préposition convenait le mieux à un verbe sans pouvoir trouver réponse à notre question dans le dictionnaire!

En effet, il nous semble que les ouvrages de référence présentent des lacunes importantes dans leur description du régime des UL et que cet aspect donne lieu à de nombreuses incohérences. Par exemple, les trois ouvrages retenus dans cette recherche décrivent PARTIR comme un verbe intransitif, alors que

sémantiquement, il contrôle trois actants (*X part de Y vers/pour Z*), que l'on peut tout à fait exprimer syntaxiquement : devant l'indication *v. intr.*, l'élève doit-il comprendre qu'il commettrait une erreur en écrivant *Il part de la maison à 7 heures tous les matins*? Ou encore conclure que *de la maison* est complément de phrase? Et n'est-il pas aberrant de dire que PARTIR est un verbe intransitif au même titre qu'ÉTERNUER, qui lui ne contrôle véritablement qu'un seul actant (*X éternue*)? On constate le même type d'incohérences lorsqu'on s'intéresse aux verbes ayant plus d'un complément. Alors que les verbes à un seul complément sont généralement qualifiés de transitifs directs ou indirects, les verbes ditransitifs sont simplement marqués comme transitifs, ce qui nous donne bien peu d'informations quant à la structure qu'ils contrôlent et nous oblige à nous en remettre encore une fois aux exemples. Difficile aussi de trouver une information systématique sur la possibilité des verbes d'apparaître en emploi absolu; on en fait parfois mention, d'autres fois pas. Et nous ne parlons ici que des verbes, mais le régime relatif à des UL appartenant à d'autres classes de mots est encore moins bien décrit.

On peut donc imaginer qu'un élève se sente un peu dépourvu s'il décide de s'en remettre au dictionnaire pour des informations liées à la structure syntaxique régie par une UL. Les dictionnaires, du moins ceux ayant une visée pédagogique, devraient proposer plus de systématisme dans la façon de présenter les informations relatives à la combinatoire syntaxique, comme le font le *Lexique actif du français* (Mel'čuk et Polguère, 2007) et le *Dictionnaire du français usuel* (Picoche et Rolland, 2001), même si la consultation de tels ouvrages demande un certain entraînement.

Sans avoir à écrire un nouveau dictionnaire, un enseignant peut toutefois faire certaines choses pour aider ses élèves à commettre moins d'erreurs lexicosyntaxiques! Tout d'abord, le simple fait de les sensibiliser à la notion de régime devrait pouvoir les amener à se poser davantage de questions lorsqu'ils emploient une UL qu'ils connaissent moins. Par ailleurs, lorsqu'un enseignant explicite le sens d'un nouveau verbe en classe, par exemple à l'occasion de la lecture d'un texte, il pourrait mettre aussi l'accent sur la structure régie par ce verbe, notamment en en fournissant d'autres exemples d'utilisation. En plus de travailler adéquatement les notions de transitivité et d'intransitivité, l'enseignant devrait faire ressortir le fait que tous les compléments ne doivent pas toujours nécessairement être exprimés et que les exemples fournis dans les dictionnaires sont souvent très révélateurs de la façon d'utiliser une UL. Travailler sur les régimes des différentes acceptions d'un vocable peut aussi aider les élèves à comprendre que le sens est souvent intimement lié à la structure syntaxique. Il peut aussi être intéressant d'avoir recours à des outils particuliers, comme des ouvrages visant précisément le choix de la préposition (Grevisse, 1977 ; Bulman, 2003), ou de contraster les structures régies par des termes synonymes.

2.2.5 LES ERREURS DE COMBINATOIRE LEXICALE

Les problèmes liés à la combinatoire lexicale, comme le montre la figure ci-dessous, sont largement dominés par les collocations boiteuses ou inexistantes (CL-Inex : 78,7 % = 100 cas). Certaines contraintes syntaxiques imposées par les collocations (CL-Cont) permettent d'expliquer 11,8 % des cas (= 15) de cette grande classe, alors que nous relevons 8 cas (6,3 %) où la collocation employée existe, mais ne véhicule pas le sens visé par l'élève (CL-FL). Quatre cas de mauvaise dérivation sémantique (CL-Dér) ont enfin été constatés.

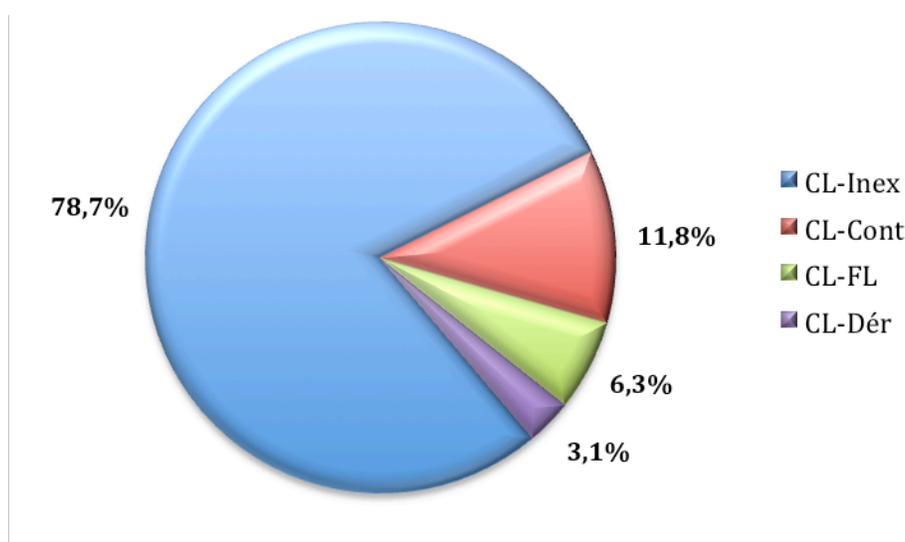


Figure 13. Répartition des 127 erreurs de combinatoire lexicale

2.2.5.1 LES COLLOCATIONS BOITEUSES OU INEXISTANTES

Comme nous l'avons mentionné, la majorité des erreurs relatives à la combinatoire lexicale sont liées à l'utilisation par les élèves de collocations qui ne correspondent pas aux cooccurrences qu'aurait spontanément produites un locuteur expert dans un même contexte d'énonciation. Pour analyser ces combinaisons boiteuses d'UL, nous avons utilisé le formalisme des fonctions lexicales (FL), présenté brièvement au chapitre II (point 3.6.2.1), de façon à pouvoir effectuer certains regroupements d'erreurs en fonction du type de lien lexical qu'elles mettent en jeu. Évidemment, tous les problèmes de cooccurrence relevés dans notre corpus ne se prêtent pas avec la même aisance à une modélisation à l'aide des FL, puisque certains ne correspondent pas à des

relations lexicales typiques et impliquent des FL non standards. Plutôt qu'une modélisation fine de l'ensemble des erreurs recueillies, nous avons donc opté pour l'encodage des erreurs correspondant aux patrons de collocations les plus typiques, c'est-à-dire celles pour lesquelles les FL standards étaient suffisantes, soit 85 des 100 erreurs relevées.

Parmi ces erreurs, la majorité (59 sur 85) concerne des collocations verbales, c'est-à-dire dans lesquelles le collocatif est un verbe; 57 de ces erreurs ont comme base un nom.

Le patron de collocation le plus fréquent est celui où la base de la collocation (nominale) joue le rôle de complément d'un verbe support qui n'ajoute aucune valeur sémantique à la base et n'a pour seul rôle que de la faire « fonctionner » syntaxiquement dans la phrase. C'est le cas dans les quatre exemples suivants¹³⁹ :

- (116) *Ils (= les publicitaires) **portent** [ont, exercent] une très mauvaise influence pour toutes les sociétés.* [82]
- (117) *Si vous achetez de la nourriture Pedigree (...), vous **contribuez** [faites, versez] un don à l'SPCA.* [85]
- (118) *... un regard terrifiant, un regard que seul notre héros aurait pu **faire** [avoir; jeter, lancer].* [23]
- (119) *J'ai regardé sur internet pour voir quel emploi je pourrait **faire** [exercer, occuper; obtenir].* [272]

Ces quatre erreurs correspondent en fait à la fonction lexicale *Oper* de la TST. Selon ce patron de collocation, un des actants sémantiques de la base remplit le rôle syntaxique de sujet, le collocatif correspond à un verbe support vide sémantiquement et la base de la collocation joue le rôle de complément direct de ce verbe; potentiellement, un autre actant sémantique de la base joue le rôle de complément indirect du collocatif. Illustrons cette configuration à l'aide de la collocation présente dans la phrase (116).

La base de la collocation présente en (116) est INFLUENCE, qui correspond à la structure actancielle suivante : *influence de X sur Y*. Dans cette phrase, l'élève cherchait un verbe prenant pour sujet le premier actant sémantique d'INFLUENCE et le mot-forme *influence* comme complément, en plus d'avoir le deuxième actant

¹³⁹ Pour faciliter la compréhension des exemples d'erreurs de collocation, nous avons souligné la base de la collocation, alors que le collocatif apparaît en gras.

sémantique d'INFLUENCE comme second complément. Cette configuration est schématisée ci-dessous.

SUJET	VERBE	COMPLÉMENT DIRECT	COMPLÉMENT INDIRECT
<i>La publicité</i>	<i>exerce, avoir, *porter</i>	<i>une influence</i>	<i>sur toutes les sociétés</i>
1 ^{er} actant sémantique d'INFLUENCE (= X)	Verbe support (sémantiquement vide)	Base de la collocation	2 ^e actant sémantique d'INFLUENCE (= Y)

Sous forme de fonction lexicale, cette collocation serait encodée comme suit, $Oper_{12}(influence) = exercer, avoir [une <de l'> \sim sur]$. En voyant cette formule, nous savons que les valeurs retournées par la fonction (items linguistiques suivant le signe =) seront des verbes supports ayant la base de la collocation comme complément; le chiffre 1 en indice à la suite du nom de la FL indique que le premier actant sémantique de la base est le sujet syntaxique du verbe support, alors que le chiffre 2 nous indique ici que les valeurs retournées par la FL auront aussi le deuxième actant sémantique de la base comme complément, en plus de la base de la collocation (qui est toujours le premier complément avec la FL $Oper$). Les éléments entre crochets indiquent la structure syntaxique contrôlée par les valeurs retournées et le tilde y représente la base de la collocation : on dit donc *avoir une influence sur* ou *exercer une influence sur*.

Dans la phrase (116), l'élève cherchait un verbe fonctionnant dans la structure $Oper_{12}(influence)$, mais a choisi le collocatif **porter*, qui ne correspond pas à un $Oper_{12}$ typique d'INFLUENCE; les collocatifs *exercer* ou *avoir* sont beaucoup plus appropriés dans cette structure. Par ailleurs, si nous disons que les valeurs retournées par la fonction syntaxique $Oper$ sont des verbes supports sémantiquement vides, c'est qu'ils n'apportent aucun ajout de sens; en effet, la phrase *La publicité exerce une influence négative sur la société*. ne signifie rien de plus que le syntagme nominal *l'influence négative de la publicité sur la société*. Bien entendu, ces verbes ont tout de même leur importance, puisqu'ils permettent de construire une phrase syntaxique autour d'une UL nominale, qui n'aurait pu constituer à elle seule le noyau syntaxique d'une phrase. Le cas de la phrase (117) est similaire,

puisque *contribuer n'est pas une valeur appropriée pour $Oper_{12}(don)$, qui correspond plutôt à *faire* et *verser*¹⁴⁰.

Les phrases (118) et (119) correspondent aussi à des cas d' $Oper$, mais ici à de simples $Oper_1$, c'est-à-dire des structures dans lesquelles le verbe support n'a qu'un seul complément, qui correspond à la base de la collocation, par exemple $Oper_1(emploi) = exercer, occuper [un \sim]$. Le cas de REGARD est intéressant, puisque la valeur $Oper_1(regard)$ n'existe qu'à la condition que REGARD soit modifié par un complément du nom; en effet, il est possible de dire *avoir un regard terrifiant/plein de compassion/qui vous glace le sang*, mais pas simplement **avoir un regard*. Les collocatifs *jeter* et *lancer* existent aussi auprès de REGARD, mais correspondent à une autre acception de ce vocable – signifiant 'action, fait de regarder; coup d'œil' plutôt qu' 'expression des yeux' –, selon nous moins compatible avec le contexte de la phrase (118) parce cette dernière ne contient pas de deuxième complément, et que ces collocatifs correspondent plutôt à la FL $Oper_{12}(regard) = jeter, lancer [un \sim à]$.

Ce ne sont pas toutes les phrases correspondant à la configuration syntaxique $Oper$ qui mettent en jeu des verbes supports vides de sens. Examinons par exemple les phrases suivantes :

(120) ... j'avais aussi "**pris**" [adopté] une partie de la mentalité européenne. [257]

(121) L'excursion **venait** [tirait, touchait] à sa fin quand tout à coup, un cheval incontrôlable vint perturber Cielo. [259]

(122) ... une fantastique idée lui **remit** [redonna] le sourire. [213]

En (120) et (121), nous avons toujours le même patron de collocation : le premier actant sémantique de la base remplit le rôle de sujet alors que la base de la collocation (respectivement MENTALITÉ et FIN) est complément du verbe. Seulement, dans ces deux phrases, le verbe véhicule le sens 'commencer à'; en (120), le narrateur commence à avoir une mentalité européenne et l'excursion commence à finir (c'est le début de la fin!) en (121). Cette idée de commencement se traduit dans les FL par la fonction Incep, qui se combine aux autres fonctions verbales. Ainsi, $IncepOper_1(mentalité) = adopter [une \sim]$ (plutôt que **prendre une*

¹⁴⁰ Bien sûr, l'erreur est double ici, puisque le verbe CONTRIBUTER n'est pas non plus un verbe transitif direct et est par conséquent incompatible avec le contexte syntaxique de la phrase (117).

mentalité) et $\text{IncepOper}_1(\text{fin}) = \text{tirer, toucher [à sa ~]}$ (plutôt que **venir à sa fin*). Dans la phrase (119), le verbe contient à la fois l'idée de causation et l'idée d'itération; nous pourrions paraphraser le sens du collocatif de cette phrase par 'causer que X ait le sourire de nouveau'. Si l'on tient pour acquis qu' $\text{Oper}_1(\text{sourire}) = \text{avoir [le ~]}$ ¹⁴¹, la FL correspondant au sens visé par l'élève en (122) serait $\text{CausDe_nouveauOper}_1(\text{sourire}) = \text{redonner [le ~]}$ (plutôt que **remettre le sourire*).

À la configuration syntaxique Oper , vide de sens, correspond la FL Real , qui véhicule le sens très général de 'faire ce qu'on doit faire (avec la base de la collocation)', ces deux FL renvoyant à une même configuration syntaxique. Par exemple, si nous considérons la lexie BAGUE au sens de 'bijou' (*bague de X*), alors $\text{Real}_1(\text{bague}) = \text{porter [une ~]}$. (= 'l'action typique que fait X avec une bague'). Si nous nous intéressons plutôt à la lexie AVERTISSEMENT (*avertissement de X à Y à propos de Z*), et plus principalement aux valeurs en lien avec son deuxième actant, nous obtenons $\text{Real}_2(\text{avertissement}) = \text{tenir compte [d'un ~], prendre en considération [un ~]}$ (= 'l'action typique que doit faire Y avec un avertissement'). Certaines des erreurs rencontrées dans notre corpus correspondaient à la FL Real .

(123) ... je vais **faire** [passer] la tondeuse à la maison. [269]

(124) Ces créatures sanguinaires s'attaquaient aux habitants de petits villages pour **réduire** [assouvir, satisfaire, éteindre] leur soif de sang. [161]

(125) Il commença à **réciter** [chanter/fredonner] les berceuses. [18]

Dans le premier cas, l'élève utilise le collocatif FAIRE, collocatif souvent qualifié de « mot passepartout » par les enseignants, qui interdisent parfois à leurs élèves d'utiliser ce verbe. Pourtant, FAIRE demeure un collocatif fréquent et tout à fait légitime auprès de plusieurs bases, mais correspond le plus souvent à la FL Oper (ex. *faire un commentaire, un pari, un témoignage, une erreur*). Or, la phrase (123) appelle plutôt une valeur de $\text{Real}_1(\text{tondeuse})$ signifiant 'faire ce que l'on doit faire avec une

¹⁴¹ En effet, on pourrait à plus forte raison considérer AVOIR LE SOURIRE comme une locution, puisque le sens de cette expression signifie davantage que l'addition du sens des éléments qui la constituent, cette expression signifiant 'être très content'. Par contre, comme sourire en ce sens apparaît en conjonction avec plusieurs verbes (*avoir, retrouver, perdre, garder le ~*), on peut aussi l'analyser comme base de collocations.

tondeuse', soit *passer* [la ~].¹⁴² Il en est de même pour la lexie SOIF de la phrase (124) ('désir passionné et impatient'), qui demande des collocatifs comme *assouvir* et *satisfaire*, voire *étancher*, qui nous paraissait plus approprié pour le sens premier de SOIF ('envie de boire'), mais que nous rencontrons même auprès de l'acception métaphorique dans les exemples fournis par *Antidote RX* (ex. *la soif de découverte que les livres nous permettent d'étancher*). En (125), finalement, les valeurs possibles de $Real_2$ (*berceuse*)¹⁴³ sont *chanter*, *chantonner*, *fredonner* plutôt que **réciter*.

Analysons maintenant une spécificité syntaxique de certaines des phrases présentant des erreurs de collocations correspondant aux FL Oper et Real : la coordination.

- (126) *Tes parents verront que tu n'es plus un enfant, donc ils te **laisseront** surment plus de liberté et [te donneront/t'accorderont plus] de permissions. [235]*
- (127) *Avec tous les efforts et les obstacles **commis** [(les efforts) faits, fournis, déployés (et les obstacles) franchis, surmontés] Zackary et ses amis réussèrent afin à aller chercher l'étoile magique. [111]*

Nous avons en effet remarqué que la coordination est en jeu dans un certain nombre d'erreurs de collocation de notre corpus. Il arrive que l'élève coordonne deux UL autour d'un même verbe, alors que seule une des deux bases fonctionne avec ce collocatif. C'est notamment le cas de la phrase (126), dans laquelle la collocation *laisser de la liberté* est tout à fait correcte, mais pas **laisser des permissions* ($Oper_{12}(permission) = accorder, donner [une \sim à]$). Il y a lieu de croire que l'élève n'aurait peut-être pas produit spontanément cette cooccurrence douteuse hors du contexte syntaxique de cette phrase et que la coordination l'a en quelque sorte embrouillé, le conduisant à commettre cette erreur. La phrase (126) est quant à elle beaucoup plus problématique, puisque l'élève y coordonne deux bases (*efforts*, *obstacles*) à un même collocatif (*commis*) alors qu'aucune des deux ne correspond à ce collocatif : on ne peut dire ni **commettre des efforts* ni **commettre des obstacles*. Ici, l'élève cherchait une valeur de $Oper_1$ pour EFFORT ($Oper_1(efforts) = fournir, faire,$

¹⁴² Il se peut bien sûr que l'élève ait été influencé par les collocations *faire le gazon* et *faire la pelouse*, qui sont équivalentes sémantiquement à *passer la tondeuse*.

¹⁴³ Si nous encodons ici ces valeurs comme $Real_2$ plutôt que $Real_1$, c'est que nous considérons pour BERCEUSE la même structure actancielle que celle de CHANSON, telle que décrite dans le DiCo, soit *chanson composée par X et interprétée par Y*; les valeurs de $Real_2$ correspondent donc aux actions typiques que fait Y (l'interprète) avec une berceuse.

déployer [des ~]) et une valeur de Real_2 pour OBSTACLE ($\text{Real}_2(\text{obstacle}) = \text{surmonter}$, *franchir* [un ~]). Difficile de comprendre, donc, pourquoi il a finalement opté pour le verbe COMMETTRE, mais il nous semble assez probable que la coordination ait, dans ce cas aussi, constitué une difficulté supplémentaire jouant contre les intuitions de l'élève, qui aurait peut-être produit des collocations plus convaincantes dans un contexte syntaxique différent.

Intéressons-nous maintenant aux collocations dans lesquelles la base joue le rôle de sujet du verbe collocatif. Cette configuration syntaxique correspond à la FL Func_0 , elle aussi sémantiquement presque vide, qui correspond à un sens très général (que l'on pourrait paraphraser par 'il y a'). Par exemple, $\text{Func}_0(\text{vent}) = \text{souffler} \rightarrow \text{le vent souffle} = \text{'il y a du vent'}$. Analysons tout d'abord des cas de Func_0 , c'est-à-dire des syntagmes dans lesquels la base de la collocation est sujet du verbe collocatif, qui n'a pas de complément.

(128) *Mais un bruit **ce fut** [retentit]. Dans le loin de la forêt. [124]*

(129) *Un grand fracas **se créer** [retentit, résonna] puis un silence **se fit entendre** [régna; s'installa, tomba]. [268]*

(130) *Un soir, où la lune **plombait** [brillait, luisait], Berhèse et ses amis se promenaient dans la forêt, loin des sentiers battus. [114]*

La phrase (128) constitue un excellent exemple de Func_0 , puisque le verbe collocatif qui fonctionne avec BRUIT permet d'exprimer le sens 'il y a un bruit'; l'élève, plutôt que de proposer ici une valeur appropriée pour $\text{Func}_0(\text{bruit})$, a utilisé le collocatif *se faire*, ce qui n'est pas sans rappeler le collocatif maladroit *se créer* de la phrase (129). BRUIT et FRACAS, termes synonymes, partagent en effet des collocatifs communs, dont *retentir* et *résonner*. La deuxième erreur en (129) – qui fait sourire! – correspond aussi à la même FL: $\text{Func}_0(\text{silence}) = \text{régner}$. Par contre, deux des corrections que nous proposons (*s'installa* et *tomba*), qui selon nous conviennent davantage au contexte, seraient mieux décrites par la fonction $\text{IncepFunc}_0(\text{silence})$ ('il commence à y avoir un silence'). Dans le cas de la phrase (130), l'élève a utilisé un collocatif qui aurait été tout à fait correct avec SOLEIL (dans un contexte québécois), mais semble étrange lorsqu'on parle de la lune; voilà un exemple de plus de l'arbitraire des collocations!

Il existe des configurations où un collocatif verbal correspondant à la fonction Func aura comme complément un des arguments de la base de la collocation. C'est ce type de configuration qui est en jeu dans les erreurs suivantes.

- (131) *Depuis ce jour d'autres rumeurs parlants d'attaques de deux Vampires **se dirent** [se repandirent, circulèrent, coururent] à travers le pays. [119]*
- (132) *Sur l'île Kyùshù, **il y avait** une menace pour [planait sur, pesait sur] tout le pays. [162]*
- (133) *Ça lui à permis de pouvoir aller au coeur du volcan, sans ressentir la chaleur immense qu'il y **avait** [régnait, faisait]. [110]*

La structure actancielle contrôlée par RUMEUR est *rumeur dans la population X à propos de Y*; ainsi, la collocation de la phrase (131) peut être modélisée comme suit : $Func_1(\text{rumeur}) = \text{circuler, courir [dans X]; se répandre [dans X]}$. La structure est similaire dans la collocation en jeu en (132), mais c'est plutôt le deuxième actant de MENACE qui est complément du verbe, puisque la structure actancielle de ce nom est *menace de X à Y (de faire Z)* : $Func_2(\text{menace}) = \text{planer, peser [sur Y]}$. Par contre, l'élève ayant produit la phrase (132) a plutôt choisi une forme impersonnelle, comme s'il avait tenté d'éviter d'avoir à utiliser un collocatif avec MENACE; on ne peut donc pas dire que cette phrase est aussi « erronée » que la précédente – dans laquelle l'élève avait clairement produit une valeur inexistante de $Func_1$ pour RUMEUR –, seulement, la tournure impersonnelle est plutôt pauvre ici et témoigne selon nous d'un certain manque de vocabulaire. L'erreur en (133) est presque identique, si ce n'est que les valeurs de $Func_1$ pour CHALEUR auraient été davantage compatibles avec la tournure impersonnelle privilégiée par l'élève (*Il fait/règne une chaleur insoutenable dans ce volcan.*, mais moins *Il plane/pèse une menace sur ce pays.*).

Tout comme c'était le cas avec la FL Oper, Func peut aussi se combiner à d'autres FL pour encoder le sens plus complexe de certains collocatifs, comme dans les phrases suivantes, dans lesquelles les collocatifs expriment la causativité.

- (134) *Celui qui l'avait **produit** [conçu, engendré] ne faisait pas d'elle (= l'enfant de son père) qui elle était devenue. [204]*
- (135) *Pour **annuler** [rompre, briser] le charme, il fallait tuer le vampire qui l'avait mordu. [37]*

Dans (134), l'élève cherche un verbe qui signifie 'causer qu'un enfant existe' : il choisit *produire, alors que les verbes véhiculant ce sens auprès d'ENFANT sont

plutôt *concevoir* et *engendrer*, qui correspondent à la fonction $\text{Caus}_2\text{Func}_0(\text{enfant})$. L'idée de causativité est aussi présente dans la phrase (135), cette fois liée à l'idée de fin; ainsi, l'élève cherche la façon d'exprimer 'causer qu'un charme cesse d'exister' ($\text{CausFinFunc}_0(\text{charme})$), et choisit le verbe **annuler*, plutôt que *rompre* ou *briser*.

Une autre catégorie de collocations semble poser problème à certains élèves : l'intensification, c'est-à-dire l'utilisation d'un collocatif pour intensifier le sens de la base, ce qui correspond à la FL Magn. La configuration la plus fréquente parmi les erreurs relevées est l'utilisation d'un adjectif pour modifier une base nominale, comme dans les exemples suivants.

- (136) ... en moins de 30 secondes, elle avait tombée dans un **grand** [profond] sommeil. [76]
- (137) Ça lui a permis de pouvoir aller au coeur du volcan, sans ressentir la chaleur immense [étouffante, accablante, intense] qu'il y avait. [110]
- (138) L'exagèrement de cette publicité est assez **intense** [poussée]. [134]
- (139) Les premiers obstacles du superhéro sont très **faciles** [faciles à surmonter, surmontables]. [54]

Dans les trois premiers exemples, les élèves ont utilisé des adjectifs qui, bien qu'ils traduisent l'idée d'intensification visée, se combinent mal avec les bases qu'ils modifient. On ne dit pas d'un *sommeil* qu'il est **grand*, mais plutôt *profond* ou *lourd* ($\text{Magn}(\text{sommeil}) = \text{profond}, \text{lourd}$), de la même façon qu'une *chaleur* est *étouffante*, *accablante*, *torride*, *suffocante*, *écrasante*, etc. et qu'une *exagération*¹⁴⁴ est *poussée*, *grotesque*, *outrancière* ou *ridicule*. Le cas de la phrase (139) est similaire, si ce n'est que le collocatif vise l'« atténuation » de la base plutôt que son intensification : $\text{AntiMagn}(\text{obstacle}) = \text{surmontable}, \text{léger}$. La plupart du temps, comme dans ces exemples, il est possible de proposer un ou plusieurs collocatifs beaucoup plus idiomatiques que ceux utilisés par les élèves, mais nous avons aussi rencontré quelques cas pour lesquels il n'existe pas de valeurs correspondant à la FL visée.

- (140) La "pub" utilise du langage courant et un lieu familier pour les **vieux adolescents** [jeunes adultes]. [146]

¹⁴⁴ Nous parlons bien sûr des collocatifs d'EXAGÉRATION, puisque **exagèrement* n'existe pas.

Dans cette phrase, par exemple, l'élève tente de désigner par une collocation les adolescents qui sont vers la fin de leur adolescence. Or, si la collocation *jeune adolescent* est tout à fait usitée pour dénoter quelqu'un qui entre dans l'adolescence, aucun adjectif ne permet de qualifier ADOLESCENT pour faire référence à une personne qui quitte cette période de la vie pour entrer dans l'âge adulte ($\text{Magn}^{\text{âge}}(\text{adolescent})^{145} = \emptyset$). Par conséquent, l'élève aurait dû utiliser l'expression *jeune adulte* plutôt que *vieil adolescent*, puisqu'il n'existe pas en français de collocation pour exprimer ce sens auprès d'ADOLESCENT.

Dans quelques cas, beaucoup plus rares, la FL Magn s'applique non pas à des noms, mais à des adjectifs ou des verbes.

- (141) *Têtue comme un cochon [comme une mule/un âne], le chef refusa. [22]*
- (142) *William s'était tristement déçu [profondément déçu] quand il a su (...) que ses parents se sont chicanés. [49]*
- (143) *En leur annonçant la nouvelle pour le moins farfelue, ses parents (...) la fixait légèrement.*

Dans la phrase (141), l'élève a sans doute été influencé par l'expression familière TÊTE DE COCHON, qui désigne une 'personne très têtue' et constitue en quelque sorte une valeur fusionnée de $\text{Magn}(\text{têtu}) = \text{comme un âne, comme une mule, comme une bourrique // tête de cochon}^{146}$. Dans la phrase (139), l'élève a utilisé l'adverbe **tristement* pour intensifier le sens de *déçu*, alors que les adverbes comme *profondément* ou *extrêmement* auraient été beaucoup plus appropriés. En (143), le problème est différent. Outre le fait que la base de la collocation soit un verbe, ce qui frappe est surtout l'incompatibilité entre le sens d'atténuation visé par l'élève (AntiMagn) et la base de la collocation. Comme le verbe FIXER signifie déjà 'regarder longuement avec une certaine intensité', il est impossible de le combiner avec une valeur d'AntiMagn, puisque son sens véhicule en lui-même l'idée d'intensification. Par conséquent, on pourrait analyser ce problème de cooccurrence comme un

¹⁴⁵ Nous indiquons *âge* en exposant après la FL, puisque l'intensification porte ici sur une composante de sens précise d'ADOLESCENT, soit l'âge.

¹⁴⁶ Il est possible de considérer TÊTE DE COCHON comme une valeur fusionnée de $\text{Magn}(\text{têtu})$ seulement si l'on admet que cette expression peut être utilisée comme un adjectif (ex. *Il est très tête de cochon.*) et non seulement comme un nom.

problème plutôt sémantique, puisque l'élève ne semble pas bien saisir le sens du verbe FIXER.

De façon générale, nous constatons donc que les problèmes liés aux collocations mettent presque exclusivement en jeu des bases nominales et correspondent à des collocatifs sémantiquement très « légers » traduisant des sens plutôt vagues ('exister', 'avoir lieu', 'commencer', 'finir', 'causer', 'intensification', etc.). Comme la façon d'exprimer ces sens vagues est contrainte par le choix de la base, on peut comprendre que les élèves produisent des cooccurrences qui arrivent à traduire le sens visé, mais qui sont déficientes d'un point de vue de la norme (très floue en cette matière d'ailleurs) : **faire son plan* [178] plutôt que *mettre son plan à exécution*, **avoir des changements* [203] plutôt que *connaître, subir des changements*, **monter les ventes* [55] plutôt que *mousser, augmenter les ventes*, **mettre une loi* [117] plutôt que *passer, adopter une loi*, **tomber en éruption* [187] plutôt qu'*entrer en éruption*, **force puissante* [199] plutôt que *force surhumaine, incroyable*. Par ailleurs, nous croyons que la hantise que peuvent développer les élèves à force de se faire répéter par des enseignants d'éviter certains verbes (AVOIR, FAIRE, PRENDRE, etc.) n'est peut-être pas étrangère à quelques-unes des erreurs observées, du type **s'offrir de grosses dépenses* plutôt que *faire de grosses dépenses* [280] ou **créer un curriculum vitae* [260] plutôt que *faire, rédiger un curriculum vitae*.

Les nombreuses erreurs présentées plus haut nous permettent de mieux saisir l'arbitraire du phénomène de la collocation. Comment, alors, les élèves auraient-ils pu trouver les « bons » collocatifs pour exprimer leur message? Est-ce que le dictionnaire ne leur aurait pas donné des pistes? Il est vrai que certains des collocatifs correspondant aux FL en jeu dans les erreurs analysées apparaissaient dans les exemples des dictionnaires – principalement dans *Le Petit Robert* et dans *Antidote RX*, le *Multidictionnaire* étant moins riche en exemples –, sous l'article de la base de la collocation bien entendu. On y retrouve principalement certaines valeurs d'Oper, de Real, de Func et de Magn. Par contre, ceci n'est pas vrai pour la majorité des erreurs observées et les élèves, même s'ils avaient consulté un dictionnaire comme *Le Petit Robert* ou *Le Larousse* lors de la rédaction, n'auraient

trouvé le collocatif approprié que dans une minorité de cas. Par contre, s'ils avaient eu accès à un dictionnaire de cooccurrences, comme l'ouvrage de Jacques Beauchesne ou *Antidote RX*, ils auraient selon nous trouvé un collocatif correspondant à la structure et au sens visé la plupart du temps. Dans certains cas – comme **un somptueux gâteau* [125] pour *un délicieux gâteau* ou **un objectif famineux* [265] pour *un objectif ambitieux* –, chercher sous l'article du collocatif aurait pu les éclairer sur le caractère problématique de leurs collocations. Par contre, cette technique n'aurait pas été concluante dans le cas de collocations verbales sémantiquement vides du type **porter une influence*, pour laquelle chercher PORTER dans le dictionnaire n'aurait pas été d'un grand secours. Quel dictionnaire propose une définition de PORTER qui justifie que l'on peut dire *porter un jugement*, mais pas **porter une opinion*?

2.2.5.2 LES CAS DE NON-RESPECT DES CONTRAINTES D'EMPLOI DU COLLOCATIF

Les problèmes dont nous traitons dans cette section sont beaucoup moins fréquents que les collocations boiteuses de la section précédente. Il s'agit en fait de collocations correctes en termes de choix du collocatif en fonction de la base et du sens visé, mais qui présentent malgré tout certains problèmes liés à des contraintes imposées par la collocation. En voici quelques exemples.

- (144) *Il poussa un cri inimaginable à son ami [lança un cri à son ami]. [32]*
- (145) *... une publicité qui attire beaucoup d' ['] attention. [83]*
- (146) *Tout à coups, une grosse bête avec de longues battes griffues et des moustaches tournoyantes fit apparition [fit son apparition]. [170]*
- (147) *Arriver devant sa porte, Bobby la [---] frappa. [255]*

En (141), le complément indirect à *son ami* est problématique après la collocation *pousser un cri*. Si nous avons classé cette erreur parmi les problèmes de collocation, c'est que cette contrainte de « saturation » des compléments possibles nous paraît liée à la collocation *pousser un cri* plutôt qu'au seul verbe POUSSER; en effet, une phrase comme *Émile a poussé le ballon à sa sœur*. – dans laquelle POUSSER n'est pas un collocatif choisi en fonction d'une base – nous semble nettement plus acceptable que la phrase (144). Dans la phrase (145), c'est l'utilisation d'un déterminant indéfini (*beaucoup de*) qui est problématique : lorsque ATTIRER agit comme collocatif d'ATTENTION, il faut que le déterminant qui introduit ATTENTION

soit un déterminant défini (ex. *Cette publicité attire l'attention.*) ou possessif (ex. *Cette publicité attire son attention.*). L'absence de déterminant dans la collocation *faire son apparition* explique l'erreur présente en (146). Encore une fois, lorsque *faire* est collocatif d'*apparition*, la base doit être introduite soit par un déterminant indéfini – auquel cas *apparition* doit être modifié par un complément (ex. *Elle fit une apparition digne des plus grandes divas.*) –, soit par un déterminant possessif (ex. *Ils font leur apparition à la tombée de la nuit.*). Il est moins clair que l'élève ayant écrit la phrase (147) a commis une erreur de nature lexicale, du fait de la pronominalisation de la base de la collocation; difficile de savoir en effet si dans un autre contexte il aurait écrit **frapper la porte* ou *frapper à la porte*.

Les contraintes en jeu dans ce type d'erreurs sont très rarement indiquées dans les ouvrages de référence et par conséquent très difficiles à expliquer ou à justifier. Encore une fois, c'est en portant attention aux exemples fournis par les ouvrages que les élèves auraient peut-être pu éviter ces erreurs, pourvu bien sûr que la collocation qui les intéresse y soit recensée.

2.2.5.3 L'EMPLOI D'UN COLLOCATIF VALABLE, MAIS QUI NE CORRESPOND PAS AU SENS VISÉ

Huit erreurs de notre corpus correspondent au cas de figure suivant : la collocation produite par l'élève existe, mais ne correspond pas au sens visé dans le contexte. En voici deux exemples.

(148) [Les héros sont en panne d'essence.] *Marc vit une petite maison abandonnée. Ils allèrent la visiter. Tout était cassé et très malpropre. Tout à coup, ils virent un tandem. Ils **firent** du vélo [continuèrent à vélo, utilisèrent le vélo], car ni l'un ni l'autre n'eut une autre idée pour se rendre à Las Vegas. [256]*

(149) *Plusieurs publicités ont suscité de l'intérêt à **mon égard** [chez moi, ont suscité mon intérêt]. [153]*

La collocation *faire du vélo* existe bien entendu, mais nous semble liée à une activité récréative, ce qui est incompatible avec le contexte de l'histoire, dans laquelle les héros tombent en panne d'essence au milieu d'une route déserte et utilisent un tandem pour se rendre à destination. Un collocatif plus neutre comme *prendre*, *utiliser* ou *continuer* aurait été plus approprié.

L'erreur de la phrase (146), quant à elle, porte vraiment à confusion. L'élève tentait d'exprimer qu'une publicité avait suscité son intérêt, mais a utilisé une structure

syntactique signifiant que cette publicité a suscité un intérêt général pour l'élève en question. La structure syntaxique d'INTÉRÊT est *intérêt de Y pour (à l'égard de) X*¹⁴⁷; or, l'élève a utilisé une préposition servant typiquement à introduire la source de l'intérêt (X) pour désigner la personne témoignant de l'intérêt (Y).

Ce type d'erreurs de collocation, même si elles sont intéressantes à analyser, sont difficiles à expliquer et demeurent plutôt marginales, ce qui fait qu'elles méritent selon nous moins d'attention didactique que les problèmes de collocations présentés plus haut. Passons donc maintenant à la présentation des erreurs liées à la dérivation sémantique.

2.2.5.4 EMPLOI D'UN MAUVAIS DÉRIVÉ SÉMANTIQUE

Nous avons décidé d'ajouter cette classe d'erreur à notre typologie à la suite de notre mémoire de maîtrise (Anctil, 2005), puisqu'une des erreurs relevées alors dans notre corpus était en lien avec la dérivation sémantique.

Nous parlons de dérivation sémantique lorsqu'une relation lexicale récurrente existe entre des UL¹⁴⁸, mais, contrairement à la collocation, les UL unies par cette relation n'apparaissent pas forcément en cooccurrence : elles se situent sur un axe paradigmatique plutôt que syntagmatique. La synonymie est une relation typique de dérivation sémantique, et ce type de relation peut aussi être modélisée à l'aide des FL (ex. *Syn(riche) = fortuné, aisé, nanti*).

Trois des quatre erreurs de dérivation relevées dans notre corpus sont en lien avec les animaux et correspondent à des FL existantes.

(150) ... le gros **félin** [*canidé, renard*] argenté créa un disque de lumière. [26]

(151) On entendit un **hululement** [*hurlement*] tout près et on vit une quinzaine de loups-garous surgir. [35]

(152) Il était une fois, dans une des magnifiques forêts du Japon, vivait un petit renard (...) Un jour, le pauvre petit 罽入 (*larme de feu*) fut abandonné par le reste de son **groupe** [*meute*]. [194]

¹⁴⁷ Si le X et le Y peuvent sembler inversés ici, c'est que nous avons établi la structure actancielle du nom INTÉRÊT à partir du verbe correspondant : *X intéresse Y* → *intérêt de Y à l'égard de X*.

¹⁴⁸ Rappelons que la dérivation sémantique, contrairement à la dérivation morphologique, n'implique pas nécessairement une parenté formelle entre les UL en jeu. La relation lexicale qui permet de « dériver » une UL d'une autre sera parfois marquée par un lien de forme (ex. UTILE ~ INUTILE), parfois non (ex. GENTIL ~ MÉCHANT).

Dans la phrase (150), l'élève tente de produire un hyperonyme pour RENARD, sans doute pour éviter de répéter ce terme encore une fois. Il choisit par contre le terme FÉLIN qui, malgré nos connaissances plutôt limitées en zoologie, nous semble plutôt être un hyperonyme pour CHAT, LION, TIGRE, etc.; la relation lexicale entre les deux termes est donc inexistante, ce qui à la limite aurait pu introduire un problème de compréhension pour le lecteur, puisque le terme erroné est utilisé comme élément de reprise, alors qu'il ne présente pas de lien de sens avec l'élément repris (*renard*).

L'erreur en (151) correspond à la FL Son, qui associe à une UL le nom typique du son produit par ce qu'elle désigne¹⁴⁹ (ex. Son(*bouilloire*) = *sifflement*, Son(*éléphant*) = *barrissement*). Dans le texte contenant cette erreur, on comprend clairement que le son qu'entendent les gens du village est le cri d'un loup (un loup-garou, plus précisément), alors que l'élève utilise HULULEMENT, qui serait plutôt une valeur de Son(*hibou*) ou de Son(*chouette*). Évidemment, la parenté formelle entre les deux UL y est sans doute pour quelque chose.

L'erreur en (152) renvoie quant à elle à la FL Mult, qui retourne comme valeur le nom typique donné à un groupe d'entités dénotées par l'argument de la FL (ex. Mult(*abeille*) = *essaim*, *colonie*, Mult(*clé*) = *trousseau*). Dans le cas du renard, la valeur de Mult appropriée serait sans doute MEUTE, comme c'est le cas pour d'autres canidés comme les loups, les chiens ou les chacals. Par contre, la collocation *meute de renards* n'apparaît dans aucun ouvrage, ce qui nous conduit à poser l'hypothèse que le renard doit être un animal plutôt solitaire... Comme quoi l'analyse d'erreurs lexicales sollicite plusieurs types de connaissances et que nous aurions avantage à nous mettre très sérieusement à la zoologie!

¹⁴⁹ Nous prenons ici un raccourci en passant directement d'un nom à un autre, puisque la FL Son retourne généralement des verbes comme valeurs. La façon correcte d'encoder le lien dont nous parlons ici serait S_{θ} Son(*bouilloire*) = *sifflement*, ou S_{θ} indique que la valeur retournée par la FL doit être l'équivalent sémantique nominal (*sifflement*) du verbe (*siffler*).

IMPLICATIONS DIDACTIQUES

Pour bien intervenir devant des erreurs liées à la combinatoire lexicale, il faut tout d'abord une bonne compréhension du phénomène de collocation. Ceci n'implique pas nécessairement d'être habile à manier les fonctions lexicales, ni même de connaître ce formalisme, mais au moins d'avoir une forte intuition, devant une expression de type collocation, nous permettant de savoir qu'un des éléments de l'expression est choisi librement (la base) et contraint le choix de l'autre (le collocatif). Il faut aussi savoir que ce n'est généralement pas en allant consulter l'article de dictionnaire du collocatif que nous pourrions justifier à l'élève qu'il a commis une erreur, puisqu'il est fréquent, comme nous l'avons vu plus haut, que le collocatif ne véhicule pas un sens très précis à l'intérieur de la collocation. Il est essentiel de garder en tête que la collocation est un phénomène très arbitraire que ne sous-tend aucune logique : c'est l'usage qui détermine la norme en cette matière et toute explication logique pour tenter d'expliquer une erreur de collocation mènera à une impasse, comme le démontre cet échange entre un élève et un enseignant tiré de l'article de Polguère (2004b: 5).

L'élève : – *Pourquoi dit-on « passer des vacances » et « faire un séjour », et pas « faire des vacances » et « passer un séjour »?*

Le professeur : – *... mais... c'est parce que les vacances c'est pour se reposer alors que dans un séjour on fait quelque chose.*

Nous avons mentionné que les dictionnaires grand public comme ceux utilisés en classe sont généralement assez pauvres en ce qui concerne les collocations. Néanmoins, il est fréquent de trouver un ou deux collocatifs parmi les plus communs dans une entrée de dictionnaire si l'on est attentif aux exemples fournis; généralement, ils correspondent aux FL dont nous avons parlé plus haut (Oper, Func, Magn, Real). Il est donc important d'encourager les élèves à lire les exemples lorsqu'ils consultent le dictionnaire. Il pourrait aussi être intéressant de disposer en classe d'ouvrages visant plus précisément les cooccurrences, comme le *Dictionnaire et thésaurus HRW*, *Le dictionnaire des cooccurrences à l'usage des écoles* de Jacques Beauchesne ou le logiciel *Antidote*, si la classe dispose d'un ordinateur.

Lors de la correction de productions écrites, il serait pertinent de suggérer à l'élève un collocatif approprié dans le cas où la collocation utilisée paraît boiteuse; on peut aussi consigner les « belles » erreurs de collocation pour les retravailler en classe en trouvant de meilleures cooccurrences. Et surtout, il ne faut pas pénaliser un élève qui emploie un verbe comme FAIRE ou AVOIR dans le cadre d'une collocation comme **faire une remarque** ou **avoir une opinion**, puisqu'il s'agit dans ces cas de collocatifs typiques tout à fait valables pour ces bases, même s'ils correspondent à des verbes que l'on qualifie de « passepartout ».

Comme nous l'avons mentionné pour plusieurs autres types d'erreurs, comparer des termes sémantiquement proches pour voir s'ils partagent certains collocatifs peut constituer un exercice très formateur qui sensibilisera les élèves au phénomène de collocation. Par exemple, il est intéressant de voir que l'on dit **faire un choix** et **prendre une décision**, mais pas ***prendre un choix** ou ***faire une décision**; pourtant CHOIX et DÉCISION sont des synonymes. De la même façon, les

élèves pourraient trouver étonnant que les antonymes BON et MAUVAIS expriment exactement le même sens auprès de GRIPPE (Magn(*grippe*) = *vilaine, bonne, mauvaise, carabinée*).

Lorsque les élèves rencontrent un nom nouveau lors de lectures, on peut aussi attirer leur attention vers les verbes qui l'accompagnent, de façon à enrichir la facette collocationnelle de leur connaissance de l'UL en question. On peut aussi leur suggérer de prendre en note, dans un aide-mémoire lexical par exemple, les combinaisons de « mots » qu'ils trouvent belles ou intéressantes pour les réinvestir dans leurs textes par la suite; cette façon de faire correspond d'ailleurs à la façon dont Jacques Beauchesne a rédigé son dictionnaire.

2.3 LES TYPES D'UNITÉS LEXICALES EN JEU DANS LES ERREURS LEXICALES DU CORPUS

Lorsque nous avons codé nos erreurs, nous avons mentionné pour chacune si elle concernait un lexème ou une locution. L'analyse de notre corpus de 1144 erreurs révèle que près d'une sur cinq concerne une locution (214 cas = 18,7 %); cette proportion est réduite à 14,5 % si l'on considère aussi les erreurs d'orthographe et de flexion (395 locutions sur 2720 erreurs). Cette proportion assez faible de problèmes liés aux locutions semble corroborer une observation que nous nous sommes faite lors de l'analyse du corpus : les élèves utilisent très peu les locutions lorsqu'ils écrivent. Bien entendu, cette affirmation repose sur notre impression seulement, et il aurait fallu considérer l'ensemble des locutions utilisées (correctement ou non) pour pouvoir quantifier le recours aux expressions par rapport au nombre de lexèmes utilisés. Il faudrait aussi être en mesure de savoir dans quelle proportion les scripteurs experts utilisent eux-mêmes des expressions, puisqu'il nous semble évident que les lexèmes et les locutions n'apparaissent pas de manière égale dans leurs textes, même dans ceux présentant un vocabulaire riche. Néanmoins, si l'on considère que les locutions sont aussi susceptibles que les lexèmes d'être touchées par des erreurs lexicales, les proportions issues de notre analyse révèlent tout de même une certaine sous-représentation des locutions dans les rédactions des élèves; les enseignants à qui nous avons exposé cette observation semblaient d'ailleurs d'accord pour affirmer que leurs élèves ont peu recours aux expressions, tant à l'oral qu'à l'écrit.

Outre les 176 erreurs orthographiques liées aux locutions, deux types d'erreurs contiennent un nombre élevé de cas et présentent une proportion importante de locutions : les barbarismes (ex. **de plein feu* [72] pour DE MILLE FEUX, **le sourire à la bouche* [241] pour [avoir] LE SOURIRE AUX LÈVRES, **sur le camps* [253] pour SUR-LE-CHAMP), où elles représentent 34,4 % des 209 erreurs, et les problèmes de registres de langue (ex. (*on l'appelait*) **de même* [59] pour AINSI, **à force* [286] pour À LA LONGUE, **faire ça vite* [228] pour SE DÉPÊCHER), où 27,1 % des 192 cas concernent des locutions. Ces observations témoignent peut-être du fait que les élèves connaissent certaines expressions, mais en ont une connaissance imparfaite. Fait remarquable : dans les 30 cas d'erreurs de construction d'un complément, c'est plus de la moitié des erreurs (53,3 %) qui concernent des locutions.

IMPLICATIONS DIDACTIQUES

Le fait que les élèves semblent peu utiliser les locutions dans leurs textes nous laisse croire qu'il serait important de mettre davantage l'accent sur ce type d'UL dans les classes de français. Il ne s'agit pas nécessairement de faire l'enseignement direct de locutions spécifiques, mais bien de sensibiliser les élèves aux expressions rencontrées dans les lectures et de faire soi-même preuve de curiosité face à ces UL, par exemple en mentionnant des nouvelles expressions entendues ou lues.

Il serait important de montrer que les locutions, comme les autres « mots », régissent certaines structures et appartiennent à des registres de langue spécifiques. Il peut aussi être intéressant de comparer des expressions québécoises (ex. LÂCHE PAS LA PATATE) et franco-françaises (ex. TENIR BON LA RAMPE) pour bien faire ressortir la dimension culturelle des locutions.

Par ailleurs, bien que les dictionnaires soient généralement assez peu développés dans leur description des locutions, il est tout de même possible d'y retrouver la plupart des expressions, ne serait-ce que pour en vérifier la forme. À cet effet, il peut être pertinent de montrer aux élèves comment les retrouver dans le dictionnaire, c'est-à-dire en cherchant sous l'article d'un des mots-formes principaux constitutifs de l'expression, généralement vers la fin de l'article; par exemple, PRENDRE SON COURAGE À DEUX MAINS devrait se trouver à la fois sous COURAGE et sous MAIN, mais probablement pas sous PRENDRE.

Si les élèves tiennent un aide-mémoire lexical, il peut être intéressant d'y consigner les nouvelles expressions qu'ils rencontrent, toujours en en indiquant le sens, par exemple à l'aide d'un synonyme.

2.4 LES CLASSES DE MOTS DES UL EN JEU DANS LES ERREURS DU CORPUS

Le tableau ci-dessous représente la répartition des 1144 erreurs dans les différentes classes de mots.

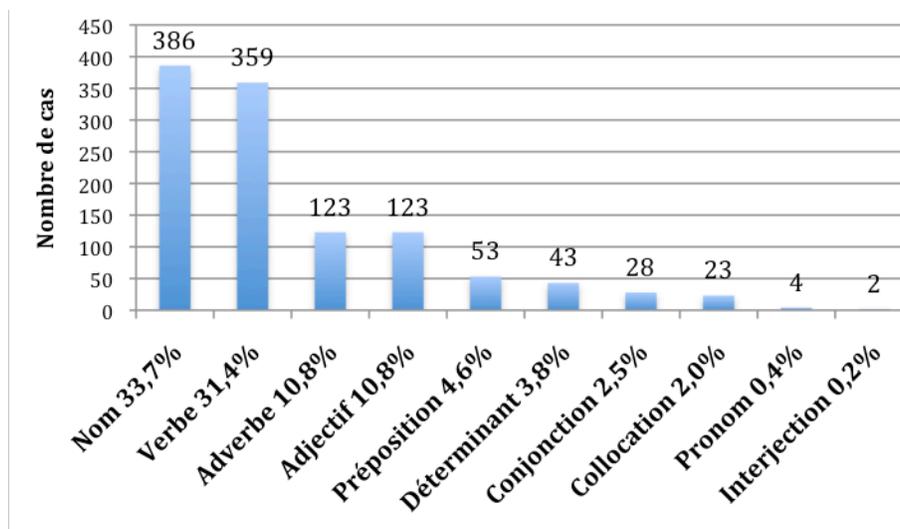


Figure 14. Répartition des erreurs en fonction des classes de mots

Sans grande surprise, nous constatons que les noms et les verbes constituent les deux classes de mots les plus touchées par les erreurs lexicales. Ceci s'explique sans doute par la fréquence élevée de ces types d'UL et par le fait qu'ils correspondent à des classes ouvertes, c'est-à-dire des classes de mots qui, contrairement aux conjonctions ou aux déterminants par exemple, contiennent un nombre imposant d'UL; par conséquent, il est normal de ne pas les maîtriser toutes. Dans des proportions égales, les adverbes et les adjectifs représentent un peu plus du cinquième de notre corpus (21,6 %). Suivent finalement les erreurs associées à des UL appartenant à des classes fermées, plus fonctionnelles, qu'on désigne parfois par le terme *mots outils*: les prépositions (4,6 %), les déterminants (3,8 %), les conjonctions (2,5 %), les pronoms (0,4 %) et les interjections (0,2 %). Si nous avons choisi d'ajouter une classe de mots *collocation* lors du codage, c'est parce que certains problèmes nous paraissaient associés directement à la combinaison de deux UL plutôt qu'à une seule (par exemple, la collocation *prendre une marche*, qui est considérée comme un anglicisme, alors que ni PRENDRE ni MARCHÉ ne constituent des anglicismes).

Les proportions présentées plus haut sont celles du corpus excluant les erreurs d'orthographe et de flexion. Par contre, si nous analysons l'ensemble des 2720 erreurs, nous constatons que les proportions demeurent très similaires (voir tableau ci-dessous), à la différence près que les verbes deviennent la classe de mots la plus fréquente; ceci s'explique sans doute par le fait que 141 des 150 problèmes flexionnels exclus de notre analyse concernaient des verbes. Il semble donc que les erreurs orthographiques affectent les différentes classes de mots presque dans les mêmes proportions que les erreurs lexicales en général.

Tableau 16. Comparaison des proportions de classes de mots avec et sans orthographe et flexion

	Pourcentage des 1144 erreurs (sans orthographe ni flexion)	Pourcentage des 2720 erreurs (avec orthographe et flexion)
Nom	33,7	32,8
Verbe	31,4	33,6
Adverbe	10,8	12,1
Adjectif	10,8	11,8
Préposition	4,6	4,0
Déterminant	3,8	2,1
Conjonction	2,4	1,6
Collocation	2,0	0,9
Pronom	0,4	0,8
Interjection	0,2	0,4

Voyons maintenant si certains types d'erreurs lexicales sont plus susceptibles d'apparaître dans certaines classes de mots. Le tableau et la figure ci-dessous représentent la façon dont se répartissent les erreurs de chaque classe de mots dans les grandes classes d'erreurs de notre typologie.

Tableau 17. Répartition des classes de mots dans chaque grande classe d'erreurs (pourcentages)

	Sens	Forme	Combinatoire pragmatique	Combinatoire syntaxique	Combinatoire morphologique	Combinatoire lexicale
Nom	25,7	19,5	33,8	9,4	70,3	74,8
Verbe	41,5	30,7	30,3	55,5	3,6	7,1
Adverbe	9,6	19,5	11,3	6,3	11,7	0
Adjectif	13,2	14,1	7,7	15,6	3,6	3,9
Préposition	7,0	5,8	5,6	0,8	1,8	0,8
Déterminant	0,3	8,3	7,2	0	7,2	0
Conjonction	1,8	1,2	2,1	11,7	0	0
Collocation	0,6	0,4	1,5	0	0	13,4
Pronom	0	0,4	0	0,8	1,8	0
Interjection	0,3	0	0,5	0	0	0
Total	100	100	100	100	100	100

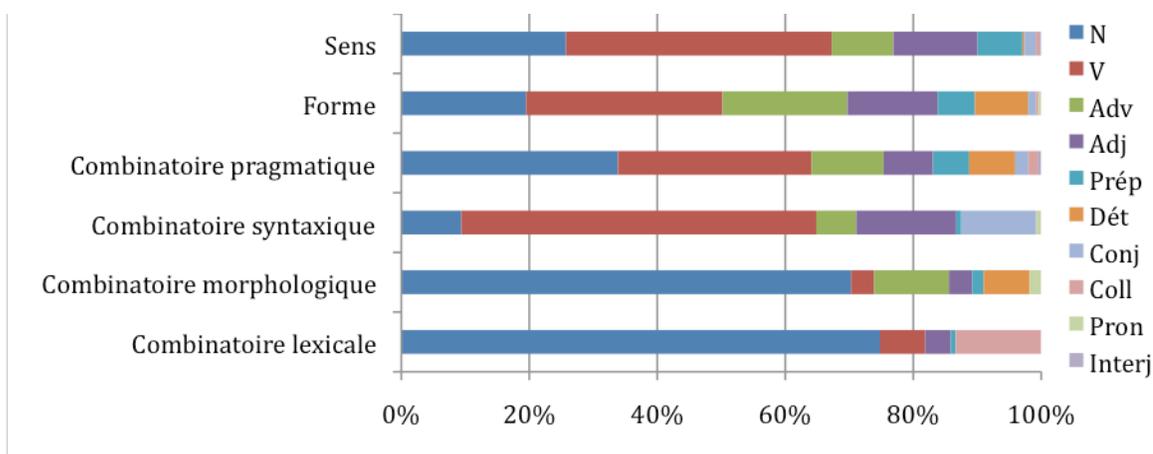


Figure 15. Proportion des classes de mots pour chaque grande classe d'erreurs

Évidemment, les problèmes de sens sont associés à des UL appartenant aux classes de mots les plus pleines sémantiquement; les verbes représentent 41,5 % des erreurs, les noms en constituent 25,7 %, alors que les adjectifs (13,2 %), les adverbes (9,6 %) et les prépositions (7 %) complètent presque entièrement le portrait.

Parmi les erreurs formelles, et en excluant celles liées à l'orthographe et à la flexion, ce sont principalement les verbes (30,7 %), les noms (19,5 %), les adverbes (19,5 %) et les adjectifs (14,1 %) qui posent problème. Si nous avons aussi considéré les erreurs orthographiques et flexionnelles, le pourcentage d'erreurs associées à des noms aurait presque rejoint celui des erreurs touchant les verbes – respectivement 30,4 % et 34,7 %.

Les erreurs de registres de langue affectent dans des proportions presque égales les noms (33,8 %) et les verbes (30,3 %); suivent les adverbes (11,3 %) et les adjectifs (7,7 %). Les 7,2 % d'erreurs de registres concernant des déterminants correspondent presque exclusivement à l'emploi du déterminant PLEIN DE (ex. *J'ai appris **plein de** nouvelles choses.* [245]).

Les erreurs lexicosyntaxiques concernent principalement les verbes (55,5 %) et présentent la particularité de toucher plus souvent les adjectifs (15,6 %) et les conjonctions (11,7 %) que les noms (9,4 %). Les erreurs lexicomorphologiques, quant à elles, présentent une très forte proportion de problèmes liés à des noms (70,3 %), ce qui découle naturellement des propriétés visées par cette classe d'erreurs, principalement le genre nominal et l'invariabilité. On observe aussi un certain nombre d'erreurs concernant des adverbes (11,7 %) et des déterminants (7,2 %).

Finalement, les erreurs de combinatoire lexicale, comme nous l'avons vu plus haut, concernent principalement des collocations dont la base est un nom (74,8 %). Bien sûr, cela ne veut pas dire que les erreurs relevées ne mettent pas en jeu des verbes et des adjectifs, qui sont les collocatifs les plus fréquents auprès des noms; une analyse approfondie des collocations erronées nous a d'ailleurs permis d'observer quels patrons de collocations sont les plus problématiques.

IMPLICATIONS DIDACTIQUES

Certaines pistes didactiques découlent de ces observations sur les classes de mots les plus touchées par les différents types d'erreurs.

Tout d'abord, nous constatons que le travail sémantique en classe – par exemple la comparaison de termes « synonymes » ou la rédaction de définitions – devrait porter en priorité sur les verbes et les noms, puis sur les adjectifs. L'ordre de priorité devrait être le même pour les problèmes formels, mais nous rappelons qu'il est plus difficile de bâtir des activités pour ce type d'erreurs.

Les chiffres présentés plus haut nous convainquent aussi de l'importance à accorder à la notion de régime verbal, qui donne lieu à de nombreuses erreurs. Il est primordial que l'enseignant attire l'attention des élèves sur les prépositions régies par les verbes et qu'il les sensibilise au fait que chaque verbe contrôle une structure syntaxique qui lui est propre, et qu'il l'impose à la phrase dans laquelle on l'utilise.

La notion de genre nominal est aussi en jeu dans plusieurs erreurs. Les élèves savent bien sûr que les noms ont un genre en français, mais il serait sans doute pertinent de le leur rappeler régulièrement afin qu'ils doutent plus facilement du genre des noms commençant par une voyelle et qu'ils n'hésitent pas à le vérifier dans les ouvrages. Des activités décrochées peuvent aussi être réalisées pour ces noms au genre plus « difficile ». En ce qui concerne l'invariabilité, il est moins aisé de proposer des activités, mais il pourrait être intéressant de faire remarquer aux élèves que des noms apparaissent très rarement au pluriel ou au singulier dans certaines de leurs acceptions (ex. ACTION au sens de 'déroulement des événements d'un récit, intrigue' et APPLAUDISSEMENTS). Pour les erreurs lexicomorphologiques, nous croyons qu'il serait aussi pertinent de réserver dans la correction des annotations particulières qui mettraient bien en évidence leur source lexicale plutôt que de les regrouper systématiquement avec les erreurs d'orthographe grammaticale : nous rappelons en effet que les élèves qui commettent ce genre d'erreurs appliquent généralement bien les règles d'accord, mais à partir d'informations lexicales erronées (ex. le mauvais genre d'un nom).

L'attention didactique portée à la dimension collocationnelle des UL devrait porter en priorité sur les collocations avant pour base des noms. Comme nous l'avons vu plus haut, les erreurs de collocation mettent surtout en jeu des collocations verbales (Oper, Func, Real) et des collocations adjectivales (Magn, AntiMagn). Il faudrait réaliser en classe des activités de sensibilisation au phénomène de collocation et d'autres visant plus spécifiquement l'accroissement du stock collocationnel associé à des UL particulières, par exemple des UL liées à une thématique faisant l'objet d'un projet d'écriture : en travaillant le conte, on pourrait se concentrer sur les collocatifs de FORÊT, SORCIÈRE, PRINCESSE, ROYAUME, etc. Montrer aux élèves qu'il existe des ouvrages nous aidant à enrichir nos textes à l'aide de combinaisons de « mots » plus recherchées et leur enseigner à les utiliser représenterait aussi pour eux un atout.

2.5 PORTRAIT DES ERREURS ORTHOGRAPHIQUES DU CORPUS

Comme nous l'avons constaté d'emblée, les erreurs orthographiques sont les plus fréquentes de notre corpus et représentent à elles seules 52,4 % des 2720 erreurs relevées. La présente section s'intéresse spécifiquement à ces problèmes. Nous rappelons par contre que l'objectif de notre thèse n'étant pas de fournir une description détaillée de ce type d'erreurs, nous nous contenterons d'en dresser un portrait plus succinct, tout en sachant que ces quelques centaines d'erreurs pourraient faire l'objet d'une analyse plus fine.

Notre corpus comprend 1426 erreurs d'orthographe. En fait, il serait plus juste de dire que 1426 mots-formes du corpus sont affectés par des erreurs d'orthographe, puisque certains en présentaient plus d'une. Seulement, nous avons comptabilisé comme une seule erreur un mot-forme qui contenait plus d'une erreur orthographique, puisque la maîtrise d'une seule dimension de l'UL est en jeu : la forme.

La majorité des 1426 mots-formes mal orthographiés (93,1 %) ne présentaient qu'une seule erreur, mais 6,4 % en contenaient deux, et 0,5 % en contenaient trois. Nous pouvons donc dire que les 1426 erreurs relevées totalisent 1528 anomalies par rapport à la norme orthographique. La figure de la page suivante montre la répartition de ces anomalies selon les dimensions touchées.

Le quart (24,9 %) des erreurs orthographiques concernent un mauvais choix de graphème parmi plusieurs correspondant au même phonème (ex. *évantualités* [9]; *chemain* [77]). Dans la très grande majorité des cas, ce choix malheureux de graphème n'affectait pas la prononciation du mot-forme affecté; par contre, dans certains cas, nous avons choisi de classer certaines erreurs affectant la forme orale du mot-forme parmi les erreurs orthographiques lorsqu'il nous semblait très improbable qu'elle révèle réellement une méconnaissance de la forme orale. Par exemple, nous sommes convaincu que l'élève qui écrit *fonsait* [103] pour *fonçait* ou *bagette* [106] pour *baguette* aurait bien prononcé ces mots-formes. Il aurait donc été incohérent de classer ces erreurs parmi les barbarismes sous prétexte qu'elles affectent la forme orale du mot-forme en jeu, puisqu'elles témoignent selon nous

de simples lacunes orthographiques et non d'une méconnaissance profonde de la forme associée à ces UL¹⁵⁰. Une des erreurs les plus fréquentes relativement à un mauvais choix de graphème est la terminaison en *-aient* pour les verbes au passé simple à la troisième personne du pluriel (ex. *sautaient* [30]; *laissaient* [188]; *acceptaient* [284]).

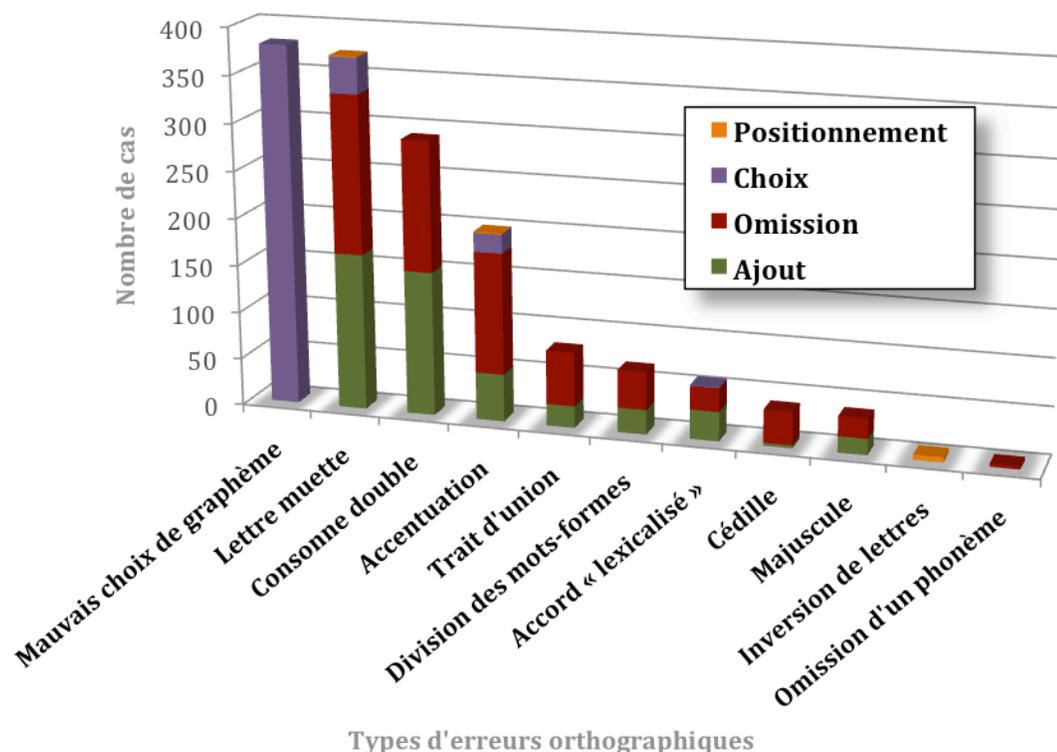


Figure 16. Répartition des 1528 erreurs orthographiques

Un autre quart des erreurs orthographiques (24,3 %) concernent les lettres muettes. Nous avons constaté beaucoup de cas d'omission et d'ajout de lettres muettes (ex. *longtemp-* [12]; *faire parti- de* [222]; *étuit* [95]; *odeures* [295]). Il arrive aussi que les élèves choisissent une lettre muette erronée (ex. *dedant* [30]; *[le] crusifie* [183]) ou encore qu'il positionne la lettre muette au mauvais endroit (ex. *entousiasthme* [300]).

¹⁵⁰ Nous avons suivi le même raisonnement pour les trois erreurs d'omission d'un phonème, qui s'apparentaient selon nous davantage à un oubli de la part de l'élève qu'à une méconnaissance de la forme de l'UL utilisée. Il s'agit des erreurs **prend* [25] pour *prendre*, **checher* [73] pour *chercher* et **patait* [97] pour *partait*.

Les consonnes doubles sont en jeu dans 19 % des erreurs d'orthographe, soit parce que les élèves, de façon à peu près égale, les omettent (ex. *accomplit* [106]; *endomagée* [164]) ou en ajoutent (ex. *ammorça* [16]; *bouille* [250]). Treize pour cent des erreurs d'orthographe ont à voir avec les accents, qui sont souvent omis (ex. *grace* à [72]; *melodieux* [174]), mais parfois ajoutés (ex. *môtel* [42]; *sélection* [155]), mal placés (ex. *fôrets* [189]) ou mal choisis (ex. *pèdestre* [31]; *requête* [164]). Nous avons l'intuition que beaucoup des cas d'omission d'accent peuvent être attribués à une certaine négligence de la part des élèves.

Les autres sous-classes d'erreurs orthographiques sont plus marginales et affectent au plus 5,2 % du total des 1528 erreurs. Les traits d'union sont le plus souvent omis (ex. *rendez vous* [80], *c'est à dire* [221, 234, 235]), mais parfois ajoutés (ex. *plus-tard* [2, 14], *tout-à-coup* [116, 175]), alors que l'omission caractérise les erreurs relatives à l'emploi de la cédille, la plupart du temps dans les formes verbales conjuguées (ex. *aperçurent* [11, 67, 78, 174], *lanca* [10, 26, 79, 102, 166, 119]), mais aussi dans certains noms (ex. *glacon* [96], *façon* [137]). Nous avons observé un certain nombre de problèmes concernant la segmentation des mots-formes : dans la majorité des cas, l'élève fusionnait des mots-formes distincts (ex. *nulpart* [66], *parceque* [167, 288]) alors que, plus rarement, certains étaient divisés (ex. *en suite* [55], *l'orsqu'il* [94]). Les erreurs en lien avec l'emploi de la majuscule affectaient principalement des UL correspondant à des noms de peuples (ex. *[des] japonais* [38]), de pays, de régions ou de villes (ex. *australie* [42], *estrie* [272], *athène* [181]), de langues (ex. *[ne parler qu'en] Anglais* [285]), de mois de l'année (ex. *[de] Septembre à Juin* [209]) et de points cardinaux (ex. *[au] Nord du Tsab* [73]). Les quelques erreurs d'inversion de lettres touchaient des mots-formes à l'orthographe un peu difficile (ex. *acceuillir* [168, 169], *curriculum vitea* [47]).

Une courte remarque mérite d'être faite à propos de la sous-classe *accord lexicalisé*. Nous avons classé sous cette rubrique les problèmes d'accord qui touchent le nombre du nom dans des locutions pour lesquelles il n'est pas toujours possible de l'établir de façon « logique ». Donnons en exemples les expressions suivantes : *en détails* [259] pour EN DÉTAIL, *sans doutes* [299] pour SANS

DOUTE, à *genou* [80] pour À GENOUX, *au dépent de* [55] pour AUX DÉPENS DE. Ces erreurs sont intimement liées à l'orthographe grammaticale, mais elles demeurent lexicales dans la mesure où il faut connaître le bon nombre du nom pour pouvoir utiliser une expression en conformité avec la norme des dictionnaires.

IMPLICATIONS DIDACTIQUES

Nous rappelons encore une fois que notre thèse ne vise pas à fournir une description approfondie des erreurs orthographiques et nous serons par conséquent succinct en ce qui concerne les pistes didactiques associées à l'orthographe lexicale.

Tout d'abord, comme les erreurs les plus fréquentes sont celles attribuables à un mauvais choix de graphème pour transcrire un phonème particulier, nous croyons que la clé du succès réside dans la vérification de l'orthographe dans le dictionnaire. Il faut donc encourager les élèves à vérifier systématiquement l'orthographe d'une UL pour laquelle ils ont le moindre doute. Il est aussi utile de savoir faire des liens entre les différentes UL d'une famille de mots morphologique lorsqu'on doute de l'orthographe d'un mot-forme. Les consonnes doubles et les lettres muettes devraient constituer pour les élèves des indices les poussant à vérifier l'orthographe des mots-formes qu'ils utilisent. Bien sûr, ceci est surtout vrai dans le cas où l'élève écrit effectivement une lettre muette ou une consonne double, correcte ou non; dans le cas où il les omet, il est moins évident pour lui que le mot-forme en question constitue un « danger ».

Nous croyons que beaucoup des cas d'omission d'accent sont attribuables à une négligence de la part des élèves (c'est du moins le cas chez nos étudiants universitaires!). Relever systématiquement ce type d'erreurs et bien faire ressortir les difficultés de lecture qui peuvent en découler conduira peut-être les élèves à faire preuve de plus de vigilance et à ne pas minimiser l'importance des accents lorsqu'ils écrivent.

Finalement, on peut demander aux élèves de consigner dans un aide-mémoire lexical les mots fréquents pour lesquels ils éprouvent des difficultés orthographiques.

2.6 RÉFLEXIONS SUR LA DESCRIPTION DES ERREURS

Maintenant que nous avons analysé 300 textes et plus de 2700 erreurs, nous jugeons approprié de faire quelques réflexions sur la méthode de description retenue et l'outil principal ayant servi au classement des erreurs : notre typologie des erreurs lexicales.

2.6.1 L'IDENTIFICATION DES ERREURS

L'une des plus grandes difficultés associées à l'analyse d'erreurs lexicales demeure de tracer une limite entre *erreur* et *variation lexicale*. En effet, nous savons que la dimension lexicale est sans aucun doute celle qui laisse le plus de latitude au locuteur. Le style d'un bon scripteur doit souvent beaucoup à sa capacité à jouer avec les « mots », à en étendre le sens, à oser, pour faire image ou pour obtenir des effets de style, des combinaisons inusitées et surprenantes. La limite entre *création lexicale* et *erreur* – que nous définissons comme *une forme linguistique ou une combinaison de formes linguistiques qui diffère de ce qu'un locuteur expert aurait selon toute probabilité produit dans un même contexte de production* (voir IV-1.4) – est donc mince. Le danger qui nous guettait, tout au long de notre analyse, était d'adopter une approche trop normative, voire castratrice, qui nous aurait conduit à considérer comme des erreurs certaines créations des élèves, sous prétexte qu'elles sont « déviantes » par rapport à la norme. Nous croyons d'ailleurs qu'une telle attitude de la part d'un enseignant aurait un effet catastrophique sur ses élèves : au contraire, nous voulons encourager les élèves à « prendre des chances » avec le vocabulaire, à jouer avec les mots, à faire preuve d'originalité lexicale.

Nous voulons donc rassurer nos lecteurs : malgré la tendance éminemment normative inhérente à tout travail sur l'erreur, nous croyons avoir réussi à analyser notre corpus en gardant une attitude d'ouverture relativement à la création lexicale. Par ailleurs, il ne faut pas oublier que nous avons lu les 300 rédactions en entier et qu'il nous était donc possible, pour chacune, de juger de façon intuitive du niveau de maîtrise lexicale de l'élève l'ayant écrite, ce qui nous

permettait d'avoir un regard plus éclairé sur les emplois plus « originaux » : la plupart du temps, il nous était possible de voir s'ils étaient maîtrisés ou s'ils résultaient plutôt d'un manque de vocabulaire. Nous avons d'ailleurs rencontré plusieurs formulations assez habiles de la part de certains élèves. Comment ne pas sourire lorsqu'un élève écrit *Devant leur incapacité totale à chanter, le metteur en scène fait une scène!* [138] ou qu'un autre, dont le texte a pour personnage principal un renard, dit de celui-ci qu'il *prend son courage à deux pattes* [205]? De telles tournures de phrases témoignent d'une capacité à jouer avec la langue et il aurait été absurde de les considérer comme erronées. Par ailleurs, nous avons rencontré plusieurs termes, tant dans les textes réussis que dans les moins bons, qui démontrent que les élèves, malgré certaines lacunes, connaissent aussi de « beaux mots » : OPALESCENT, DISSIPER, **s'engager** dans un tunnel, un froid **de loup**, l'**enceinte** du château, une odeur **pestilentielle**, etc.

Par contre, parallèlement à ces « succès lexicaux », plusieurs emplois nous semblaient nettement moins heureux, et c'est sur ceux-là que nous avons concentré nos analyses. Bien sûr, certains cas « limites », à propos desquels il est plus difficile de trancher (ex. *un sourire accroché aux deux oreilles* [295]), pourraient faire l'objet d'avis divergents de la part de correcteurs différents. Il ne fait aucun doute que toutes les erreurs de notre corpus n'auraient peut-être pas été considérées comme telles par d'autres analystes; néanmoins, nous croyons qu'avec les balises que nous nous étions fixées (utilisation de dictionnaires, recours à des moteurs de recherche) et l'attitude d'ouverture dont nous avons fait preuve tout au long de notre travail, ces cas sont somme toute assez marginaux.

2.6.2 LE DOUBLE CODAGE DES ERREURS

Comme nous l'avons vu dans notre méthodologie (VII-3.1.2), certains chercheurs ayant travaillé sur l'analyse d'erreurs lexicales ne voient pas d'inconvénients au double codage des erreurs (Zimmermann, 1986 ; Agustín Llach, 2005a ; Milićević et Hamel, 2007), c'est-à-dire le fait d'associer une même erreur à deux classes de la typologie utilisée. Dans le cadre de cette thèse, nous avons préféré classer chaque erreur dans une seule des classes de notre typologie. Cette décision a été

prise à postériori, une fois que nous avons constaté le nombre limité de cas ayant fait l'objet d'une double description ; en effet, sur les 2720 erreurs recueillies, nous avons attribué deux étiquettes à seulement une trentaine de cas. Nous avons remarqué que deux cas de figure nous amenaient à proposer un double codage.

Le premier est celui où nous avons de la difficulté à choisir entre deux codes similaires, par exemple S-Syn (synonyme) ou S-Proc (sens proche); une telle hésitation ne devait pas selon nous subsister dans les résultats finaux; nous avons donc repris ces cas et les avons comparés à des erreurs ayant reçu l'une ou l'autre des étiquettes avant de choisir de classer l'erreur parmi une seule des classes de la typologie.

L'autre raison pouvant expliquer un double codage est le fait qu'une même erreur peut être classée à deux endroits différents, selon notre façon d'appréhender le problème. Par exemple, on peut considérer qu'un élève qui emploie *ROCHE* au sens de 'caillou' commet une erreur sémantique, puisque *ROCHE* désigne un 'bloc de matière minérale' et qu'un caillou est un 'petit fragment de pierre'. Par ailleurs, on pourrait plutôt dire qu'il s'agit d'un problème de registre de langue, puisque dans le vocabulaire québécois familier, *ROCHE* est un synonyme de *CAILLOU*, ce qu'indiquent d'ailleurs les dictionnaires. Dans de tels cas, il nous semblait toujours qu'une des façons d'envisager le problème était plus plausible du point de vue de l'élève, compte tenu de ses connaissances. Dans ce cas par exemple, il est évident que l'élève connaît exactement le sens familier de *ROCHE* (= 'caillou') et il nous paraît étrange de dire qu'il commet ici une erreur sémantique; classer cette erreur parmi les problèmes de registre de langue nous semble plus approprié. Pour tous les cas similaires, nous avons donc conservé seulement le classement qui nous apparaissait le plus logique, ce qui fait que nous avons pu éviter le double codage.

Outre le fait de faciliter le traitement des données aux fins de notre analyse (une erreur = un seul code), nous croyons qu'éliminer le double codage a aussi permis de produire un codage plus homogène. En effet, il est très difficile d'être

systématique lorsqu'on se donne la possibilité d'attribuer plus d'un code à chaque erreur, et comme notre travail ne faisait pas l'objet d'un contre-codage, nous croyons qu'il était plus prudent de nous en tenir à un seul code par erreur.

2.6.3 LA TYPOLOGIE DES ERREURS LEXICALES

La typologie que nous avons utilisée pour classer les erreurs constitue l'un des apports importants de notre travail. Les améliorations que nous avons apportées par rapport à la grille proposée dans notre mémoire de maîtrise (Anctil, 2005), entre autres à la lumière des différents travaux consultés (voir V-2) se sont révélées utiles, puisqu'elle s'est avérée opérationnelle pour classer l'ensemble des erreurs rencontrées dans notre corpus. En effet, aucune classe d'erreurs n'a dû être ajoutée en cours d'analyse; en fait, seule la classe d'erreurs résiduelles parmi les erreurs en lien avec la combinatoire grammaticale aurait pu être subdivisée, mais le nombre limité d'erreurs s'y retrouvant (3) ne justifie pas selon nous l'ajout de classes supplémentaires.

Un des avantages de notre typologie est aussi de ne contenir aucune hypothèse quant aux sources possibles des erreurs qu'elle sert à classer, problème fréquent dans beaucoup de grilles d'analyse d'erreurs. Ceci nous aura notamment permis de bien distinguer les niveaux de description et d'explication des erreurs de façon à éviter les chevauchements souvent reprochés aux méthodologies employées dans les travaux sur l'analyse d'erreurs.

Notre choix d'opter pour une approche holistique dans le classement des problèmes sémantiques, comme le suggéraient Zimmermann (1986) et Milićević et Hamel (2007), nous a aussi permis de classer beaucoup plus aisément ce type d'erreurs, sans avoir à systématiquement faire ressortir des composantes sémantiques précises. Bien sûr, la distinction entre les erreurs de sens proches (S-Proc) et celles dans lesquelles le sens visé est synonyme de celui de l'UL erronée (S-Syn) n'étaient pas toujours faciles à départager, tout comme les erreurs de typage des actants (S-Act), qui peuvent souvent mettre en jeu des synonymes. Néanmoins, ces trois types de problèmes nécessitent des interventions didactiques similaires, et il n'est pas réellement gênant de ne pas arriver à les

différencier facilement. Cette division des problèmes sémantiques, particulièrement l'ajout de la classe S-Proc, nous aura évité de retomber dans le piège auquel nous avons dû faire face dans notre maîtrise : le fait que la classe *impropriété* (S-Fict) devienne une classe fourretout pour l'ensemble des problèmes sémantiques que nous avons de la difficulté à expliquer. Malgré tout, les erreurs sémantiques demeurent les erreurs lexicales les plus difficiles à classer et mériteraient de faire plus spécifiquement l'objet d'un travail approfondi, peut-être à partir d'un corpus plus grand d'erreurs.

Une dernière remarque mérite d'être faite à propos de notre typologie. Malgré l'intérêt indéniable des observations qu'elle permet et des pistes didactiques que suggèrent ces dernières, la typologie des erreurs lexicales que nous proposons dans cette thèse demeure un outil de recherche et non pas un outil d'aide au diagnostic des erreurs lexicales qui pourrait être utilisé en l'état par les enseignants. Cet outil est complexe et demande une bonne compréhension de plusieurs concepts linguistiques; c'est d'ailleurs ce qui a justifié que nous ne fassions pas contre-coder le volet descriptif de notre analyse. Pour que les enseignants puissent l'utiliser en classe, tout un travail de simplification et d'explicitation devrait être fait sur la typologie. Nous croyons par contre qu'une version allégée de cet outil pourrait s'avérer intéressante pour les enseignants et avons bien l'intention de nous atteler à la tâche dans un avenir rapproché!

3. LA GRAVITÉ DES ERREURS

La présente section s'intéresse à l'évaluation de la gravité des erreurs du corpus. Rappelons que cette évaluation reposait sur trois critères (voir VII-3.2) : l'acceptabilité linguistique selon les ouvrages de référence, la compréhensibilité et la perception de l'intégration de l'erreur à l'usage. Comme deux de ces critères ont fait l'objet d'un contre-codage pour une partie des erreurs, nous commencerons par en présenter les résultats avant d'exposer ceux relatifs à l'ensemble du corpus.

3.1 LE CONTRE-CODAGE ET L'ÉVALUATION DE LA GRAVITÉ DES ERREURS

Des trois critères retenus pour l'évaluation, nous avons choisi d'en faire contre-coder deux : la compréhensibilité et la perception de l'intégration de l'erreur à l'usage, soit les deux critères les plus subjectifs. Pour alléger la tâche des contre-codeuses, nous avons omis le critère d'acceptabilité linguistique; comme la valeur attribuée pour ce critère était déterminée en fonction de ce que nous trouvions dans les ouvrages de référence, nous avons jugé que le contre-codage aurait fourni les mêmes résultats dans la mesure où les ouvrages consultés auraient été les mêmes.

Le contre-codage portait sur 130 erreurs, soit un dixième de notre corpus sans les erreurs d'orthographe (1294 erreurs). Les erreurs étaient présentées aux contre-codeuses dans un fichier *Excel*, où elles apparaissaient dans leur contexte avec la correction que nous avons proposée. Les contre-codeuses devaient attribuer à chacune un score de 0 à 2 pour la compréhensibilité ainsi que pour l'intégration à l'usage (critères 2 et 3). Elles devaient aussi proposer une explication sur la source de l'erreur lorsque cela leur semblait possible; les résultats relatifs à l'explication des erreurs seront présentés à la section 4.1.

Nous établirons maintenant les taux d'accord interjuges, de façon à voir dans quelle mesure notre propre évaluation de la gravité correspond à celle qu'auraient fournie d'autres codeurs. Nous parlerons d'*accord complet* lorsque le

score attribué pour un critère par les deux contre-codeuses est le même que le nôtre. Par *accord partiel*, nous faisons référence au cas de figure où notre score est le même que celui de seulement une des contre-codeuses; dans le cas où la contre-codeuse « dissidente » a attribué un score se situant à l'autre bout du spectre de valeurs par rapport au score des deux autres codeurs (soit 0 plutôt que 2 ou vice versa), nous parlerons d'*accord partiel avec écart maximal*. Finalement, nous emploierons le terme *désaccord* pour faire référence aux cas où le jugement des deux contre-codeuses diffère de celui du chercheur principal.

Le tableau qui suit présente les taux d'accord inter-juges pour les deux critères évalués.

Tableau 18. Accord inter-juges pour les critères de compréhensibilité et de perception de l'intégration de l'erreur à l'usage

	Accord complet 3 valeurs identiques		Accord partiel 2 valeurs identiques sur 3, dont celle du chercheur principal			Désaccord valeur du chercheur principal différente des 2 autres ou 3 valeurs différentes	
	nb cas	%	nb cas	%	Écart maximal	nb cas	%
Compréhensibilité	82	63	36	27,7	0	12	9,2
Perception de l'intégration à l'usage	42	32,3	54	41,5	1	34	26,2

On constate que pour la compréhensibilité, 63 % des cas ont donné lieu à un accord complet entre les trois juges, alors qu'il y a accord partiel dans 27,7 % des cas. Finalement, la valeur attribuée par le chercheur principal différait de celles données par les contre-codeuses dans 9,2 % des cas. Il est important de spécifier qu'aucun des 36 cas correspondant à un accord partiel ne présentait un écart maximal entre les jugements identiques du chercheur principal et d'une des contre-codeuses et celui de l'autre contre-codeuse, par exemple un cas où l'une des codeuses et nous aurions jugé que l'erreur ne compromet en rien la compréhension (0) alors que l'autre codeuse considèrerait qu'elle pose un problème grave de compréhension (2). Nous pourrions donc dire que l'accord interjuges pour la compréhensibilité est soit complet soit satisfaisant (accord partiel sans écart maximal) dans 90,7 % des cas.

En ce qui concerne la perception de l'intégration de l'erreur à l'usage, l'accord interjuges est moins élevé, ce qui était prévisible vu le caractère très subjectif de ce critère. Seulement un tiers des cas (32,3 %) ont donné lieu à un accord complet, tandis que 41,5 % ont conduit à un accord partiel, dont 1 seul cas présentait un écart maximal entre le jugement du chercheur principal et celui des contre-codeuses (nous considérons une des erreurs comme complètement déviante par rapport à l'usage [2], alors qu'elles le percevaient comme intégré à l'usage [0]¹⁵¹). Notre jugement différait de celui des deux contre-codeuses dans un peu plus du quart des cas (26,2 %), ce qui aurait pu remettre en doute notre évaluation de la gravité des erreurs si nous n'avions retenu que ce critère. Malgré tout, nous pouvons dire que l'accord interjuges pour la perception de l'intégration de l'erreur à l'usage est soit complet soit satisfaisant (accord partiel sans écart maximal) dans 73,1 % des cas.

Nous avons aussi mesuré dans quelle mesure nos jugements globaux de gravité – c'est-à-dire la somme des trois critères – divergent par rapport à ceux des contre-codeuses. Pour ce faire, nous avons, pour chacune des erreurs, fait la moyenne des scores attribués à chaque critère par les contre-codeuses. Nous avons ensuite additionné ces scores moyens à la valeur que nous avons attribuée à chaque erreur pour l'acceptabilité linguistique, ce qui nous procurait un score sur 6 reflétant l'évaluation moyenne des deux contre-codeuses. Nous avons finalement établi l'écart entre notre propre score et celui des contre-codeuses. L'écart-type est de 0,67 et les écarts se répartissent comme suit :

¹⁵¹ Avec le recul, nous trouvons notre jugement un peu sévère et pensons qu'il aurait été plus approprié d'attribuer la valeur 1 dans ce cas.

Tableau 19. Écart entre l'évaluation de la gravité des contre-codeuses et celle du chercheur principal

	nb cas/130	%
Scores identiques	26	20
Écart de 0,5	58	44,6
Écart de 1	29	22,3
Écart de 1,5	12	9,2
Écart de 2	3	2,3
Écart de 2,5	2	1,5

Bien que seulement 20 % des cas aient donné lieu à des scores identiques de la part du chercheur principal et des contre-codeuses, nous constatons que l'écart était d'un point ou moins dans 86,9 % des cas. Compte tenu du caractère subjectif des critères retenus dans la présente recherche pour opérationnaliser le concept de gravité, nous jugeons qu'un écart d'un point ou moins sur 6 (16,7 % ou moins) dans 86,9 % des cas est tout à fait satisfaisant. Ce contre-codage procure donc une certaine validité à l'évaluation de la gravité des erreurs que nous avons faite sur l'ensemble du corpus, puisqu'il démontre que d'autres codeurs seraient arrivés à une évaluation semblable; nous utiliserons donc nos propres scores pour la suite de l'analyse. Bien sûr, le concept même de gravité demeure flou et difficile à évaluer; sans doute aurait-il été intéressant d'observer si une évaluation holistique des erreurs aurait produit des jugements similaires, mais là n'est pas l'objet de notre thèse.

Voyons maintenant la répartition de l'ensemble des erreurs du corpus selon leur gravité.

3.2 RÉPARTITION DES ERREURS SELON LA GRAVITÉ

Le tableau ci-dessous présente la répartition des 1144 erreurs dans les six niveaux de gravité alors que la figure 2 l'illustre de façon plus schématique.

Tableau 20. Gravité des 1144 erreurs lexicales du corpus

	1	2	3	4	5	6	Total
Sens	5	38	84	94	90	31	342
Forme	2	33	33	134	34	5	241
Combinatoire pragmatique	13	51	102	27	2	0	195
Combinatoire syntaxique	2	14	40	65	7	0	128
Combinatoire lexicale	2	11	60	40	7	7	127
Combinatoire morphologique	0	6	19	81	5	0	111
Total	24	153	338	441	145	43	1144
Pourcentage	2,1	13,4	29,5	38,5	12,7	3,8	

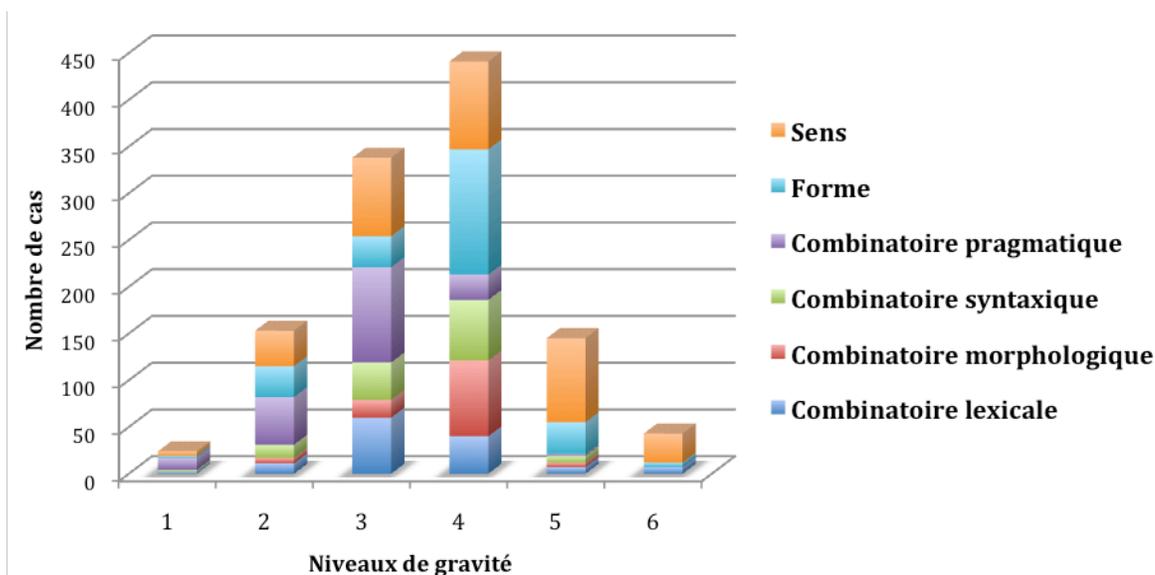


Figure 17. Gravité des 1144 erreurs lexicales du corpus

On remarque tout d'abord que la très grande majorité des erreurs (68 %) ont été considérées comme moyennement graves (niveaux 3 et 4). Le nombre d'erreurs classées comme peu graves (niveaux 1 et 2) ou plutôt graves (niveaux 5 et 6) est à peu près identique (15,5 % et 16,5 %).

Si l'on s'intéresse à la répartition sur l'échelle de gravité en fonction des types d'erreurs, certaines informations sautent aux yeux. Tout d'abord, on constate que les problèmes de combinatoire pragmatique (en mauve) reçoivent presque toujours un score inférieur ou égal à 3 sur l'échelle de gravité; cela s'explique sans doute par le fait que ce type d'erreurs n'entrave jamais la compréhension

du message et que les UL familières sont intégrées à l'usage et ont donc reçu des scores de 0 ou 1 pour la perception de l'intégration à l'usage.

Les erreurs liées aux propriétés de combinatoire morphologique, syntaxique et lexicale se concentrent de façon évidente dans les niveaux 2, 3 et 4 de l'échelle, les problèmes morphosyntaxiques occupant de façon plus marquée le niveau 4. Ceci peut être facilement expliqué, puisque les problèmes morphosyntaxiques sont principalement constitués d'erreurs de genre nominal et d'invariabilité, qui ont presque systématiquement été considérées en désaccord complet avec les ouvrages [= 2] et perçues comme très déviantes par rapport à l'usage (dans un contexte de production écrite) [= 2], sans toutefois compromettre la compréhension du message [= 0]. La norme étant plus floue relativement aux collocations, et les ouvrages beaucoup moins précis en cette matière, il est normal que les erreurs de combinatoire lexicale (principalement l'emploi de collocatifs douteux) aient reçu un score plus faible pour l'acceptabilité linguistique (souvent 1 – *ne figure pas dans les ouvrages*), ce qui explique qu'elles se concentrent au niveau 3.

Les problèmes sémantiques sont présents à tous les niveaux, mais sont plus souvent considérés comme moyennement graves ou plutôt graves; de façon générale, plus l'écart sémantique est grand entre le sens visé et celui de l'UL utilisée, plus l'erreur est grave (sens fictif > sens proche > synonyme). Les pléonasmes et les anglicismes sont moyennement graves (niveaux 2, 3 et 4), encore une fois parce qu'ils ne constituent pas généralement des obstacles à la compréhension. Finalement, les problèmes formels couvrent aussi l'ensemble du spectre de gravité; les problèmes formels les plus graves sont les signifiants inexistantes qui ne découlent pas d'une mauvaise dérivation et les cas de déformation du signifiant qui produisent une autre UL existante (ex. *vit* au lieu de *fit*, *au bout* pour *à bout*).

3.3 RÉPARTITION DES ERREURS SELON CHACUN DES CRITÈRES

Le tableau 21 présente la répartition des 1144 erreurs selon chacun des trois critères utilisés dans l'évaluation de la gravité, ce qui permet certaines observations intéressantes, présentées dans les sections qui suivent. Pour faciliter la lecture du Tableau 21, nous recopions l'échelle de gravité utilisée pour le codage et présentée au point 3.2 du chapitre VII.

Tableau 13. Échelle de gravité des erreurs

Acceptabilité linguistique selon les ouvrages		est attesté par les ouvrages	0
		n'apparaît pas dans les ouvrages	1
		est en contradiction avec les ouvrages de référence (<i>téléphoner qqn</i>) ou y est explicitement condamné (<i>*téléphone</i> au sens 'appel téléphonique')	2
Jugement du correcteur	Impact sur la compréhension	n'a aucun impact sur la compréhension	0
		peut être interprété, mais pas avec assurance	1
		compromet la compréhension	2
	Perception de l'intégration de l'erreur à l'usage	la formulation semble tout à fait naturelle en français (intégration au code, ex. anglicisme comme <i>opportunité</i>)	0
		la formulation semble maladroite et marquée par rapport à l'usage normal	1
		la formulation semble s'écarter fortement de l'usage normal	2

Tableau 21. Répartition des 1144 erreurs dans chacun des trois critères de gravité

	Acceptabilité linguistique			Compréhensibilité			Intégration au code		
	0	1	2	0	1	2	0	1	2
Sens	10	37	295	192	116	34	31	87	224
Forme	3	23	215	195	40	6	29	26	186
Combinatoire pragmatique	7	37	151	191	4	0	41	113	41
Combinatoire syntaxique	2	26	100	115	13	0	9	32	87
Combinatoire lexicale	1	76	50	100	19	8	2	20	105
Combinatoire morphologique	1	15	95	103	8	0	3	11	97
Total	24	214	906	896	200	48	115	289	740

3.3.1 L'ACCEPTABILITÉ LINGUISTIQUE

En ce qui concerne l'acceptabilité linguistique selon la norme descriptive des dictionnaires, nous constatons que les ouvrages fournissent l'information – de façon explicite ou par le biais des exemples fournis – permettant d'expliquer

79,2 % des erreurs relevées (en omettant encore une fois les erreurs orthographiques et flexionnelles; si l'on considère ces erreurs, le taux monte à 91 %). Vingt et huit dixièmes pour cent (20,8 %) des erreurs ne peuvent être expliquées à l'aide des dictionnaires retenus, soit parce que l'information requise en est absente (18,7 %), soit parce que notre intuition de correcteur va à l'encontre de ce qui apparaît dans les ouvrages (2,1 %)¹⁵².

C'est pour les erreurs de nature formelle et sémantique que les dictionnaires s'avèrent le plus utiles, contenant autour de 88 % des informations recherchées; les propriétés morphologiques des UL y sont aussi assez bien décrites, ce qui a permis de trouver réponse à 85,6 % de nos questions. Dans un peu plus des trois quarts des cas, les informations relatives à la combinatoire syntaxique et pragmatique (78,1 % et 77,4 %) ont aussi pu être trouvées. C'est pour les problèmes touchant à la combinatoire lexicale que les dictionnaires se révèlent le moins utiles : 59,8 % de nos interrogations sont demeurées sans réponse à la suite de leur consultation.

3.3.1.1 UN MOT SUR LES DICTIONNAIRES COMME OUTILS DE DIAGNOSTIC DE L'ERREUR LEXICALE

Les longues semaines durant lesquelles nous avons analysé les centaines d'erreurs de notre corpus nous ont permis de nous plonger, comme jamais auparavant, dans les ouvrages fascinants que sont les dictionnaires et de prendre le temps de nous y perdre un peu, laissant notre curiosité nous porter d'un article à l'autre ou, technologie oblige, d'un clic à l'autre. En quête de réponses pouvant étayer nos intuitions, nous avons scruté méticuleusement des centaines d'articles, comparé et confronté les ouvrages, pris la mesure de la richesse de l'information qui y est présentée, mais aussi réalisé à quel point les dictionnaires commerciaux présentent des lacunes, ne constituent qu'une actualisation extrêmement fragmentaire de la connaissance complexe qu'exige la maîtrise d'un « mot ». Voici donc quelques observations à propos de l'utilisation des dictionnaires comme outils de diagnostic des erreurs lexicales.

¹⁵² Par exemple, l'emploi de FACE dans *les traits de sa face* [139] nous semble familier, même si aucun des ouvrages consultés ne fait de remarque à cet effet.

L'INCOMPLÉTUDE DES DICTIONNAIRES : Tout d'abord, il est arrivé fréquemment que l'information que nous recherchions n'apparaisse pas dans les ouvrages consultés. Par exemple, en cherchant le terme CHASSE-MOUSTIQUES afin de vérifier l'emploi du trait d'union, nous nous sommes aperçu que cette UL n'est répertoriée dans aucun de nos dictionnaires. Pourtant, ce terme est très répandu et apparaît 57 000 fois sur *Google*, c'est-à-dire quatre fois plus que *répulsif à moustiques*, le terme équivalent recensé par les ouvrages. Par conséquent, il nous aurait semblé absurde de considérer cet usage comme une erreur, même si cette UL n'a pas droit de cité dans les dictionnaires. Même constat pour le verbe RELAXER employé de façon non pronominale (ex. *Maintenant, je vais pouvoir relaxer.* [244]), qui, même s'il appartient sans doute au registre familier, est si fréquent (plus encore que l'emploi pronominal) qu'il devrait selon nous apparaître (ou tout au moins faire l'objet d'une remarque) dans les dictionnaires. Voilà deux exemples d'UL complètement absentes des dictionnaires, mais il est encore plus fréquent de ne pas trouver de réponse à notre questionnement lorsqu'on s'intéresse à la combinatoire d'UL disposant d'une entrée. Par exemple, la construction *être motivé à* + verbe infinitif (ex. *Elle n'est pas très motivée à étudier.*) est-elle une correction acceptable pour la phrase *Sarah se lève motivée d' [à] aller jouer de la musique.* [274]? Est-ce que TRAVAIL au sens d' 'emploi' peut être utilisé à la forme plurielle (*Alex a deux travaux en même temps.*)? Est-ce que le déterminant GRAND en contexte négatif est familier¹⁵³ (ex. *Je ne mis pas grand [beaucoup de] temps à regretter d'avoir voulu cet emploi.* [267])? Si on peut dire *entreprendre une aventure*, peut-on aussi dire qu'on **a accompli** une aventure [110]? Voilà autant de questions pour lesquelles les ouvrages ne nous ont été d'aucun recours; dans tous ces cas, comme dans plusieurs autres, nous avons dû nous fier à nos propres intuitions et les valider à l'aide de moteurs de recherche du type *Google*.

¹⁵³ La grammaire d'*Antidote* comporte un article sur cet emploi de GRAND, considéré comme un déterminant parce qu'il peut être remplacé par BEAUCOUP DE. On y mentionne que dans le reste de la francophonie, cet emploi se limite aux expressions *pas grand monde* et *pas grand-chose*, mais qu'il est très productif au Québec.

LES AVIS DIVERGENTS D'UN OUVRAGE À L'AUTRE : Au-delà des nombreuses informations absentes des dictionnaires, nous avons été frappé par le nombre élevé de cas pour lesquels les ouvrages étaient en désaccord. Bien sûr, cela n'a pas représenté pour nous un immense problème, puisque nous avons choisi de statuer sur l'acceptabilité d'un usage en fonction de l'ouvrage le plus permissif pour cet usage précis. Néanmoins, nous avons trouvé curieux d'être confronté aussi souvent à des avis divergents. Ces différences touchaient principalement les marques d'usage relatives aux registres de langue et à la variation régionale. Par exemple, ÉPEURANT est considéré comme un québécoisme par les trois ouvrages consultés, mais *Le Multidictionnaire* mentionne que cet adjectif est de registre familier, alors que les deux autres ne font aucune remarque à cet effet. SACRER au sens de 'blasphémer' est marqué comme familier dans *Antidote RX*, tandis que *Le Petit Robert* et *Le Multidictionnaire* indiquent seulement qu'il s'agit d'un québécoisme, ce qu'*Antidote RX* ne spécifie pas. Nous avons d'ailleurs été surpris de constater qu'*Antidote* était parfois plus normatif que *Le Multidictionnaire* et *Le Petit Robert*, considérant comme familier des UL comme CHAUDIÈRE au sens de 'seau', ENFILER [*un vêtement*] et SE CHICANER. Dans le même ordre d'idées, *Antidote RX* spécifie qu'AIR CLIMATISÉ est une impropriété, alors que les deux autres ouvrages présentent cette UL comme un synonyme d'AIR CONDITIONNÉ. Dans d'autres cas, c'est *Le Petit Robert* qui était l'ouvrage dissident, par exemple en marquant comme familier le verbe ZAPPER. *Le Multidictionnaire* présente aussi certains avis divergents, notamment en condamnant la conjonction DE FAÇON À CE QUE, présente dans les autres ouvrages, puis en mentionnant que PUISQUE ne s'élide que devant *il(s), elle(s), en, on, un(e)* et *ainsi*, alors qu'*Antidote RX*, à la suite de Grevisse, permet l'élision partout devant les voyelles et les *h* muet. Ces quelques exemples, qui sont légion lorsqu'on prend le temps de comparer des ouvrages, permettent de prendre pleinement conscience du caractère fluctuant de la norme dans les dictionnaires et de constater que ces « bibles », qui véhiculent pour beaucoup LA vérité, sont le fruit de décisions éditoriales.

LES INFORMATIONS ALLANT À L'ENCONTRE DE NOS PROPRES INTUITIONS : Dans certains cas, plus rares tout de même, nos propres intuitions de locuteur allaient à l'encontre de l'information présentée dans les dictionnaires. À quelques reprises, nous avons dû revenir sur notre décision de considérer comme une erreur un usage qui pourtant nous semblait problématique parce que les ouvrages statuaient clairement sur son acceptabilité. En voici deux exemples.

(153) *À la toute fin de l'été et de ça dernière tonte (= 'fait de tondre le gazon'), son père est venu le chercher. [292]*

(154) *Le monde [gens] entend tellement parler du Coca-Cola qu'ils finissent par acheter du Coca n'importe où. [158]*

La phrase (153) provient d'un texte à propos d'un adolescent qui gagne de l'argent de poche en tondant des pelouses et l'emploi de TONTE y figurant nous paraissait vraiment bizarre. Pourtant, *Le Petit Robert* décrit le nom TONTE comme 'action de tondre' et donne *tonte des gazons* en exemple : difficile alors de considérer que l'élève a fait une erreur! Pourtant, encore maintenant, cette utilisation de TONTE nous paraît étrange et nous serions curieux de la soumettre à quelques locuteurs natifs afin de voir leur réaction.

Dans la phrase (154), c'est MONDE qui nous posait problème, parce que cet emploi nous semble familier. Pourtant, les ouvrages retenus décrivent MONDE comme un synonyme de GENS et ne mentionnent aucune contrainte d'utilisation. Nous sommes d'accord pour dire qu'une phrase comme *Il y avait beaucoup de monde sur la plage dimanche dernier.* est tout à fait correcte, mais lorsque MONDE au sens de 'gens' est en position sujet (*Le monde attend l'ouverture du magasin.*), ou tout simplement lorsqu'il est employé avec un déterminant défini (*Les gardiens de sécurité redirigent le monde vers les sorties.*), il nous semble appartenir au registre familier.

Dans pareils cas, s'agit-il de mauvaises intuitions de notre part ou de lacunes dans la description proposée par les dictionnaires? Ou encore – et ce scénario est rarement envisagé par l'utilisateur moyen d'un dictionnaire – s'agit-il carrément d'erreurs de la part des lexicographes? Difficile de le dire...

Nous constatons donc que le dictionnaire peut être un outil précieux dans le diagnostic d'erreurs lexicales, mais qu'il n'est pas une panacée. Même nous, malgré notre expérience lexicographique et notre connaissance approfondie de la structure de ce type d'ouvrages et la nature des informations qui y sont présentées, n'avons pas pu trouver réponse à nos questions aussi souvent que nous l'aurions souhaité. On peut donc imaginer la situation d'un élève, à qui on a rarement enseigné comment exploiter pleinement un dictionnaire, devant un gros livre aussi intimidant. Et celle des enseignants, qui, devant ces informations lacunaires, ces contradictions (si tant est qu'ils aient le temps de consulter plus d'un dictionnaire... et même un seul!) et ces informations allant parfois à l'encontre de leurs propres intuitions, ne voient peut-être pas dans le dictionnaire un allié si précieux lorsque vient le temps d'expliquer à leurs élèves leurs « erreurs de vocabulaire ». D'ailleurs, il serait intéressant de voir si les enseignants ont cette perception naïve du dictionnaire comme le détenteur de la vérité linguistique absolue ou si, au contraire, ils possèdent l'assurance linguistique nécessaire pour remettre en question ce qui entre en conflit avec leurs intuitions de professionnels de la langue. On peut aussi se demander s'ils connaissent assez bien ce type d'ouvrages pour s'y retrouver facilement et les utiliser à leur plein potentiel. Notre deuxième objectif de recherche nous donnera quelques pistes à ce propos (voir X-3.2.3).

Mais pour l'instant, revenons à l'analyse des critères de gravité en nous intéressant plus précisément à l'impact des erreurs analysées sur la compréhension du message.

3.3.2 L'IMPACT DE L'ERREUR SUR LA COMPRÉHENSION DU MESSAGE

L'impact sur la compréhension est souvent considéré comme l'un des critères ayant le plus d'influence sur les jugements de gravité des locuteurs; les travaux mentionnent aussi souvent que les erreurs lexicales sont celles qui entravent le plus la compréhension. Or, notre analyse fait ressortir que des 1144 erreurs lexicales conservées pour l'analyse, 78,3 % ne compromettent aucunement la compréhension du message, et ce chiffre ne tient pas compte des erreurs

d'orthographe et de flexion; en prenant en compte ces erreurs dans notre calcul, le taux monterait à 89,7 %. Dans 17,5 % des cas, la compréhension du message était possible, mais demandait un certain effort d'inférence; seulement 4,2 % des erreurs entravaient de façon importante la compréhension du message et apparaissaient dans un contexte qui ne permettait pas de rétablir le sens visé avec assurance¹⁵⁴. On pourrait donc dire que 21,7 % des erreurs analysées constituaient un obstacle à la compréhension. Comme on pouvait s'y attendre, ce sont les erreurs sémantiques qui entraînent le plus de problèmes de compréhension¹⁵⁵ : 33,9 % de ces problèmes altèrent de façon plus ou moins importante la compréhension, alors que 9,9 % la compromettent complètement. Outre les problèmes sémantiques, ce sont les problèmes de collocation (collocations inexistantes ou emploi d'un collocatif ne correspondant pas au sens visé) et les erreurs formelles (essentiellement les barbarismes et les formes inexistantes) qui occasionnent le plus de problèmes de compréhension, avec respectivement 21,3 % et 19,5 % des cas observés qui altèrent complètement ou en partie le message visé. Soulignons qu'aucune des erreurs relatives à la combinatoire syntaxique, morphologique ou pragmatique n'a donné lieu à un problème de compréhension majeur.

3.3.3 LA PERCEPTION DE L'INTÉGRATION DE L'ERREUR À L'USAGE

Ce dernier critère d'évaluation de la gravité est, rappelons-le, le plus subjectif des trois et celui pour lequel nous avons obtenu le taux d'accord interjuges le moins élevé. Le score à ce critère était attribué sur la base d'une impression de la

¹⁵⁴ Il est d'ailleurs intéressant de faire le parallèle entre ces chiffres et le critère « degré de certitude relatif à la correction proposée » attribué à chaque erreur lors du codage (voir VII-2). En effet, il n'y a que 87 cas sur les 2710 erreurs relevées pour lesquels nous n'étions que moyennement certain de la pertinence de la correction proposée et seulement 10 cas pour lesquels nous n'avons suggéré aucune correction parce qu'il nous était impossible de comprendre le message visé par l'élève. Ces informations viennent étayer le constat que l'erreur lexicale constitue rarement un véritable obstacle à la compréhension, ce qui rejoint les résultats de Juhász (1970) rapportés par Johansson (1978) (voir IX-3.3.2).

¹⁵⁵ C'est d'ailleurs parce que *l'erreur lexicale* est assimilée presque exclusivement à l'erreur sémantique dans la plupart des recherches que les erreurs lexicales sont souvent considérées comme les problèmes linguistiques les plus graves.

part du codeur (et des contre-codeuses) quant au degré de l'intégration de l'erreur à l'usage.

Soixante-quatre et sept dixièmes pour cent (64,7 %) des erreurs ont été perçues comme très déviantes par rapport à l'usage normal (cote 2) dans un contexte de texte écrit, qui exige un français de niveau standard. Les erreurs liées à la combinatoire morphosyntaxique ont très souvent été jugées comme telles, principalement les erreurs concernant l'invariabilité (92 % ont reçu la cote 2) et le genre nominal (85 %), la construction d'un complément (87 %) et les prépositions régies (78 %). Les collocations inexistantes (84 %), les formes inexistantes (81 %) et les barbarismes (78 %) ont aussi été considérés « hors norme » dans de fortes proportions, tout comme les impropriétés (81 %), les erreurs de sens proche (79 %) et les problèmes sémantiques en lien avec le typage des actants (77 %).

Un quart des erreurs du corpus (25,3 %) ont reçu la cote 1 pour le critère *perception de l'intégration à l'usage*, c'est-à-dire qu'elles semblaient maladroitement et marquées par rapport à l'usage sans être particulièrement choquantes, ou encore qu'elles semblaient naturelles, mais inappropriées dans un contexte de production écrite. Les classes présentant les proportions les plus fortes d'erreurs s'étant vues attribuer la cote 1 sont au nombre de trois : l'utilisation d'une UL synonyme du sens visé mais incompatible avec le contexte (60 %), l'emploi d'une UL appartenant au registre familier (58 %) et les anglicismes (44 %).

Finalement, 10,1 % des erreurs relevées sont si fréquentes ou intégrées à l'usage qu'elles ont été classées comme parfaitement naturelles pour un locuteur natif. Parmi ces erreurs, on retrouve principalement des UL appartenant au registre de langue familier, mais qui ne sont pas nécessairement perçues comme telles même par des locuteurs experts (ex. [un] JOB [39, 43, 47, 230, etc.], VITRE [254, 256] au sens de 'verre', RELAXER [262, 294] au sens de 'se détendre', emploi de la préposition *à* avec le déterminant CHAQUE dans des expressions comme *à chaque fois/jour/mercredi). Plus du tiers des anglicismes ont aussi reçu la cote 0, encore une fois parce qu'ils

sont complètement intégrés à l'usage (ex. [le] FUTUR [39, 40, 297], PRATIQUER [184, 281, 282, 300] au sens de 'répéter (une œuvre)', FAN [282] au sens d' 'admirateur', la collocation *prendre une marche* [168]).

IMPLICATIONS DIDACTIQUES

Les travaux sur l'analyse d'erreurs s'intéressent à la gravité relative des problèmes observés dans le but d'établir des priorités dans l'enseignement des langues : les erreurs les plus graves devraient être celles méritant le plus d'attention didactique.

Selon notre échelle de gravité, il semble que ce soient les erreurs sémantiques mettant en jeu une UL ayant un sens éloigné du sens visé qui sont les plus graves et méritent par conséquent le plus l'attention de la part des enseignants. Ainsi, de telles erreurs devraient systématiquement être relevées par les enseignants sur les rédactions des élèves et une correction toujours proposée. Dans le cas où l'erreur induit une trop grande incompréhension chez le correcteur, il serait important de demander à l'élève de reformuler autrement son idée afin de trouver avec lui une correction qui corresponde vraiment au sens qu'il visait.

Par contre, nous ne croyons pas que seules les erreurs considérées comme graves (niveaux 5 et 6) selon notre calcul de gravité devraient être pénalisées. En effet, il nous semble que ce sont tous les emplois allant à l'encontre de ce que l'on trouve dans les dictionnaires (critère 1) ou très déviants par rapport à l'usage (critère 3) qui devraient être relevés systématiquement par l'enseignant. Même si une erreur de genre nominal ou de régime verbal n'entraîne généralement pas de problème de compréhension, il est du devoir de l'enseignant d'indiquer à l'élève qu'il l'a commise. Pour ce qui est des erreurs considérées comme peu graves – emploi de termes familiers, erreurs sémantiques subtiles (quasi-synonymes) et formulations maladroitement –, c'est à l'enseignant de choisir lesquelles méritent d'être relevées.

Bien sûr, pointer une erreur à l'élève ne signifie pas nécessairement qu'il faille la pénaliser. Les critères de gravité retenus dans cette recherche nous offrent certaines pistes quant aux erreurs à pénaliser. D'abord, nous sommes d'avis que toute erreur causant un problème de compréhension devrait influencer la note de l'élève. Ensuite, il nous semble évident qu'on ne peut pénaliser que ce que les ouvrages de référence nous permettent de justifier; par exemple, il serait gênant de pénaliser un élève pour une collocation boiteuse si le collocatif visé n'apparaît pas dans les dictionnaires auxquels l'élève a accès.

Bien sûr, le choix de pénaliser ou non une erreur lexicale revient bien sûr à l'enseignant : si ce dernier trouve que l'emploi du nom PUB (au sens de 'publicité') est approprié dans le contexte d'une rédaction sur la description d'une publicité, il peut choisir d'ignorer la mention *familier* associée à ce terme dans les dictionnaires. En définitive, c'est sa propre norme évaluative (III-2) qui guide l'enseignant correcteur.

3.4 RÉFLEXION MÉTHODOLOGIQUE SUR NOTRE ÉCHELLE DE GRAVITÉ DES ERREURS

Le second niveau de notre démarche d'analyse des erreurs lexicales aura permis de dégager certaines tendances relatives aux types d'erreurs lexicales et à la perception de gravité qui y est associée, et ces observations nous ont permis de faire certaines propositions didactiques. Par contre, la méthodologie associée à ce deuxième axe d'analyse, en particulier l'échelle de gravité des erreurs utilisée, présente certaines limites qu'il importe d'exposer ici.

Rappelons tout d'abord que nous avons voulu proposer une démarche d'évaluation de la gravité qui permette de positionner avec un minimum de précision les erreurs analysées sur un continuum de gravité allant d'un emploi tout à fait acceptable à une erreur unanimement considérée comme telle. Nous avons en effet constaté, dans le cadre de notre mémoire de maîtrise (Anctil, 2005), qu'une simple dichotomie *erreur* ~ *maladresse* ne permettait pas de rendre compte de la gradation dans le poids relatif des problèmes analysés et que, par conséquent, une échelle comportant plusieurs niveaux de gravité serait préférable pour traiter la question de la gravité avec plus de finesse. Comme il nous semblait plutôt arbitraire d'attribuer directement une cote de gravité (par exemple de 1 à 5) à une erreur sur la base d'une impression générale, nous avons cherché à opérationnaliser le concept de gravité en différents critères. Nous avons donc consulté plusieurs travaux portant sur l'évaluation de la gravité et conservé les trois critères qui nous semblaient les plus pertinents par rapport à l'erreur lexicale, pour ensuite les réunir dans la grille d'évaluation de la gravité utilisée dans la thèse.

D'un point de vue méthodologique, il est évident que le fait que la gravité des erreurs n'ait été appréciée que par un seul codeur – sauf pour le sous-corpus destiné au contre-codage – réduit le poids que nous pouvons accorder à nos observations, ou du moins nous oblige à nuancer grandement leur potentiel de généralisation, en dépit d'un contre-codage qui démontre des jugements somme toute assez convergents. La plupart des travaux portant sur la gravité des erreurs font appel à plusieurs locuteurs natifs (souvent jusqu'à une dizaine)

pour pouvoir produire des résultats plus solides. Idéalement, il aurait fallu qu'au moins trois personnes, outre le chercheur principal, évaluent la gravité de l'ensemble des erreurs du corpus, et il aurait été intéressant qu'au moins deux de celles-ci soient des enseignants de français.

Le fait d'avoir recours à plusieurs juges nous aurait aussi permis d'éviter certaines fluctuations dans nos jugements de gravité. Au moment de l'analyse de nos données, nous avons en effet rencontré certaines incohérences dans les scores de gravité attribués à des erreurs similaires. Ces faiblesses dans notre codage peuvent être expliquées de différentes façons. Premièrement, le processus de codage des erreurs s'étant étalé sur plusieurs semaines, il se peut que notre perception de certains problèmes ait évolué durant cette période. Certaines erreurs qui nous semblaient peut-être très graves en début de parcours nous paraissaient sans doute plus acceptables après en avoir analysé des centaines, soit parce que nous nous sommes aperçu qu'elles étaient fortement présentes dans l'ensemble du corpus, soit parce que notre impression se tempérait au fur et à mesure que nous nous familiarisions avec le niveau d'écriture d'élèves de troisième secondaire. Deuxièmement, l'appréciation de la gravité demeure subjective, et le fait que les erreurs analysées apparaissent noyées dans des dizaines d'erreurs relevant d'autres dimensions du code – particulièrement l'orthographe grammaticale – a sans doute influencé notre jugement dans les rédactions d'élèves plus faibles, de façon plus ou moins prévisible en fonction de notre degré d'empathie ou d'irritation. Enfin, il faut aussi souligner que, de la même façon que notre odorat se brouille lorsqu'on respire plusieurs parfums, nous avons remarqué la perte d'une certaine acuité linguistique lors de périodes de travail prolongées au cours desquelles nous essayions de mettre à profit nos intuitions de locuteur expert sur des dizaines d'erreurs. Dans un monde idéal, il aurait fallu poser nos jugements de gravité à tête reposée, en laissant s'écouler un certain laps de temps entre chaque erreur... ce qui nous aurait conduit à reporter le dépôt de cette thèse de plusieurs mois!

D'un point de vue un peu plus théorique, nous nous interrogeons maintenant sur la pertinence de notre opérationnalisation du concept de gravité en trois critères. Le critère d'acceptabilité linguistique a été assez facile à gérer, mais, comme le score attribué pour ce critère dépendait de ce que nous trouvions dans les ouvrages, nous nous demandons s'il a véritablement un poids dans le jugement que pose un locuteur confronté à une erreur. Si ce n'est pas le cas – ce que nous ne sommes pas en mesure de vérifier –, notre décision de l'inclure dans notre calcul de gravité fausserait les scores obtenus par rapport à d'autres qui s'appuieraient sur les critères strictement subjectifs¹⁵⁶. Le critère de compréhensibilité, quant à lui, devait être, selon les travaux consultés, le plus déterminant dans un jugement de gravité; or, nous avons vu qu'un nombre très limité d'erreurs lexicales induisant un problème de compréhension, ce qui fait que ce deuxième critère n'a influencé le score global que dans une minorité de cas. Le critère de perception de l'intégration de l'erreur à l'usage, si subjectif soit-il, semble avoir été le plus déterminant dans les jugements posés. Par contre, nous nous sommes aperçu, notamment en travaillant avec les contre-codeuses, que ce critère flou servait en fait à ajuster le score obtenu à l'aide des autres critères pour le rendre plus conforme à une cote de gravité globale déterminée de façon implicite, un peu comme si ce critère subsumait les autres pour devenir le seul vraiment déterminant.

Ces constats nous poussent à nous interroger sur la pertinence de vouloir opérationnaliser le concept de gravité. En définitive, peut-être qu'une évaluation globale de la gravité, à l'aide d'une échelle de Likert de 1 à 5, aurait été aussi valable, comme le suggère Johansson (1978) lorsqu'il affirme que les jugements de gravité posés de façon subjective et globale correspondent très fortement à ceux issus de mesures objectives.

Pour conserver une méthodologie reposant sur une opérationnalisation du concept de gravité, l'idéal aurait été de disposer de plusieurs jugements de

¹⁵⁶ Ainsi, une erreur serait considérée plus grave non pas parce que la norme prescriptive des ouvrages condamne cet usage, mais bien parce que notre propre norme évaluative (voir chapitre III) le considère hors norme.

gravité, posés par des enseignants pour un nombre important d'erreurs, puis, par le biais d'entrevues, de tenter de comprendre les aspects en jeu dans leur classement. Cela nous aurait permis alors dégager les critères véritablement impliqués dans ce type de jugements et de mieux saisir leur poids relatif¹⁵⁷ afin de proposer une grille d'évaluation de la gravité plus fidèle à leur façon de fonctionner. Bien entendu, cette entreprise constituerait à elle seule un projet de recherche doctorale et n'aurait pas été réalisable dans le cadre d'une thèse s'intéressant aux multiples facettes de la question de l'erreur lexicale. C'est donc conscient des limites de la démarche choisie ici, mais fort des observations qu'elle a permises, que nous nous tournons maintenant vers le troisième axe de notre analyse des erreurs lexicales : l'explication de leurs sources.

¹⁵⁷ Rappelons que dans notre grille, les trois critères correspondaient à un score de 0 à 2, ce qui implique qu'ils influençaient tous de façon égale le score de gravité global.

4. L'EXPLICATION DES ERREURS LEXICALES

Cette partie du chapitre s'intéressera au troisième niveau d'analyse de notre démarche : l'explication des erreurs du corpus. Toutes les fois où l'origine d'une erreur semblait pouvoir être expliquée d'une façon ou d'une autre, nous avons consigné nos hypothèses. Une multitude d'explications ont émergé au fil du codage des erreurs, et nous nous concentrerons ici sur celles qui sont les plus récurrentes, donc susceptibles d'expliquer un certain nombre d'erreurs lexicales. Comme plusieurs hypothèses étaient parfois proposées pour une même erreur, nous tâcherons aussi de voir lesquelles apparaissent ensemble de façon privilégiée et si elles constituent des hypothèses concurrentes, ou plutôt des sources d'erreurs qui agiraient en synergie. Mais d'abord, faisons un détour du côté des données produites par les contre-codeuses pour voir dans quelle mesure les hypothèses que nous proposons s'apparentent à celles qu'elles ont avancées.

4.1 LES HYPOTHÈSES DES CONTRE-CODEUSES À PROPOS DES SOURCES DES ERREURS

Lors du contre-codage des critères de gravité des erreurs, dont les résultats ont été présentés plus haut, nous avons demandé aux contre-codeuses de proposer des hypothèses d'explication des erreurs du corpus lorsque de telles hypothèses leur venaient spontanément à l'esprit. Notre but n'était pas ici de produire un contre-codage systématique pour ce niveau d'analyse, mais plutôt de nous fournir des données pour que nous puissions observer si leurs hypothèses se rapprochent des nôtres. Rappelons qu'aucune liste de sources d'erreurs n'était proposée dans le guide de contre-codage, dans le but de ne pas limiter les hypothèses possibles : y étaient seulement mentionnées quelques pistes, comme la notion d'analogie et l'influence de la langue orale.

Mentionnons d'entrée de jeu que les contre-codeuses ont posé beaucoup moins d'hypothèses d'explication que nous. Pour les 130 erreurs à analyser, nous avons proposé des scénarios d'explication dans 95 cas (73 %), alors qu'elles en ont avancé respectivement 55 (42 %) et 36 (26 %). Notons aussi qu'elles ont

toujours émis une seule hypothèse pour les cas traités, alors que nous en avons posé deux pour certaines des erreurs de ce corpus. Diverses raisons peuvent expliquer ces différences. Tout d'abord, les contre-codeuses n'avaient accès qu'à un court extrait du texte pour chaque erreur, ce qui rendait sans doute plus difficile la tâche d'explication du problème observé. De plus, le fait que nous ayons vu des centaines d'erreurs a sans doute eu une incidence sur la facilité avec laquelle nous avons pu poser certaines hypothèses; en ne travaillant que sur 130, les codeuses ont sans doute trouvé la tâche plus difficile. Il faut aussi souligner que le manque de balises pour ce volet du contre-codage (aucune d'obligation d'émettre une hypothèse et absence de liste de sources possibles) peut sans doute expliquer en grande partie que les codeuses aient été moins enclines à se concentrer sur cette dimension du codage. Par ailleurs, une partie non négligeable des commentaires produits par les contre-codeuses ne visaient pas réellement l'explication des erreurs en jeu, mais constituaient plutôt des remarques sur leur perception de ces erreurs (ex. *J'aurais accepté, même si la formulation est maladroite.* ou *Selon moi, le sens visé était plutôt X que Y.*), ce qui fait qu'ils ont dû être écartés; sans doute une plus longue séance de travail commun avant le contre-codage aurait-elle permis aux codeuses de mieux se familiariser avec cette dimension de leur tâche.

Des 95 erreurs pour lesquelles nous avons nous-même proposé une explication, 55 ont fait l'objet d'une hypothèse de la part d'au moins une des deux contre-codeuses. Seize de ces 55 cas (29 %) ont donné lieu à des scénarios d'explication similaires de la part des trois codeuses : la plupart du temps, l'hypothèse avancée était une influence de la langue orale. Dans 20 cas (36 %), seule une des codeuses avait proposé une explication de l'erreur, mais son explication était similaire à la nôtre. C'est seulement pour cinq cas (9 %) que les scénarios d'explication proposés – soit par une seule codeuse soit par les deux – différaient complètement de notre propre hypothèse.

Ces observations suggèrent donc que les scénarios d'explication ont tendance à converger lorsque deux personnes ou plus tentent d'établir la source présumée

d'une erreur; mais bien sûr, le nombre restreint de cas en jeu ici ne nous permet pas de mesurer de façon très précise cette convergence. Si nous avions exigé de nos codeuses qu'elles posent une hypothèse sur l'origine de chacune des erreurs pour lesquelles nous en avons nous-mêmes avancé une, nous aurions déjà eu un plus grand nombre de cas pour tirer des observations plus concluantes. Par contre, cela aurait été à l'encontre de la méthodologie que nous avons nous-même appliquée tout au long de notre codage, soit de ne proposer des hypothèses que lorsqu'elles se présentent spontanément.

Évidemment, il aurait été encore mieux de demander à des enseignants d'expliquer les erreurs, eux qui sont confrontés quotidiennement aux erreurs lexicales de leurs élèves. Pouvoir poser des questions directement aux élèves ayant produit les textes nous aurait sans doute permis d'écarter certaines de nos hypothèses. Mais une telle méthodologie aurait grandement alourdi l'étape de l'analyse des données et sans doute compliqué énormément le recrutement de sujets.

Malgré les conclusions limitées qu'elles permettent, les hypothèses d'explication d'erreurs issues du contre-codage nous rassurent sur deux points. Tout d'abord, nous constatons que nos propres scénarios d'explication ne diffèrent de ceux proposés par d'autres que dans un nombre limité de cas. Deuxièmement, les contre-codeuses n'ont pas avancé d'hypothèses auxquelles nous n'avions pas pensé et qui auraient pu modifier de façon importante la liste d'explication des erreurs lexicales que nous avons construite au fur et à mesure de notre analyse. Comme nous, elles font beaucoup référence à l'influence de l'oral, à l'intégration de l'erreur à l'usage et, dans une moindre mesure, à l'analogie. Nous utiliserons donc nos propres scénarios d'explication pour explorer l'ensemble du corpus d'erreurs, tout en gardant en tête qu'ils ne constituent que des hypothèses et qu'ils sont sans doute incomplets en regard des hypothèses qui auraient émergé si nous avions pu réfléchir à chaque erreur à plusieurs, notamment avec des enseignants.

Passons maintenant aux principales sources d'erreurs présumées.

4.2 LES PRINCIPALES SOURCES D'ERREURS LEXICALES

Des 1144 erreurs du corpus¹⁵⁸, 793 ont fait l'objet d'au moins une hypothèse quant à leur origine; 695 n'ont donné lieu qu'à une seule, alors que 93 correspondent à deux et que cinq cas nous semblaient pouvoir être expliqués de trois façons différentes. Ceci représente un total de 896 scénarios d'explication pour les 793 erreurs « diagnostiquées ». Il est tout de même important de rappeler que nous n'avons proposé des explications seulement dans le cas où une hypothèse nous apparaissait spontanément en cours de codage. Il ne faut donc pas lire les chiffres présentés ici comme des absolus et comprendre que seulement 793 erreurs sur les 1144 du corpus pouvaient être expliquées d'une façon ou d'une autre. À postériori, il aurait sans doute été possible d'avancer des scénarios d'explication pour certaines erreurs pour lesquelles nous ne l'avons pas fait lors du codage et un autre codeur aurait potentiellement proposé des explications pour d'autres erreurs, ou d'autres explications pour les mêmes erreurs. Ce qu'il faut retenir des chiffres présentés, c'est que nous disposons d'hypothèses d'explication pour 793 des erreurs du corpus, ce qui constitue un volume de données assez important pour pouvoir en tirer des observations ayant une certaine validité.

Comme nous l'avons mentionné dans notre chapitre de méthodologie, nous ne disposions pas d'une liste fermée de sources possibles d'erreurs pour le volet « explication des erreurs ». Ainsi, à partir des quelques pistes que nous proposait la littérature, et au fur et à mesure de notre codage, nous avons créé une liste de sources possible d'erreurs qui s'est stabilisée en cours d'analyse. Bien entendu, certaines sources étaient beaucoup plus fréquentes que d'autres, et c'est sur celles-là que nous nous concentrerons dans les sections qui suivent.

La figure qui suit indique dans quelle proportion chaque hypothèse d'explication a été avancée, de la plus fréquente à la moins fréquente. Nous rappelons que

¹⁵⁸ Comme notre travail ne s'intéresse aux erreurs d'orthographe lexicale que de façon périphérique, nous n'avons pas proposé de scénarios d'explication pour ce type de problème, ni pour les erreurs flexionnelles d'ailleurs, d'où le fait que seulement 1144 des 2720 erreurs du corpus auraient potentiellement faire l'objet d'une hypothèse quant à leur(s) source(s).

certaines erreurs ont fait l'objet de plus d'une hypothèse; c'est ici les pourcentages des 896 scénarios d'explication qui apparaissent dans la figure 18.

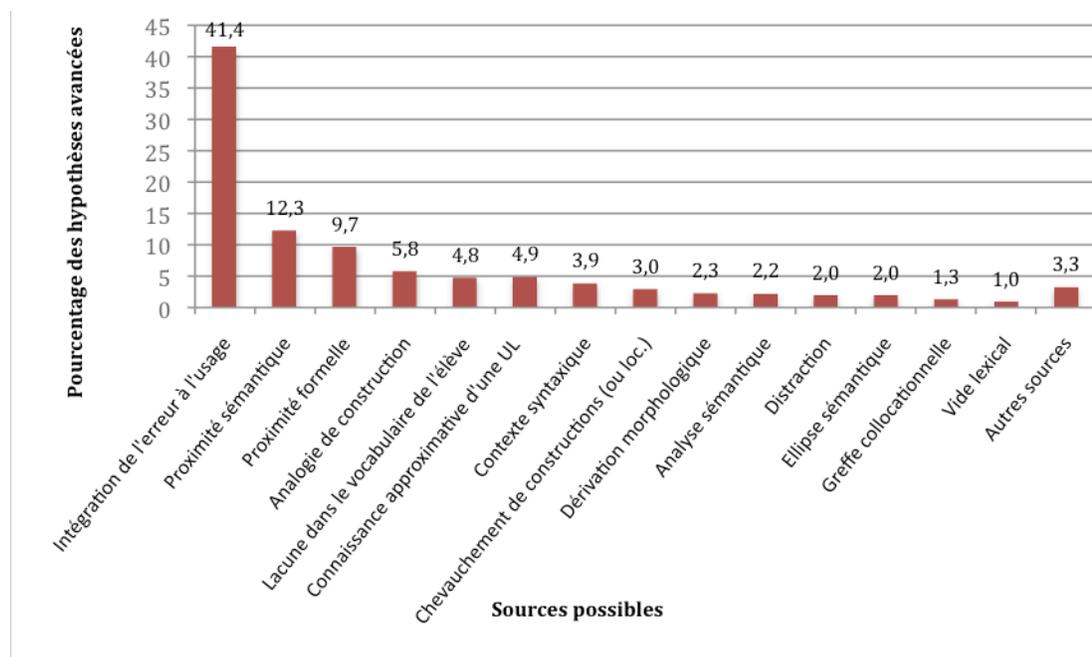


Figure 18. Représentation des différents types hypothèses sur les 896 posées

4.2.1 L'INTÉGRATION DE L'ERREUR À L'USAGE

L'intégration de l'erreur à la langue, principalement à la langue familière utilisée à l'oral, est de loin la source d'erreurs la plus récurrente dans les explications que nous avons proposées; elle représente 41,4 % de nos hypothèses (371 cas sur 896) et concerne 46,8 % des erreurs. Nous avons d'ailleurs remarqué que c'est aussi l'explication la plus fréquente fournie par les contre-codeuses pour leur partie du corpus.

Ainsi, si un élève commet une erreur lexicale, c'est souvent parce qu'il a déjà lu, ou à plus forte raison entendu, cette erreur plusieurs fois. Bien sûr, cette explication correspond dans de très nombreux cas à des erreurs de registres de langue (92 % des erreurs de ce type correspondaient à cette hypothèse) : l'élève entend [un] MONSIEUR [158], AVOIR DE LA MISÈRE [173] et SUPER [3, 6] et utilise tout simplement ces UL dans ses rédactions, sans même savoir qu'elles sont considérées comme familières. Cette même hypothèse permet d'expliquer

l'emploi d'anglicismes¹⁵⁹ (ex. LOCATION [15] au sens d' 'endroit', LOOK [131] au sens d' 'allure', SUPPOSÉ [19] au sens de 'censé') dans des proportions semblables (91 %), ainsi que les trois-quarts (77 %) des erreurs ayant trait au genre nominal (ex. [*une] ÉTÉ [44, 230, etc.], [*une] AUTOBUS [256], [*une] PÉTALE [169]). Près du tiers des barbarismes observés (32 %) sont aussi très répandus (ex. *à *chaque fois* [57, 64, 146, 194, 245, 269] plutôt que CHAQUE FOIS, *avoir de *besoin [de]* [77] plutôt qu'AVOIR BESOIN [de], ARRÊTER [292] au sens de 's'arrêter'), tout comme la plupart des pléonasmes rencontrés (ex. *comme par exemple [70, 128], *entrer à *l'intérieur* [107]). Dans une moindre mesure, l'influence de la langue orale est sans doute aussi à l'origine de certaines erreurs liées à la combinatoire syntaxique, par exemple celles présentes dans les phrases suivantes.

(155) *La semaine suivante, nous **quittions** déjà [quittions déjà ce camping] pour le Rocher Percé. [271]*

(156) *Alors qu'ils ne **s'en attendaient** [s'y attendaient] pas, un gros éclair déchira le ciel. [202]*

(157) *... si ils leurs **seraient** [étaient] arrivé quelque chose... [275]*

IMPLICATIONS DIDACTIQUES

Une prise de conscience de la part des élèves quant à l'existence de la variété des registres de langue – notamment le langage familier typique de la langue orale – pourrait sans doute diminuer le nombre d'erreurs attribuables à une influence de la langue orale. Montrer aux élèves comment identifier les termes familiers à l'aide des dictionnaires et surtout leur apprendre à douter de l'emploi standard de plusieurs UL utilisées fréquemment dans le langage courant les aiderait sans doute à éviter certaines erreurs de registres. Lors de la correction de productions écrites, il pourrait être pertinent de réserver une annotation particulière pour les termes familiers (ex. *fam.*) et de proposer à l'élève un terme standard équivalent à l'UL familière. On peut aussi demander à l'élève de trouver lui-même cet équivalent à l'aide d'un dictionnaire dans le cas où la rédaction fait l'objet d'un travail de correction. Il est aussi possible de récolter les termes familiers utilisés par les élèves de façon récurrente pour bâtir une activité sur les registres de langue.

Quant aux erreurs bien intégrées à l'usage, il faut les présenter aux élèves comme des emplois à éviter, par exemple en construisant avec eux une liste d'anglicismes fréquents ou de noms masculins et féminins, ou en leur faisant consigner dans un aide-mémoire lexical la structure contrôlée par certains verbes.

¹⁵⁹ Ceci vient en quelque sorte étayer notre discussion à propos des anglicismes (voir V-3.2.2) : si un élève utilise un anglicisme, ce n'est souvent pas sous l'influence de la langue anglaise, mais bien parce que l'UL problématique est tout à fait passée dans l'usage des locuteurs de sa langue et qu'il y est exposé régulièrement.

4.2.2 LA PROXIMITÉ SÉMANTIQUE ENTRE L'UL ERRONÉE ET L'UL VISÉE

La présence de composantes sémantiques communes ou d'un lien conceptuel entre l'UL erronée et l'UL visée est la deuxième hypothèse d'explication la plus fréquente et concerne 13,9 % des erreurs. Évidemment, cette source d'erreurs est principalement évoquée dans le cas de problèmes sémantiques, comme ceux des phrases suivantes.

- (158) *Il voulut y rentrer, mais il n'y avait pas d'issus [accès, ouverture] pour rentrer à l'intérieur. [107]*
- (159) *J'espert que mon texte explicatif sur cette publicité vous seras **éficace** [utile] pour votre reportage. [85]*
- (160) *... les couleurs de la **propagande** [publicité] varient entre mauve, rose, rouge et orange. [165]*
- (161) *Il courrait pour se rendre au sommet où le colérique Lithium **se situait** [se trouvait]. [61]*

Dans la phrase (158), l'UL ISSUE est un quasi-synonyme d'OUVERTURE, sauf qu'elle dénote une 'ouverture offrant la possibilité de sortir', ce qui n'est pas compatible avec le contexte de la phrase, dans laquelle le personnage principal cherche plutôt à entrer. Néanmoins, l'élève a sans doute confondu les deux UL sur la base d'une forte proximité sémantique, ou du moins – parce qu'il est impossible de savoir avec assurance que l'élève a confondu les deux UL –, a utilisé ISSUE parce que son sens se rapprochait du sens visé, celui d'OUVERTURE. Le lien de sens entre EFFICACE et UTILE (phrase (159)) est sans doute moins étroit, mais il demeure aisé de le percevoir : si quelque chose est efficace, cela produit des résultats utiles. On note d'ailleurs que certains ouvrages fournissent UTILE parmi la liste des synonymes d'EFFICACE. Le lien est encore plus ténu entre PUBLICITÉ et PROPAGANDE, mais on sait que la publicité est un outil de propagande important; ce lien conceptuel a peut-être poussé l'auteur de la phrase (160) à commettre cette erreur. À plus forte raison, peut-être a-t-il rencontré le terme PROPAGANDE dans ses lectures préparatoires et l'a-t-il assimilé à PUBLICITÉ ou, plus simplement, peut-être a-t-il simplement choisi un peu au hasard un « synonyme » de PUBLICITÉ dans un dictionnaire des synonymes en n'ayant aucune idée du nombre limité de

contextes où ces deux UL sont interchangeable¹⁶⁰. Dans la phrase (161), l'utilisation du verbe SE TROUVER au lieu de SE SITUER peut sans aucun doute être expliquée par le fait que ces deux UL sont des synonymes exacts (ex. *La Sagrada Família se situe à Barcelone.* = *La Sagrada Família se trouve à Barcelone.*), si ce n'est que SE SITUER est plus contraint quant à ses actants sémantiques : il ne peut avoir de sujet vivant, comme le magicien Lithium de la phrase (161).

Au-delà des problèmes purement sémantiques, il arrive que des composantes sémantiques communes permettent d'expliquer certaines autres erreurs. Par exemple, l'élève écrivant **contribuer un don* [85] est sans doute conforté dans son choix de collocatif par le fait que CONTRIBUER est très lié sémantiquement à l'idée de don, tout comme l'élève produisant la collocation **sa puissante force* [199] est probablement convaincu que PUISSANT peut intensifier FORCE, parce qu'on peut dire d'une personne forte qu'elle est puissante. On peut aussi penser que la personne ayant produit le barbarisme **deux torches enflanbées* [116] s'est embourbée entre l'adjectif ENFLAMMÉ et le verbe FLAMBER, sémantiquement très proches; de la même façon, l'élève ayant écrit **tournade* [170], faute de maîtriser la forme correcte de ce nom, sait au moins qu'une tornade, ça tourne!

Voilà donc les principaux types d'erreurs lexicales tirant selon nous leur origine d'une proximité sémantique entre deux UL. On peut remarquer, à la lecture des deux dernières erreurs que nous venons de présenter, qu'une proximité formelle existe aussi, en plus de la proximité sémantique, entre les deux UL en jeu dans l'explication de ces barbarismes. Il est effet frappant de constater que nous avons proposé une double explication *parenté formelle ~ parenté sémantique* pour un nombre d'erreurs assez important (42). Nous présenterons plus bas quelques-uns de ces cas, mais intéressons-nous d'abord aux erreurs lexicales ayant pour source une parenté formelle entre l'UL erronée et celle visée.

¹⁶⁰ En fait, l'acception de PUBLICITÉ visée ici par l'élève ('annonce publicitaire') n'est jamais, selon nous, synonyme de PROPAGANDE. C'est une acception plus large de PUBLICITÉ ('campagne publicitaire') qui peut parfois être remplacée par PROPAGANDE.

IMPLICATIONS DIDACTIQUES

Dans le cas où l'erreur est attribuable à une proximité sémantique entre l'UL erronée et l'UL visée, il est important de fournir à l'élève une proposition de correction plutôt que de seulement souligner l'erreur. Dans le cas où un retour sur la production écrite est possible, il serait intéressant de faire ressortir les nuances de sens entre les deux termes pour expliquer en quoi l'UL erronée est problématique dans le contexte; pour ce faire, il peut être intéressant d'utiliser le dictionnaire avec l'élève, tout en étant conscient, toutefois, qu'il est probable que l'ouvrage ne présente pas les composantes de sens recherchées de façon explicite. Dans ce cas, c'est par des exemples d'utilisation des deux UL que l'on pourra illustrer en quoi elles diffèrent.

De façon un peu plus formelle, réaliser en classe des activités portant sur des familles de synonymes peut s'avérer très pertinent pour mettre en lumière les nuances de sens existant entre des UL sémantiquement proches. Travailler sur un groupe de mots (par exemple un ensemble de cinq synonymes fournis pour un adjectif dans un dictionnaire de synonymes) peut permettre à l'élève, par des oppositions et des rapprochements, d'élargir et de structurer son vocabulaire.

Composer des phrases contenant les termes travaillés et tenter de les permuter peut faire ressortir des contraintes sémantiques liées à certaines UL et mettre en évidence le spectre plus large de contextes dans lesquels d'autres sont appropriées.

4.2.3 LA PARENTÉ FORMELLE ENTRE L'UL ERRONÉE ET L'UL VISÉE

La parenté formelle entre l'UL erronée et l'UL visée est le troisième scénario d'explication le plus fréquent et nous l'avons considérée comme une source possible pour 11 % des erreurs diagnostiquées.

C'est encore une fois pour la classe des erreurs sémantiques que la parenté formelle est le plus souvent évoquée comme scénario d'explication. En voici quatre exemples.

(162) ... le conducteur commença à toucher la cuisse de Marie. Elle était dégoûtée et elle fit aussitôt un **signal** [signe] pour débarquer. [256]

(163) Sa mère lui passa, donc, le téléphone pour qu'elle puisse écouter un message qu'elle avait recue. Il **énonçait** [annonçait, l'informait] que suite à la demande d'emploi qu'elle avait envoyé, à la piscine, pour être sauveteure, elle était engagée. [280]

(164) Elle est une jeune fille **bordée** [débordante] d'imagination. [178]

(165) Ce royaume **habitait** [abritait] un jeune nain du nom de Frali. [97]

Le fait que l'élève utilise SIGNAL plutôt que SIGNE dans la phrase (162) n'est probablement pas étranger à la ressemblance entre ces deux formes, tout

comme nous pouvons croire que la parenté formelle entre ÉNONCER et ANNONCER peut expliquer la confusion de la phrase (163). L'erreur en (164) est beaucoup plus surprenante, mais il est très probable que l'élève ait produit l'adjectif *bordé sous l'influence du verbe DÉBORDER, collocatif fréquent d'IMAGINATION (ex. *Elle déborde d'imagination.*). Dans la phrase (165), il semble que l'élève ait inversé les actants du verbe HABITER. Plutôt que d'écrire *Un jeune nain habitait ce royaume.*, il a écrit **Ce royaume habitait un jeune nain.*; cette inversion ne pourrait-elle pas être expliquée par une confusion avec le verbe ABRITER, qui serait tout à fait compatible avec ce contexte (*Ce royaume abritait un jeune nain.*)? On remarque encore une fois que dans la plupart des cas, la parenté formelle s'accompagne d'une proximité de sens : un signal est un type particulier de signe, ÉNONCER et ANNONCER partagent des composantes de sens évidentes ('transmission d'informations') et ABRITER et HABITER pourraient dans certains cas être considérés comme des conversifs.

Un certain nombre de barbarismes (environ 20 %) peuvent aussi être expliqués par une parenté formelle avec une autre UL. Nous retrouvons bien sûr quelques cas de confusion entre des formes paronymes – c'est-à-dire très proches d'un point de vue phonétique sans être des homophones –, comme dans les phrases (166) et (167).

(166) ... chercher du bois (...) pour faire un feu **enfin** [afin] de les rechauffer. [48]

(167) Il revint et tout les habitants **firent** [furent] une fête pour féliciter de l'exploit de leur jeune prince. [95]

(168) Tous les villageois [...] ont tous remerciés Radolfe pour son courage et sa **tenacité** [ténacité]. [98]

(169) S'il ne buvait pas de l'eau à temps, il allait tomber **sur la connaissance** [sans connaissance]. [18]

En (168), c'est sans doute le lien avec l'adjectif TENACE, correspondant au nom TÉNACITÉ, qui est à l'origine de l'erreur; cette erreur est en fait logique et l'élève n'a en fait qu'uniformisé une famille morphologique un peu « dysfonctionnelle »!

À notre avis, l'erreur de la phrase (169) découle d'une mauvaise compréhension de la forme orale de l'expression. Il est vrai qu'à l'oral, l'expression *tomber sans connaissance* pourrait s'entendre comme **tomber sur la connaissance* prononcée en français québécois familier (*s'a connaissance* – forme contractée de *sur la*

connaissance). Il se peut donc que l'élève ait toujours décodé **tomber sur la connaissance* lorsqu'il entend cette expression, ce qu'il reproduirait à l'écrit. D'ailleurs, il ne serait pas le premier à s'y faire prendre : on a qu'à relire les dialogues du cycle des *Chroniques du Plateau Mont-Royal* de Michel Tremblay pour s'en convaincre!

Comme nous l'avons mentionné plus haut, beaucoup d'erreurs ayant une parenté formelle entre l'UL utilisée et l'UL visée comme source présumée présentent aussi un lien sémantique. Ce sont 42 erreurs qui ont reçu la double hypothèse *parenté formelle ~ parenté sémantique*, la plupart étant classées parmi les problèmes sémantiques, comme les suivantes.

(170) ... une géant tellement gros, qu'il pouvait **confronter** [affronter] le vent. [20]

(171) Vol'jin **renonçait** [refusait] de garder les bras croisés. [61]

(172) Un homme à l'allure délabré l'interpella pour lui demander quelques sous. **Sentimental** [Sensible] comme il l'est, Gregory s'empressa de lui donner la monnaie qui lui restait au fond de ses poches. [219]

Dans la phrase (170), le verbe **CONFRONTER** est utilisé pour exprimer le sens 'affronter'. Or, il existe un lien sémantique fort entre **AFFRONTER** – 'aller au-devant de quelque chose, de quelqu'un avec courage' – et une des acceptions pronominales de **CONFRONTER** – 'faire face à quelqu'un, à quelque chose' (ex. *Les policiers se sont confrontés aux contestataires*). En outre, ces verbes sont très proches d'un point de vue formel : même nombre de syllabes (3), dont les deux finales identiques; Laufer (1992) parlerait de *synforms* de type « préfixes », puisque même si les premières syllabes des mots-formes en jeu ne sont pas ici des préfixes (ils ne véhiculent aucun sens précis), ce sont elles qui permettent la distinction entre les UL auxquelles ces mots-formes correspondent.

La confusion entre **RENONCER** et **REFUSER** de la phrase (171) peut aussi, selon nous, être expliquée par une double parenté formelle et sémantique. Phonétiquement, les deux mots-formes en jeu présentent le même nombre de syllabes (3), dont la première est identique et la dernière présente seulement une différence de voisement entre les fricatives /z/ et /s/. Sémantiquement, les deux termes renvoient à l'idée de 'ne pas faire quelque chose (que l'on avait prévu faire ou que

quelqu'un aurait voulu nous voir faire)'. La confusion nous semble d'autant plus probable que la préposition DE utilisée avec RENONCER correspond en fait à la préposition régie par le verbe REFUSER.

La ressemblance formelle entre SENTIMENTAL et SENSIBLE est peut-être moins frappante, mais le fait que les deux mots-formes commencent par la même syllabe jumelé au fait que ces termes sont des quasi-synonymes (SENSIBLE = 'qui se laisse facilement émouvoir', SENTIMENTAL = 'qui se plait à se laisser aller à des sentiments doux') permettent selon nous d'expliquer l'origine de l'erreur de la phrase (172).

Outre un nombre important d'erreurs sémantiques, certaines erreurs appartenant à d'autres classes peuvent aussi être expliquées par une double parenté formelle et sémantique, dont les deux suivantes.

(173) *Athina dut passer l'éternité à ensorceler les marins et les entrainer vers la mort, pour **assoupir** [assouvir] sa vengeance. [181]*

(174) *Cette un chef-d'oeuvre pour **insister** [inciter] les gens à acheter. [135]*

L'erreur de collocation présente dans la phrase (173) découle selon nous principalement d'une confusion entre les paronymes ASSOUIR et ASSOUVIR, mais une certaine proximité formelle entre ces deux verbes la légitime en quelque sorte. En effet, ASSOUVIR signifie 'satisfaire, calmer un besoin', alors qu'ASSOUIR signifie 'affaiblir, calmer un sentiment, une sensation'. La principale différence sémantique entre ces verbes réside selon nous dans le caractère plus définitif d'ASSOUVIR : quelqu'un qui ne fait qu'*assoupir sa vengeance* se retrouvera tôt ou tard avec ce même besoin à combler, ce qui ne peut être le cas d'Athina dans le texte de l'élève, puisqu'elle est condamnée à ensorceler les marins pour l'éternité. L'erreur de la phrase (173), que nous avons classée parmi les barbarismes (ajout d'un phonème), est à notre avis attribuable à une confusion entre les paronymes INCITER et INSISTER. Pourtant, il existe encore une fois un lien sémantique assez fort entre ces deux verbes : lorsqu'on incite quelqu'un à faire quelque chose, c'est souvent en insistant auprès de cette personne.

Il semble donc qu'une confusion entre des UL sur la base d'une parenté formelle soit souvent accompagnée de certains liens de sens entre les deux termes.

IMPLICATIONS DIDACTIQUES

À notre avis, toute erreur attribuable à une parenté formelle entre deux UL devrait donner lieu à une intervention tenant compte à la fois de la forme et du sens des deux UL en jeu. Il ne s'agit pas simplement de fournir à l'élève la forme correcte ou de lui demander de la trouver par lui-même; il est important de mettre en lumière les différences sémantiques existant entre ces dernières. Ceci est particulièrement vrai pour les erreurs découlant d'une double parenté formelle et sémantique : comme pour les problèmes sémantiques, il est essentiel de faire ressortir les nuances de sens entre deux termes formellement semblables. Encore une fois, il peut être intéressant de produire des phrases contenant ces UL pour contraster les contextes dans lesquels elles peuvent être utilisées.

Au-delà de l'attention soutenue accordée aux homophones grammaticaux, il pourrait être pertinent de présenter en classe des homophones lexicaux et d'explorer la notion de paronymie. L'enseignant peut par exemple proposer une liste de paronymes donnant fréquemment lieu à des erreurs; à cet effet, il serait d'ailleurs pertinent que l'enseignant construise lui-même une telle liste en consignnant les UL formellement proches confondues par les élèves. En effet, nous avons constaté que si les dictionnaires présentent certaines mises en garde quant à la confusion possible entre des UL, ces remarques concernent rarement celles en jeu dans les erreurs observées dans notre corpus; par exemple, *Antidote RX* liste ASSOUIR et ASSOULIR comme des paronymes, mais ne fait aucune mention d'ASSOUVIR, UL très semblable à ASSOUIR d'un point de vue formel (un seul phonème les distingue) et beaucoup plus proche sémantiquement qu'ASSOULIR. C'est donc à l'enseignant d'élargir la notion de paronymie à des paires d'UL problématiques pour ses élèves.

4.2.4 L'ANALOGIE DE CONSTRUCTION ENTRE DEUX UL

La quatrième source d'erreurs la plus souvent avancée pour expliquer les erreurs de notre corpus est la transposition de la structure syntaxique contrôlée par une UL à une autre : elle représente 5,8 % de nos hypothèses et concerne 6,6 % des erreurs, principalement celles liées à la combinatoire syntaxique.

(175) *La police avait mis une loi qui interdirait les [aux] paysans de s'aventurer la nuit. [117]*

(176) *Le jour vint succéder [succéder à] la nuit. [200]*

(177) *Le chef s'empressa à [d'] aller chercher cette formule dans une gigantesque bibliothèque. [18]*

(178) *L'horloger se dirigea à une vitesse éclair à [vers] la caverne. [88]*

Dans toutes les phrases ci-dessus, la structure erronée correspond en fait au régime d'une UL synonyme de l'UL erronée : **interdire les paysans de s'aventurer pour empêcher les paysans de s'aventurer* (175), **succéder la nuit pour suivre la nuit* (176), **s'empresse de pour se dépêcher de* (177) et **se diriger à pour aller, courir à* (178). Les analogies de construction observées semblent donc sémantiquement motivées, puisque le terme dont provient la structure erronée est toujours un (quasi) synonyme de l'UL problématique.

Dans certains cas, plus rares, l'erreur de construction apparaît dans une collocation, comme dans les deux phrases suivantes.

(179) ... de la poudre de rêve qui endormait toute personne **dans** [d'] un profond sommeil. [100]

(180) Dès que le soleil se leva **en** [à la] fin mai, il fit un immense jardin. [166]

L'emploi de la préposition **DANS** de la phrase (179) peut sans doute être expliqué par des expressions comme *tomber dans un profond sommeil* ou *somber dans un profond sommeil*; seulement, lorsque [profond] **SOMMEIL** a **S'ENDORMIR** comme collocatif, il faut plutôt dire *s'endormir d'un profond sommeil*. Et si l'élève en (180) a écrit **en fin mai* plutôt qu'*à la fin mai*, c'est sans doute qu'on dit généralement *en mai* lorsqu'on veut exprimer qu'un événement se déroule 'durant le mois de mai'. Dans ces deux phrases, l'analogie de construction n'est pas nécessairement entre deux UL différentes, mais entre des constructions différentes d'une même UL apparaissant dans des contextes syntaxiques différents.

Une autre hypothèse d'explication d'erreur peut être associée à l'analogie de construction : la greffe collocationnelle. Ce terme a été proposé par Polguère (2007) pour désigner les « collocations (expressions semi-idiomatiques) déviantes, où un des éléments semble avoir été "emprunté" à une autre collocation valide de la langue. » (p.1) Voici quatre erreurs de collocations pour lesquelles cette hypothèse nous semble valable.

(181) Un soir, où la lune **plombait** [brillait], Berthèse et ses amis se promenaient dans la forêt, loin des sentiers battus. [114]

(182) Pour l'aider à **accomplir** [entreprendre, triompher dans] cette dangereuse aventure, elle amena avec elle sa potion bleue, sa plume sortilège et son remède anti-chaleur. [110]

- (183) *En second lieu, la pratique [avoir un] d'emplois d'été permet d'acquérir des compétences, de l'expérience et de l'argent. [218]*
- (184) *J'ai regardé sur internet pour voir quel emploi je pourrait faire [exercer, occuper; obtenir], il y en a deux qui m'ont marqué. [272]*

Dans le cas de l'erreur de la phrase (181), il nous apparaît évident que l'élève a choisi le collocatif PLOMBER sur la base de la collocation *le soleil plombe*, ce que Polguère appelle *collocation source (de la greffe collocationnelle)*. La collocation visée ou *collocation cible* devrait plutôt être *la lune brille, luit*. Les deux cas correspondent à la FL Func₀.

Le cas de (182) est similaire : l'élève a selon nous greffé un collocatif de MISSION (*accomplir une mission*) à l'UL AVENTURE, dont elle est sémantiquement proche. Par contre, dans ce cas-ci, il n'existe pas selon nous de valeur correspondant à la FL Real₁ pour AVENTURE, comme il en existe pour MISSION (Real₁(mission) = réussir, remplir, accomplir, mener à bien [une ~]); ceci explique peut-être pourquoi l'élève aurait emprunté un collocatif à une autre UL, d'autant plus que l'aventure dont il est question dans le texte s'apparente vraiment à une mission : délivrer un village d'un mauvais sort.

L'erreur de la phrase (183) est plus difficile à expliquer, mais s'apparente aussi à une greffe collocationnelle. L'utilisation du nom PRATIQUE avec EMPLOI découle selon nous de la collocation erronée **pratiquer un emploi*, qui elle-même serait une greffe collocationnelle ayant pour source *pratiquer un métier*. Cette erreur est d'ailleurs très similaire à celle de (184), où le collocatif FAIRE, souvent associé à METIER, apparaît ici avec EMPLOI. Soulignons qu'encore une fois, les bases des collocations cible et source sont sémantiquement très proches (il s'agit ici de synonymes) et que le sens à exprimer est le même dans les deux collocations (FL Oper₁).

IMPLICATIONS DIDACTIQUES

Les interventions didactiques liées aux erreurs attribuables à la transposition du régime d'une UL à une autre devraient surtout porter sur les verbes et, encore une fois, reposer sur la comparaison d'UL sémantiquement proches.

Plutôt que d'indiquer simplement à l'élève la bonne structure syntaxique à employer avec l'UL en jeu dans son erreur, par exemple en biffant la préposition À pour la remplacer par DE, il serait pertinent de lui indiquer aussi quel synonyme il aurait pu utiliser en préservant le même contexte syntaxique.¹⁶¹

Pour les collocations erronées, il est aussi intéressant, lorsque c'est le cas, de pointer à l'élève que le collocatif erroné qu'il a employé aurait été approprié auprès d'une UL synonyme, par exemple *pratiquer un métier* plutôt que **pratiquer un emploi*. De la même façon, on peut aussi indiquer certains collocatifs communs à des termes synonymes (ex. *exercer un métier, un emploi*).

4.2.5 UNE LACUNE DANS LE VOCABULAIRE DE L'ÉLÈVE

Dans 5,4 % des cas, nous avons considéré que l'erreur observée pouvait être imputable à une lacune dans le vocabulaire de l'élève, un « trou lexical ». Cette explication correspond principalement à des erreurs de collocation, comme les suivantes.

(185) ... un travail **a** [*procure; garantit*] un salaire. [218]

(186) Mais un bruit **ce fut** [*retentit*]. Dans le loin de la forêt. [124]

(187) Ces créatures sanguinaires s'attaquaient aux habitants de petits villages pour **réduire** [*assouvir, satisfaire*] leur soif de sang. [161]

(188) J'avais aussi "**pris**" [*adopté*] une partie de la mentalité européenne. [257]

Dans plusieurs cas correspondant à des FL sémantiquement vides, comme les phrases (185) et (186), qui correspondent à la fonction Func¹⁶², l'élève utilise des verbes passepartouts comme AVOIR et FAIRE¹⁶³ comme collocatifs. L'utilisation de tels verbes, dans des cas où ces collocatifs ne sont pas appropriés¹⁶⁴ nous laisse

¹⁶¹ Dans la phrase (175), par exemple, il est important que l'élève apprenne que le deuxième actant sémantique du verbe INTERDIRE est introduit par la préposition À (*X interdit à Y de Z*), mais il serait encore plus efficace qu'il se souvienne que cette structure est aussi celle du verbe synonyme DÉFENDRE (*X défend à Y de Z*) et que le verbe EMPÊCHER, pourtant synonyme lui aussi, correspond quant à lui au régime utilisé par l'élève dans sa phrase (*X empêche Y de Z*).

¹⁶² Rappelons que la FL Func prend pour sujet la base de la collocation et a pour collocatif un verbe support sémantiquement vide.

¹⁶³ En effet, même si l'élève a produit la forme **ce fut*, nous croyons qu'il voulait en fait utiliser le verbe FAIRE et produire *se fit*.

¹⁶⁴ Nous soulignons de nouveau que ces verbes « passe-partout » correspondent à des collocatifs tout à fait valables auprès de nombreuses UL.

croire que l'élève ne dispose tout simplement pas d'une valeur plus satisfaisante pour la base en jeu. Ceci devient encore plus évident pour des erreurs comme celle de la phrase (187), qui correspond à une FL sémantiquement pleine (Real₁); plutôt que de fournir le collocatif approprié, l'élève paraphrase en quelque sorte le sens de la valeur visée à l'aide d'un verbe, dans ce cas-ci RÉDUIRE (faire ce que l'on doit faire avec sa soif = causer qu'elle devienne moins importante → *réduire sa soif). En (188), les guillemets utilisés autour du collocatif erroné expriment selon nous la perplexité de l'élève devant le sens à exprimer (IncepOper₁) et son incapacité à trouver un verbe plus satisfaisant.

L'hypothèse d'un trou lexical dans le vocabulaire de l'élève est aussi évoquée dans les sept cas de formes analytiques rencontrés dans notre corpus. Plutôt que de fournir un terme précis, l'élève produit une paraphrase plus ou moins satisfaisante, comme dans les phrases suivantes.

- (189) ... la tondeuse, qui se trouvait dans un entrepôt avec le **coupeur de gazon** [taille-bordure] et le souffleur. [292]
- (190) Elle n'était pas très sur de ce qu'elle devait faire pour transformer Alexandre en **homme sirène** [triton]. [174]
- (191) Elles consentirent à se départir d'un grand nombre d'objets désuets pour atteindre leur famélique objectif : poupées, **bicyclettes à trois roues** [tricycles], livres de contes... [265]
- (192) Plus tard, on voit un autre groupe de filles, à **une autre location** [ailleurs], courir vers la même place que les autres. [157]
- (193) ... ils nous demandent si nous avons besoin d'une **intervention d'âge** [cure de rajeunissement]. [82]

Dans ces phrases, l'utilisation d'une UL précise – comme dans les corrections suggérées – aurait été plus élégante, mais nous pouvons supposer que l'élève n'y avait pas accès. Il se peut tout d'abord que cette UL ne fasse tout simplement pas partie du vocabulaire de l'élève, comme c'est sûrement le cas pour les phrases (189) et (190). Par contre, il est aussi possible que l'élève, temporairement, n'ait pas pu récupérer une UL qu'il connaît par ailleurs et qu'il aurait en d'autres circonstances utilisée sans problème; par exemple, il y a tout lieu de croire qu'un élève francophone de troisième secondaire connaît les UL TRICYCLE (phrase (191)) et AILLEURS (phrase (192)). Cependant, il se peut qu'au moment de la rédaction, il n'ait pas pu récupérer ce terme en mémoire, ou encore qu'il n'ait simplement

pas trouvé problématique la formulation produite. Le cas de la phrase (193) est un peu différent, puisque **intervention d'âge* ne constitue pas véritablement une paraphrase de *cure de rajeunissement*, mais plutôt une approximation lexicale. Il semble en effet que l'auteur de cette phrase connaisse bien le concept de 'cure de rajeunissement', mais pas l'étiquette qui lui correspond; il aurait donc produit une formulation boiteuse à partir d'UL connues pour désigner ce concept.

Dans certains cas, l'emploi d'anglicismes peut aussi être justifié par l'absence d'une UL française correspondant au sens visé dans le vocabulaire de l'élève; c'est par exemple le cas de l'élève qui utilise *DUNK [86] plutôt que SMASH¹⁶⁵. Le fait de ne pas disposer d'une UL standard pour exprimer un sens pourrait aussi conduire un élève à avoir recours à une UL de registre familier. Par exemple, l'élève qui écrit *J'étais épuisée et complètement tannée d'eux!* [91] emploie-t-elle l'UL TANNÉ parce qu'elle ignore que cet adjectif est familier ou parce qu'elle ne dispose pas d'une UL de registre standard pour traduire ce sens? D'ailleurs, quel serait le bon adjectif pour remplacer TANNÉ dans ce contexte? Il s'agit pour un Québécois de la façon la plus naturelle d'exprimer ce sens.

IMPLICATIONS DIDACTIQUES

Les interventions didactiques liées à une lacune dans le vocabulaire de l'élève sont plus limitées, car il n'est pas toujours simple de trouver une forme à partir d'un sens à véhiculer.

Dans le cas de termes concrets, l'utilisation d'un dictionnaire visuel peut permettre à l'élève, voire à l'enseignant, de trouver comment on nomme telle ou telle réalité; il serait donc intéressant de disposer d'un tel ouvrage en classe et de montrer aux élèves de quelle façon on le consulte. Une autre méthode, dont l'efficacité est beaucoup moins sûre, est de consulter un dictionnaire en cherchant sous l'article d'un terme sémantiquement lié à celui que l'on cherche un renvoi à ce dernier, par exemple chercher BATEAU pour tenter de trouver quel nom désigne 'la partie du bateau sur laquelle on circule' (PONT) ou quel nom exprime 'le fait qu'un bateau coule' (NAUFRAGE). De tels renvois peuvent être explicites ou « se cacher » dans les exemples. Dans le cas où l'élève est conscient qu'une UL lui fait défaut pour exprimer un sens précis, il peut aussi

¹⁶⁵ DUNK n'apparaît dans aucun des trois dictionnaires que nous avons utilisés; c'est dans le *Grand dictionnaire terminologique* de l'OQLF que nous avons vu que le terme recommandé est SMASH. Ceci laisse croire que même si l'élève avait cherché ce terme dans le dictionnaire, il n'aurait pas trouvé l'UL appropriée pour exprimer ce sens.

demander à l'enseignant ou à des camarades de classe s'ils connaissent l'UL visée.

Lorsque l'élève utilise en production écrite une paraphrase lourde pour exprimer un sens qui correspond à une UL précise, l'enseignant devrait proposer cette UL à l'élève sur sa copie corrigée. Dans le cas où l'élève emploie un anglicisme ou un terme familier parce qu'il ne connaît manifestement pas le terme standard correspondant, il est aussi important de lui fournir ce terme, ou tout au moins de l'orienter vers les ressources pertinentes, en lui demandant par exemple de chercher l'UL problématique dans le dictionnaire.

4.2.6 UNE CONNAISSANCE PARTIELLE DU SENS OU DE LA FORME D'UNE UL

Un certain nombre des erreurs analysées (5,6 %) nous semblaient pouvoir être expliquées simplement par une connaissance partielle du sens ou de la forme de l'UL erronée.

Cette hypothèse permet entre autres d'expliquer la confusion entre des termes sémantiquement proches, synonymes ou non, comme dans les phrases présentées plus bas. Cette explication rejoint en quelque sorte l'idée de proximité sémantique évoquée au point 4.2.2; seulement, dans les problèmes sémantiques classés ici, l'élève utilise la plupart du temps un hyponyme du sens visé, mais incompatible avec le contexte, ce qui nous laisse croire que l'élève ne maîtrise pas ses nuances de sens.¹⁶⁶

(194) ... elle vit Mickaël du haut de son bateau plus **sexuel** [sensuel, attirant] que jamais. [184]

(195) Le lendemain, il a **aperçut** [vit] une jeune femme pleurée. Il décida d'aller la voir. [168]

(196) Du haut de ses quinze ans, elle se laissa **enjôlée** [charmée] par cupidon et le coup de foudre entame alors son oeuvre. [263]

(197) Imaginez-vous que par-dessus tout ça, je suis obligée d'**escorter** [accompagner] mes parents dans un chalet miteux perdu au fin fond d'un vide total. [278]

Dans ces quatre phrases, le choix lexical opéré par l'élève n'est pas complètement mauvais, mais certaines composantes de sens de l'UL sont problématiques en regard du contexte; sans doute ces erreurs résultent-elles

¹⁶⁶ En fait, nous constatons à posteriori que les deux scénarios d'explication (*proximité sémantique* et *connaissance approximative du sens*) sont très similaires et auraient peut-être gagné à être regroupés pour ce qui concerne les erreurs de sens.

d'une connaissance imparfaite du sens de l'UL employée. L'auteure de la phrase (194) voulait sûrement exprimer que Mickaël, par sa sensualité, stimulait les instincts (notamment sexuels) de l'héroïne du texte; bien sûr, la parenté formelle entre SEXUEL et SENSUEL est sans doute aussi en jeu dans cette erreur, que l'on pourrait aussi expliquer par une influence du terme anglais SEXY, qui signifie 'sensuel'. En (195), l'élève ne semble pas conscient qu'APERCEVOIR, c'est 'voir, mais **après un effort d'attention et plus ou moins nettement, ou encore de façon assez brève**'.¹⁶⁷ L'élève qui utilise ENJÔLER en (196) ne semble pas percevoir l'intention de duperie incluse dans le sens de ce verbe, tout comme celle qui croit escorter ses parents (phrase (197)), mais qui, en réalité, ne fait que les accompagner, puisqu'il n'est jamais question dans le texte de les protéger ou de les surveiller. Dans tous ces cas, la personne ayant commis l'erreur possède une bonne connaissance du sens général de l'UL employée, mais cette connaissance demeure à raffiner, ce qui peut se faire soit par plusieurs rencontres successives de l'UL en contexte, soit par un apprentissage plus formel, comme la comparaison de termes sémantiquement proches à l'aide d'un dictionnaire.

Beaucoup d'erreurs liées aux contraintes sémantiques pesant sur les actants des UL ont aussi été associées à cette hypothèse. Souvent, le sens de l'UL utilisée convient assez bien au contexte dans lequel elle apparaît, mais l'élève ne semble pas conscient qu'elle est incompatible avec l'un de ses actants.

(198) *Les attaquant se virent anéantis, réduits en poussière et **détériorés** [affaiblis, blessés]. [63]*

(199) *Elle fabriqua un panier d'osier et y **enferma** [plaça] de la viande faisandée. [160]*

(200) *L'oiseau leur donna une de ses plumes **en guise de** [pour leur, en guise de reconnaissance] courage. [116]*

En (198), *détérioré* traduit assez bien le sens visé par l'élève; seulement, ce dernier semble ignorer que DÉTÉRIORER signifie 'mettre **quelque chose** en mauvais état' et ne peut donc pas avoir de complément direct dénotant un être animé. Le cas d'ENFERMER présente un problème de compatibilité avec deux de ses actants, selon le sens que l'on retient. Une des acceptions de ce verbe à trois actants (X

¹⁶⁷ Le désir de ne pas répéter deux fois le verbe VOIR, présent dans la phrase suivante, a peut-être aussi poussé l'élève à choisir un « synonyme ».

enferme Y dans Z) signifie ‘mettre dans lieu d’où il est impossible de sortir’ (ex. *Ils ont enfermé l’araignée mortelle dans un bocal.*); or, la viande faisandée, qui correspond ici au deuxième actant, n’a pas la faculté de sortir d’où que ce soit (à moins bien sûr d’être vraiment très faisandée!). Une autre acception d’ENFERMER signifie ‘mettre en lieu sûr’ (ex. *Les millions de la vieille dame sont enfermés dans un coffre dont elle seule a la combinaison.*); dans ce cas, c’est plutôt le troisième actant qui pose problème, puisqu’on peut s’interroger sur le caractère sécuritaire d’un panier d’osier. La mauvaise utilisation de la locution prépositionnelle EN GUISE DE, rencontrée deux fois dans notre corpus [109, 116], peut aussi être expliquée par une maîtrise partielle de son sens. L’élève comprend qu’EN GUISE DE signifie que l’on offre quelque chose à quelqu’un ou que l’on pose un geste à son endroit pour manifester un sentiment, mais il utilise comme complément de la préposition, ce qui suscite ce sentiment plutôt que le sentiment lui-même; il faudrait plutôt dire *en guise de reconnaissance pour son courage.*

Deux élèves ont utilisé dans leur rédaction le nom VIEUX [224, 240] au sens de ‘vieillard’. Cet emploi est problématique d’une façon un peu différente des erreurs présentées plus haut. En fait, l’UL VIEUX_N correspond tout à fait au sens visé par l’élève; cependant, d’un point de vue pragmatique, il est un peu gênant, vu sa connotation un peu péjorative, de l’employer dans un contexte de narration dans une production écrite. Pas très politiquement correct, en effet, de parler de *vieux* lorsqu’on raconte sa première expérience de travail dans un centre pour personnes âgées! Dans de tels cas, l’élève maîtrise bien le sens de l’UL, mais n’est pas conscient de son poids connotatif, ce qui peut poser problème d’un point de vue pragmatique.

Comme nous venons de le voir, une connaissance approximative du sens d’une UL peut être à l’origine d’erreurs sémantiques ou pragmatiques. De la même façon, plusieurs des barbarismes observés ont selon nous comme source une connaissance approximative de la forme de l’UL visée. Citons en exemples **dédaigner* [60] pour DÉGAINER, **tornitruant* [265] pour TONITRUANT et les locutions

déformées **au périple de sa vie* [60] pour AU PÉRIL DE SA VIE et **faire une pierre deux coups* [139] pour FAIRE D'UNE PIERRE DEUX COUPS.

IMPLICATIONS DIDACTIQUES

Notre connaissance du sens d'une UL se raffine au fil des rencontres que nous avons avec cette dernière; c'est donc principalement par la lecture que les élèves approfondiront leur compréhension du sens des UL qu'ils utilisent.

Néanmoins, lorsqu'une erreur découle de la connaissance imparfaite du sens d'une UL, il est important de le mentionner à l'élève. Il est aussi possible de faire avec lui un travail sur le sens du terme problématique à l'aide du dictionnaire; dans certains cas, la définition proposée permettra de mettre en évidence les composantes de sens problématiques, c'est-à-dire celles qui entrent en conflit avec le contexte dans lequel l'UL est utilisée. Lorsque ce n'est pas le cas, l'enseignant peut lui-même tenter d'explicitier le sens de l'UL erronée pour faire ressortir les caractéristiques sémantiques qui expliquent l'erreur, ce qu'il peut ensuite illustrer à l'aide d'exemples d'utilisation (phrases tirées du dictionnaire, ou simplement du Web). Encore une fois, la comparaison avec des synonymes peut aider à faire ressortir certaines propriétés sémantiques tout en aidant l'élève à structurer son vocabulaire. Bien qu'elle n'ait pas donné lieu à un nombre important d'erreurs de notre corpus, il pourrait être intéressant d'aborder la notion de connotation en classe, notamment en montrant comment des termes synonymes peuvent avoir un poids connotationnel différent (ex. VIEUX_N ~ VIEILLARD).

Dans le cas d'erreurs attribuables à une connaissance approximative de la forme d'une UL, la seule intervention possible est à notre avis de fournir la forme correcte à l'élève, puisque lui demander de la trouver par lui-même à l'aide du dictionnaire risque d'être trop difficile.

4.2.7 LE CONTEXTE SYNTAXIQUE

Certaines des erreurs de notre corpus (4,4 %) ont selon nous été induites par le contexte syntaxique dans lequel elles apparaissent, ce qui nous permet de supposer que l'élève qui les a commises ne les reproduirait pas dans un environnement syntaxique différent. Ces erreurs concernent principalement la combinatoire syntaxique et les collocations. Trois contextes syntaxiques semblent favoriser ces types d'erreurs : la coordination, la présence d'un écran entre un élément régi et l'UL qui le régit et la pronominalisation d'un complément.

La coordination de verbes semble constituer la configuration syntaxique la plus propice aux erreurs. En voici quatre exemples.

- (201) *Je vous ai **décrit** et informé **des** techniques publicitaires utilisées [ai décrit la publicité et informé des techniques]. [136]*
- (202) *Nathan ce décida donc à faire des C.V et **d'** [à] aller les porters à une vingtaine d'endroits. [45]*
- (203) *J'**enfilai** immédiatement mes souliers verts et ____ [pris, saisis] mon sac à poids, puis partis distribuer mon CV partout. [284]*
- (204) *Tes parents verront que tu n'es plus un enfant, donc ils te **laisseront** surment plus de liberté et [te donneront, t'accorderont plus] de permissions. [235]*

En (201), nous observons une erreur fréquente de coordination de verbes n'ayant pas le même régime; en coordonnant ainsi les verbes DÉCRIRE et INFORMER, l'élève les traite comme deux verbes transitifs indirects, alors que seul le second l'est. Par contre, il nous apparaît peu probable que l'élève croie vraiment que la structure syntaxique régie par DÉCRIRE soit *X décrit **de** Y à Z. Généralement, lorsqu'on observe ce genre d'erreur, c'est le régime du dernier verbe coordonné, soit le plus proche du complément, qui s'impose.

Le cas de la phrase (202) est différent, puisque ce sont ici deux compléments régis par un même verbe qui sont coordonnés. Étrangement, l'élève introduit le premier complément avec la préposition À, ce qui est correct, et le second avec la préposition erronée DE. Le choix de la préposition s'avère plus compliqué lorsque le verbe qui la régit se trouve à distance, et on observe encore une fois que c'est le complément le plus près du verbe qui est bien construit. Dans ce cas, on peut aussi se demander s'il n'y a pas eu confusion entre l'emploi pronominal de DÉCIDER (X se décide à Y.) et son emploi non pronominal (X décide **de** Y.).

La coordination de la phrase (203) entraîne une erreur de nature sémantique plutôt que syntaxique, puisqu'il y a incompatibilité entre le verbe ENFILER et le deuxième de ses compléments coordonnés (*sac à pois*), qui ne dénote pas quelque chose que l'on peut enfiler. Encore une fois, il y a fort à parier que l'élève n'aurait jamais écrit **enfiler un sac*, mais que c'est la coordination et la distance entre les deux éléments qui est à la source de l'erreur.

La phrase (204) reprend une erreur de collocation que nous avons déjà présentée dans la section traitant de cette classe d'erreurs, et qui est selon nous attribuable à la coordination de deux bases (*liberté, permissions*) pour un même collocatif (*laisser*) alors que ce collocatif n'est pertinent que pour l'une des deux bases : la plus proche.

Si plusieurs erreurs de coordination sont observées dans le corpus, c'est surtout parce qu'il existe une certaine distance, dans les phrases problématiques, entre l'élément régi par une UL et cette UL. Ainsi, il n'est pas nécessaire qu'il y ait coordination pour favoriser l'apparition d'erreur. Pour certaines des erreurs observées, un simple élément faisant écran entre une UL et son complément pouvait occasionner une erreur. Cet élément était parfois d'une longueur assez importante, comme le modificateur *plus que tout au monde* dans l'exemple (205), qui interfère entre le verbe DÉSIRER, transitif direct, et la subordonnée infinitive qui lui sert de complément¹⁶⁸. Dans d'autres cas, même un élément aussi court qu'un simple adverbe peut selon nous brouiller l'élève quant au choix de la préposition, comme dans l'exemple (206).

(205) *Elle désirait plus que tout au monde d' [---] aller à Londres. [214]*

(206) *En plus, ils réussissent souvent de [à] nous convaincre... [55]*

Certaines des erreurs liées à la combinatoire syntaxique des verbes présentaient des compléments pronominalisés. Dans la majorité des cas, il nous semblait probable que l'élève connaisse le régime approprié, mais que ce soit précisément la pronominalisation qui pose problème. En voici trois exemples.

(207) *Pourquoi j'ai choisi cette publicité est une très bonne question et je vais vous la [y] répondre. [235]*

(208) *Il l'aimait aussi cette fille, mais il avait décidé de ne plus la voir, car il l'avait **avoué** [lui avait avoué] qu'il était vampire. [172]*

(209) *Ses efforts allaient **lui récompenser** [le récompenser] car, il appercu une énorme montagne plus loin devant lui.¹⁶⁹ [78]*

¹⁶⁸ Il n'est pas tout à fait exclu, toutefois, que l'élève croie réellement que la structure du verbe DÉSIRER est *X désire de faire Y.

¹⁶⁹ La correction la plus élégante pour cette phrase serait bien sûr *Ses efforts allaient être récompensés car, il appercu une énorme montagne plus loin devant lui.*, mais nous avons choisi de rester le plus près possible de la phrase produite par l'élève.

L'erreur de la phrase (207) nous permet de poser l'hypothèse que l'élève croit que le verbe RÉPONDRE est transitif direct; par contre, nous ne sommes pas convaincu que, sans pronominalisation, l'élève aurait vraiment produit **Je vais vous répondre la question*¹⁷⁰. Et il nous semble encore plus improbable que l'auteur (208) pense réellement que la structure régie par le verbe avouer est **X avoue Y Z* plutôt que *X avoue Y à Z*; à notre avis, c'est lorsque l'élève a voulu pronominaliser le troisième actant du verbe que l'erreur s'est produite. Le même raisonnement vaut aussi pour l'erreur de la phrase (209), dans laquelle le deuxième actant de RÉCOMPENSER prend la forme d'un complément indirect plutôt que direct (*X récompense Y pour Z*).

Un sous-groupe d'erreurs, cette fois liées à l'absence d'un complément obligatoire, nous semble aussi en lien avec le contexte syntaxique. Dans les exemples suivants, il nous apparaît tout à fait probable que les élèves aient omis le complément non pas parce qu'ils ignoraient la structure des verbes en jeu, mais plutôt parce que le contexte dans lequel ils apparaissaient permettait de rétablir si facilement le complément manquant qu'ils ne voyaient sans doute pas la pertinence de le reprendre par un pronom. Ce phénomène est d'ailleurs très fréquent à l'oral et ne révèle selon nous aucune lacune lexicale chez les élèves; il s'agit de problèmes purement syntaxiques.

(210) *Alice revait d'approcher Mickaël, l'homme le plus sensuel de l'île, mais elle n'___ [en] avait pas l'opportunité. [184]*

(211) *Je voudrais lui dire qu'il est beau, mais je n'ai pas eu l'occasion de le maquiller. Non, au lieu de **lui dire** [le lui dire] je vais lui demander son numéro de téléphone. [272]*

IMPLICATIONS DIDACTIQUES

Les notions d'actant sémantique et de structure prédicative sont cruciales pour éviter les erreurs attribuables à des configurations syntaxiques précises et peuvent même, selon nous, aider les élèves à faire de meilleures analyses syntaxiques. Il faudrait donc à notre avis présenter ces notions

¹⁷⁰ Un autre problème permet peut-être d'expliquer cette erreur : le fait que le verbe RÉPONDRE n'admet généralement qu'un seul complément. En effet, il nous paraît étrange d'écrire une phrase comme *Je lui ai répondu à la question.* ou *J'ai répondu à la question à Marie.*; on dirait plutôt *J'ai répondu à sa question.* et *J'ai répondu à la question de Marie.*

lexicosyntaxiques en classe et montrer aux élèves comment trouver les informations relatives au régime dans les ouvrages de référence.

Pour bien analyser une phrase, il est primordial que l'élève puisse rétablir la structure régie par le verbe qui constitue le noyau de cette dernière. Cette structure doit demeurer intacte, peu importe le contexte syntaxique dans lequel le verbe apparaît. En dégageant cette structure, il devient plus simple d'analyser une phrase en retirant les éléments non essentiels, qui souvent font écran entre le verbe et ses compléments; le choix d'une préposition devient alors plus limpide, et l'oubli d'une préposition ou d'un complément entier beaucoup plus évident.

Il faut aussi sensibiliser les élèves à la coordination, phénomène fréquent qui peut favoriser l'apparition d'erreurs. Encore une fois, en établissant la structure régie par chacun des éléments de la coordination, l'élève pourra davantage éviter de produire des phrases dans lesquelles un même complément est introduit par deux verbes contrôlant une structure différente. Il serait aussi pertinent d'apprendre aux élèves à être attentifs lorsqu'ils pronominalisent un complément, afin qu'ils choisissent le bon type de complément; ceci touche principalement l'opposition LE (= CD) ~ LUI (= CI). Dans de tels cas, le recours à la phrase de base peut s'avérer un outil précieux.

4.2.8 LA DÉRIVATION MORPHOLOGIQUE

L'application de règles de dérivation morphologique, parfois productives par ailleurs, semble aussi être à l'origine de certaines erreurs du corpus (2,6 %). Bien sûr, cette hypothèse de source d'erreurs concerne principalement les formes inexistantes produites par les élèves. Dans les trois exemples ci-dessous, l'élève a produit une forme inexistante par l'application d'une règle de dérivation du français.

(212) ... toute heureuse d'avoir enfin terminé cette dernière journée d'école
infinissable [interminable]. [280]

(213) L'**exagèremment** [exagération] de cette publicité est assez intense. [134]

(214) Il aperçu le roi qui **inquentait** [prononçait] des formules magique. [12]

Le cas de la phrase (212) est le plus simple. L'élève a utilisé le suffixe *-able*, avec lequel de nombreux adjectifs ont été créés à partir de verbes en leur ajoutant le sens 'qui peut être'; cette règle de dérivation pourrait être modélisée comme suit :

Radical_V + -ABLE → Radical_{ADJ} 'qui peut être'¹⁷¹

Par contre, comme la dérivation morphologique n'est pas un phénomène très productif en français de façon diachronique, l'élève a généré – sans doute sur le modèle d'INTERMINABLE – une forme inexistante qui est par conséquent considérée comme une erreur, malgré le fait qu'elle correspond à une règle de formation des adjectifs tout à fait valable.

Le cas de la phrase (213) n'est pas quant à lui associé à une règle de dérivation aussi évidente que celui de la phrase précédente. En effet, aucun ouvrage de référence, même *Le bon usage*, ne répertorie le suffixe *-ment* comme un suffixe permettant de former un nom à partir d'un verbe, à la façon du suffixe *-ation* (ex. ADAPTATION, EXAGÉRATION)¹⁷². Pourtant, c'est sans doute en s'appuyant sur des noms comme AFFRONTEMENT (*affront*-_v + *-(e)ment*), BRANCHEMENT (*branch*-_v + *-(e)ment*), ÉTIREMENT (*étir*-_v + *-(e)ment*), JUGEMENT (*jug*-_v + *-(e)ment*) et TREMBLEMENT (*trembl*-_v + *-(e)ment*) que l'élève en est arrivé à produire *EXAGÈREMENT. Encore une fois, cette forme, bien que justifiable d'un point de vue linguistique, n'est attestée nulle part et ne fait donc pas partie du lexique français.

La phrase (214) présente quant à elle une erreur qui semble produite selon la règle inverse. Plutôt que de produire un nom à partir d'un verbe à l'aide d'un suffixe, il semblerait que l'élève ait créé un verbe en retirant ce qu'il croyait être le suffixe *-ation* du nom INCANTATION. Ainsi, de la même façon que l'on peut associer DÉCORATION à DÉCORER, CONDAMNATION à CONDAMNER et ADAPTATION à ADAPTER, l'élève a produit le verbe *INCANTER, ce qui paraît une fois de plus cohérent avec le système de la langue, mais n'est pas conforme à la norme.

¹⁷¹ Le suffixe *-ABLE* apparaît en majuscules pour exprimer qu'il s'agit d'un archimorphème, c'est-à-dire un morphème qui peut prendre des formes différentes selon le contexte dans lequel il apparaît, mais qui ajoute le même sens au radical. Les formes *-able* (ex. *jetable* = 'qui peut être jeté', *inoublable* = 'qui ne peut être oublié') et *-ible* (ex. *lisible* = 'qui peut être lu', *indescriptible* = 'qui ne peut être décrit') sont des variantes de l'archimorphème *-ABLE* et l'ajout de sens est le même dans les deux cas.

¹⁷² On retrouve bien sûr dans les ouvrages le suffixe *-ment* permettant la formation d'adverbes à partir d'adjectifs (ex. GENTIMENT, SOURNOISEMENT).

Ces trois erreurs témoignent en fait d'une conscience linguistique certaine de la part des élèves les ayant commises. Ces derniers ont appliqué des règles de dérivation tout à fait exactes pour produire des formes visant à exprimer un sens précis. Malheureusement pour eux, la cible de leur dérivation, de façon presque accidentelle, n'appartient pas au système lexical de la langue. Comment expliquer, en effet, que l'adverbe *MORNEMENT [265] n'existe pas, alors que l'on trouve TRISTEMENT, SOMBREMENT, MÉLANCOLIQUEMENT et MAUSSADEMENT? Tout cela demeure très arbitraire et les élèves ayant produit les erreurs présentées auraient pu appliquer les mêmes règles à partir d'autres radicaux et auraient produit des formes correctes. Observons maintenant quelques autres erreurs, cette fois classées parmi les problèmes sémantiques, mais qui découlent encore une fois selon nous de dérivations erronées.

(215) **Dernièrement** [Finalement], un slogan apparaît également au bas de la page. [159]

(216) Regolroh sortit sa toge de son sac et l'enfila. Cette toge protégeait le **porteur** [celui qui la portait] contre les ondes de l'horloge. [105]

(217) Pourquoi certaines publicités sont-elles **influçables** [efficaces, persuasives; nous influencent-elles] tandis que d'autres ne le sont pas? [128]

(218) ... il est temps de se **formater** [former? s'occuper?] avec un emploi d'été. [233]

Dans ces quatre phrases, les élèves ont utilisé des formes existantes, mais dans un sens qui ne leur correspond pas, à notre avis sur la base de mauvaises dérivations. En (215), notamment, il nous paraît évident que l'élève a produit **dernièrement* au sens de 'finalement' sur le patron de PREMIÈRE, DEUXIÈMEMENT, TROISIÈMEMENT. Dans la phrase (216), l'élève a sans doute dérivé le nom PORTEUR du verbe PORTER, mais sans savoir qu'il est bizarre de parler du *porteur d'un vêtement*, hormis peut-être dans le cas du porteur du maillot jaune au Tour de France. L'auteur de la phrase (217) a quant à lui produit un adjectif tout à fait correct à partir du verbe INFLUENCER en lui accolant le suffixe *-able*; cependant, il ne semble pas maîtriser le sens de ce suffixe ('qui peut être...'), puisqu'il l'a utilisé pour exprimer 'qui a la capacité de (influencer)'. Il est vrai qu'il n'existe pas en français d'adjectif dérivé d'INFLUENCER pour exprimer ce sens, sauf INFLUENT, qui s'applique seulement à des personnes ou des groupes de personnes. Le cas de la phrase (218) nous a laissé plus perplexe, car nous n'arrivions vraiment pas à

comprendre pourquoi l'élève avait utilisé le verbe FORMATER dans ce contexte... jusqu'à ce que nous relisions la consigne d'écriture fournie par l'enseignant : « En quoi un emploi d'été est-il formateur pour un jeune? »!

IMPLICATIONS DIDACTIQUES

Ces erreurs en lien avec la dérivation démontrent l'importance de travailler ce thème en classe avec les élèves. Un travail sur le sens des principaux préfixes et suffixes du français peut sans doute s'avérer très utile pour les élèves, particulièrement pour ce qui est de la compréhension en lecture lorsque des mots inconnus les contenant sont rencontrés.

En ce qui concerne l'écriture, nous croyons qu'il est surtout important de faire comprendre aux élèves que, malgré l'existence d'un petit nombre de préfixes et de suffixes récurrents, le français n'est pas une langue dans laquelle la dérivation est très productive de façon synchronique. C'est pourquoi, lorsqu'on « créer un mot nouveau » à l'aide d'un préfixe ou d'un suffixe, mieux vaut vérifier son existence dans le dictionnaire.

Cela dit, il ne faut pas oublier que des erreurs comme celles présentées plus haut témoignent d'une certaine compréhension du système morphologique de la langue et qu'une intervention mettant en lumière la logique derrière l'erreur, notamment en fournissant des exemples pour lesquels le même préfixe ou le même suffixe aurait fonctionné, pourra encourager les élèves les ayant commises. De façon plus ludique, on pourrait même proposer des activités de création lexicale aux élèves, à partir de préfixes et de suffixes fréquents.

4.2.9 QUELQUES HYPOTHÈSES PLUS MARGINALES

Les huit sources possibles d'erreurs que nous avons présentées jusqu'à maintenant sont celles qui sont revenues le plus souvent parmi nos hypothèses – elles totalisent 86,7 % des scénarios d'explication proposés – et constituent selon nous les plus pertinentes pour expliquer les erreurs lexicales d'élèves francophones du secondaire. Bien entendu, certaines hypothèses plus marginales ont aussi été avancées et, sans les présenter en détail, nous jugeons pertinent d'en survoler quelques-unes.

Certaines erreurs, principalement liées à l'invariabilité, semblaient motivées par une analyse sémantique de la part de l'élève. Par exemple, quand un élève écrit *Plus *personnes ne souffrent de la chaleur.* [60], c'est sans doute parce qu'il comprend que 'toutes les personnes du village ont cessé de souffrir de la chaleur' qu'il fait varier le pronom indéfini PERSONNE en nombre. Le même type de raisonnement

vaut aussi pour l'élève qui accorde ENSEMBLE avec le groupe sujet dans la phrase *Billy et Camille passe tout leurs temps *ensembles*. [168], puisqu'il est vrai que sémantiquement, cet adverbe s'apparente beaucoup à un adjectif.

Évidemment, nombre d'erreurs du corpus sont potentiellement attribuables à la distraction, voire à la négligence, diraient peut-être certains enseignants. Par contre, il est difficile, à la seule lumière de la production finale de l'élève, de déterminer que la distraction est à l'origine d'une erreur. Néanmoins, il nous semblait plutôt improbable que des élèves francophones produisent certaines erreurs, que nous croyons difficile d'expliquer autrement que par l'inattention. Beaucoup de ces erreurs étaient liées à la forme : **transforma* [113] pour *transforma*, **lendemain* [284] pour *lendemain*. Dans d'autres cas, un nom utilisé correctement plusieurs fois ailleurs dans le texte apparaissait avec le mauvais genre (**[une] emploi* [211], **[cette] été* [269]).

Des ellipses dans certains textes d'élèves permettent aussi d'expliquer quelques-uns des dysfonctionnements lexicaux observés. Ces raccourcis sémantiques se traduisent par l'omission de certains éléments pourtant essentiels dans la phrase, comme dans les exemples suivants.

(219) ... *les produits de beauté qui se font publiés pour être vendus [que l'on publicise dans les revues, dont les publicités paraissent dans les revues].* [82]

(220) *Les premiers obstacles du superhéro sont très faciles [faciles à surmonter].* [54]

La confusion entre des expressions ou le chevauchement de certaines constructions nous ont aussi semblé des sources possibles d'erreurs. Par exemple, l'élève qui a écrit **Des cordes de pluie se mirent à tomber*. [241] a sans doute confondu les collocations *la pluie tombe* et *pleuvoir des cordes*, tout comme celui qui a produit **au profond de l'eau* [185] a manifestement amalgamé *au fond de l'eau* et *au plus profond de l'eau*. Des chevauchements similaires apparaissent lorsque l'élève tente d'exprimer le sens de deux verbes à la fois, comme dans les phrases suivantes.

(221) *Rendu à destination ils appâtèrent leurs lignes à l'eau [appâtèrent leurs lignes et les mirent à l'eau].* [115]

(222) *Ce truc professionnel est pour [...] nous divertir à se procurer le produit annoncé [nous divertir tout en nous poussant à se procurer...].* [82]

Nous avons parlé plus haut de lacunes dans le vocabulaire de l'élève, de « trous lexicaux », comme source possible d'erreurs. Mais il se peut aussi, beaucoup plus rarement il va sans dire, que ces vides lexicaux soient inhérents au système de la langue elle-même plutôt qu'à l'élève. Reprenons le cas de l'élève qui a employé TRAVAIL au sens d' 'emploi' au pluriel (*un très bon tournant pour tes autres **travails** dans le futur.* [39]). En utilisant la forme inexistante *travails*, cet élève a (consciemment ou non) rempli un vide du système lexical français; en effet, le pluriel « normal » de TRAVAIL, soit *travaux*, ne peut renvoyer au sens visé dans ce contexte ('emploi'). Ces vides lexicaux nous apparaissent spécialement lorsqu'on devient sensible au phénomène de collocation. Combien de fois nous arrive-t-il, lors d'une conversation, de chercher une valeur pour une FL précise à exprimer auprès d'une base, et de constater que la langue ne nous en offre aucune! Par exemple, existe-t-il un nom correspondant à l'adjectif COURT de la même façon que LONGUEUR correspond à LONG?

La dernière source possible d'erreurs que nous voulons mentionner est très peu représentée dans notre corpus, mais le serait sans doute beaucoup plus si nous avions analysé les textes d'élèves n'ayant pas le français pour langue première : l'influence d'une L2. Cette hypothèse n'a été retenue que dans les cas où l'élève ayant commis l'erreur avait indiqué sur son formulaire de consentement qu'il parlait couramment une autre langue à la maison, en plus du français. Les quelques erreurs pouvant être expliquées par une influence d'une L2 nous semblaient provenir de l'anglais, langue parlée à la maison par un certain nombre d'élèves. En voici trois exemples : **mosse verte* [10] (ang. MOSS = 'mousse'), *tous les travailleurs de l'*office* [146] (ang. OFFICE = 'bureau'), **pour le bon Dieu!* [152] plutôt que *pour l'amour de Dieu!*, sans doute sous l'influence de *for God's sake!*

4.3 LA MULTIPLICITÉ DES SOURCES D'ERREURS POSSIBLES

Comme nous l'avons déjà mentionné au point 4.2, deux hypothèses ont été avancées pour 93 erreurs et cinq cas nous semblaient pouvoir être expliqués de trois façons différentes. La présente section s'intéressera précisément à ces

hypothèses multiples dans le but de déterminer s'il s'agit d'hypothèses concurrentes ou de sources d'erreurs qui agiraient en interaction. Mais d'abord, analysons trois erreurs ayant donné lieu à plus d'une hypothèse d'explication.

(223) *Il la voyait dans le cimetière à tuer ses **ressemblants** [semblables], les vampires. [172]*

Dans cette phrase, l'élève utilise le nom fictif *RESSEMBLANT, qui, selon toute vraisemblance (!), signifie 'semblable'. La proximité formelle et sémantique entre SEMBLABLE et RESSEMBLER ne peut être ignorée, ces deux UL étant issues d'une même famille étymologique; cette double proximité formelle et sémantique constitue donc une piste d'explication de l'erreur. Par contre, si l'élève a produit le nom inexistant *RESSEMBLANT, c'est selon nous à partir du verbe RESSEMBLER, sans doute parce qu'il se rappelait qu'il existait un nom apparenté à ce verbe, mais sans pouvoir accéder au nom SEMBLABLE. Nous pouvons donc ajouter à nos scénarios d'explication une difficulté d'accès à un terme possiblement connu et une dérivation morphologique abusive.

(224) *Mais quels sont les avantages de travailler et d'avoir des activités **pertinentes** [stimulantes; enrichissantes] pour votre été? [223]*

L'emploi de PERTINENT est problématique dans cette phrase, puisqu'on ignore en quoi les activités auxquelles l'auteur du texte fait référence sont pertinentes. Il est d'ailleurs difficile de rétablir avec certitude le sens visé par l'élève. Différentes hypothèses peuvent être évoquées pour expliquer cette erreur. Tout d'abord, on peut comprendre que l'élève a une connaissance très approximative du sens de l'adjectif PERTINENT; il semble comprendre que cela fait référence à quelque chose de plutôt positif, d'utile. Par ailleurs, il existe tout de même un certain lien de sens entre PERTINENT et les corrections que nous proposons : il est pertinent (au sens de 'judicieux') pour le développement et le bien-être d'un jeune qu'il pratique des activités enrichissantes. On pourrait aussi croire que l'élève utilise PERTINENT ici parce que ce terme est de plus en plus présent dans l'usage courant et est utilisé, depuis quelques années, un peu à toutes les sauces.

(225) *... une grosse boule blanche mesurant environ deux mètres 65 **velu d'un poil blanc** [recouverte d'un poil blanc; velue... dont le pelage blanc était] magnifique. [167]*

Dans la phrase (225), l'élève ajoute un complément à VELU, alors que cet adjectif n'a qu'un seul actant (*X est velu*); il fait cet ajout dans le but de modifier une des composantes de sens de VELU : 'poils' (VELU = 'couvert de poils'). Nous posons l'hypothèse que l'élève a produit cette structure sous l'influence de l'adjectif VÊTU, qui contrôle une structure à deux actants (ex. *L'homme est vêtu d'une longue tige.*) et ressemble à VELU d'un point de vue formel (quatre lettres, deux syllabes, mêmes phonèmes en début et en fin de mot). Les deux adjectifs partagent aussi des composantes de sens communes : l'idée de quelque chose qui recouvre et cache le corps. Cette erreur serait donc attribuable à une triple analogie : analogie formelle, sémantique et de construction.

On constate donc que les sources d'erreurs avancées ne sont pas mutuellement exclusives, mais peuvent plutôt agir en synergie, le meilleur exemple restant une parenté à la fois formelle et sémantique qui concourt à la confusion de deux termes. Une seule hypothèse nous paraît ne pas nécessairement agir en interrelation avec les autres : l'intégration de l'erreur à l'usage. En effet, lorsqu'un emploi erroné s'impose dans l'usage, c'est d'abord sur la base d'une parenté formelle ou sémantique avec un autre terme, d'une influence d'une autre langue, etc., mais si on l'utilise par la suite, c'est sans doute parce qu'il fait partie du vocabulaire de notre communauté linguistique. Cela revient donc à dire que le fait qu'un élève utilise le verbe SE RAPPELER de façon transitive indirecte (ex. **Il se rappelle de son enfance.*) n'est pas nécessairement attribuable à une analogie sémantique avec le verbe SE SOUVENIR, dont il aurait transposé le régime à SE RAPPELER; il reproduit sans doute simplement le régime de SE RAPPELER qu'il entend le plus fréquemment. Ceci rejoint encore une fois l'idée selon laquelle la plupart des anglicismes utilisés couramment le sont non pas sur la base d'une influence de l'anglais, mais simplement parce qu'ils sont intégrés à l'usage.

CONCLUSION

Voici donc ce qui conclut notre réflexion sur les sources possibles des erreurs lexicales rencontrées dans notre corpus. Outre l'intégration de l'erreur à la langue orale, nous retenons l'analogie comme la source principale d'erreurs lexicales, que cette analogie soit sémantique ou formelle.

Il semble donc que la clé pour des interventions didactiques soit la comparaison d'UL qui entretiennent des liens de sens (principalement des synonymes) ou qui présentent des ressemblances d'un point de vue formel. Ces UL pourront être contrastées relativement aux nuances de sens qui les distinguent – ce qui développera leurs habiletés d'analyse du sens, notamment par une meilleure compréhension des définitions des dictionnaires –, mais aussi relativement aux constructions syntaxiques et aux collocations qu'elles contrôlent. Mettre en lumière les caractéristiques communes partagées par des UL liées sémantiquement, par exemple un régime identique ou des collocatifs communs, pourra aussi aider les élèves à faire des liens entre ces UL et à organiser leur lexique mental.

DEUXIÈME OBJECTIF - MÉTHODOLOGIE

IX. MÉTHODOLOGIE – DEUXIÈME OBJECTIF

Notre deuxième objectif de recherche est le suivant :

2. Décrire le rapport d'enseignants de français du secondaire à l'erreur lexicale en ce qui concerne

- *leur conception de ce qu'est une erreur de nature lexicale;*
- *leurs pratiques d'évaluation du vocabulaire en production écrite;*
- *leurs modes de rétroaction par rapport à des erreurs lexicales.*

Différentes méthodologies de recherches sont utilisées pour mettre au jour les éléments constitutifs du rapport d'un sujet à un objet. Nous proposons d'abord un survol de ces différentes méthodes d'accès aux représentations/conceptions/pratiques avant de présenter la méthodologie associée à notre deuxième objectif de recherche.

1. LES MÉTHODOLOGIES LIÉES AU RECUEIL DES REPRÉSENTATIONS

La présente section s'intéressera aux méthodologies visant à recueillir les représentations de sujets relativement à un objet. Considérant la difficulté de distinguer les différents éléments composant le « rapport à » – représentations/conceptions, attitudes, opinions, etc. –, nous présenterons des méthodologies susceptibles de recueillir tous les types de manifestations des différents aspects du rapport à un objet (incluant les pratiques) en utilisant le terme *méthodologies liées aux représentations*, qui est celui employé dans la littérature de façon générale. Nous mettrons évidemment l'accent sur les méthodologies les plus susceptibles d'intéresser la didactique des langues et donc de nous être utiles dans le cadre de cette thèse; nous ferons par conséquent abstraction des techniques plus figuratives, reposant sur l'utilisation d'images ou la production de dessins ou de schémas.

Comme le souligne Abric (1994b) dans un article traitant précisément des différentes méthodologies associées au recueil de représentations, il existe deux

grands types de méthodes : les *méthodes interrogatives* et les *méthodes associatives*. Les premières consistent à susciter, par des questions ou des consignes, une expression des sujets (texte oral ou écrit, dessin) en relation avec l'objet de représentation visé. Les secondes reposent aussi sur une production verbale des sujets, mais recueillie de façon plus libre et spontanée; les données ainsi produites devraient donc être plus authentiques. À ces deux familles de méthodologies, nous pourrions en ajouter une troisième, que l'on pourrait qualifier de *méthodes d'observation*, qui consisterait à recueillir les manifestations des représentations des sujets par la seule prise en compte des traces et observables (Vermersch, 2006) produits par les sujets dans une situation normale. De telles méthodologies conduisent à une analyse reposant sur des données invoquées, plutôt que sur des données provoquées ou suscitées (Van der Maren, 1999), propres aux deux autres types de méthodologies.

1.1 LES MÉTHODES INTERROGATIVES

Les méthodes interrogatives, ou enquêtes (Lessard-Hébert, Goyette et Boutin, 1996), sont les plus utilisées dans les recherches en didactique s'intéressant aux représentations, voire dans les travaux sur les représentations en général. Il s'agit de l'entretien et du questionnaire, que nous aborderons tour à tour.

1.1.1 L'ENTRETIEN

L'entretien représente sans doute la méthode la plus évidente d'accès aux représentations et il existe des dizaines de types d'entretiens développés dans différents domaines des sciences humaines. Un entretien peut être plus ou moins structuré, selon le type d'information visé; comme le dit Van der Maren (1999), l'entrevue se situe « quelque part entre la conversation et le questionnaire ». Selon Lüdi et Py (2002), l'entretien semi-directif, à cause de son caractère plus naturel, est celui qui convient le mieux pour la mise au jour des représentations. Dans ce type d'entretien, l'intervieweur prévoit les thèmes à aborder et quelques questions, mais la manière d'amener les thèmes et de

formuler les questions ainsi que l'ordre de déroulement ne sont pas fixés d'avance (Legendre, 2005: 598).

L'entretien présente l'avantage, par un accès direct au sujet, de permettre un approfondissement de ses propos de façon dynamique et interactive, ce qui n'est pas le cas du questionnaire. Le contact avec le sujet facilite aussi la prise en compte de dimensions comportementales (hésitations, intonations, etc.) qui seraient occultées par d'autres méthodes de recherche.

Par contre, l'entretien présente aussi certains inconvénients bien connus. Tout d'abord, il exige un investissement de temps important (passation des entrevues, transcription, etc.) et peut nécessiter le recours à des moyens techniques (magnétophone, caméra, etc.) qui peuvent connaître des ratés (Van der Maren, 1996). Ensuite, la situation même d'entretien n'est pas, malgré tous les efforts fournis par le chercheur, une situation naturelle; elle implique une dissymétrie intervieweur/interviewé (voire novice/expert)¹⁷³ et des conditions plus ou moins typiques de celles dans lesquelles évolue généralement le sujet, ce qui peut bien sûr avoir une influence sur le produit de l'entretien (stress, désir de répondre ce que l'enquêteur « veut » entendre, etc.). Il peut donc être difficile de départager ce qui tient du contexte d'entretien de ce qui révèle une attitude ou opinion stable du locuteur (Abric, 1994b). Charlot (1997: 97-98) illustre très bien cette limite, en pointant un piège à éviter lors de l'interprétation des données :

[...] la situation de recueil de données tend à cristalliser sous forme de « représentation » ce qui n'existait peut-être pas comme contenu explicite de conscience avant que le chercheur ne pose la question. Tout jeune sera capable de dire ce qu'est un « bon élève » mais il n'est pas sûr pour autant qu'il avait déjà construit, « dans sa tête », une représentation explicite du bon élève avant que la question ne lui soit posée. Or, le chercheur sera tenté, après avoir « recueilli » cette représentation, d'expliquer certains comportements du jeune par sa représentation du bon élève. Dans ce cas, non seulement le chercheur

¹⁷³ Par la méthode de l'entretien compréhensif, Kaufmann (1996) tente de réduire cette dissymétrie par une plus grande implication de l'intervieweur à l'entretien (par exemple par la possibilité de ne pas rester neutre aux propos du sujet), qui se rapproche alors d'une conversation « normale », ce qui devrait pousser l'interviewé à se livrer davantage (Berger, 2006).

provoque la construction d'une image mentale, mais il confère à cette image le pouvoir d'être cause de conduites. [...] Certains jeunes ont sans doute une représentation du bon élève – ils y ont déjà réfléchi et peuvent délivrer au chercheur une image déjà construite et disponible. Mais le plus souvent le « bon élève » est une signification latente, au carrefour de divers rapports (aux exigences des enseignants, aux relations entre camarades, aux disciplines scolaires, à ce que l'on est prêt à sacrifier à l'école, etc.).

De plus, comme le souligne encore Abric (1994b: 61), l'utilisation du langage, sur laquelle repose l'entretien, implique des « mécanismes psychologiques, cognitifs et sociaux qui rendent problématiques la fiabilité et la validité des résultats : rationalisations, scotomisations, contrôle, obligation plus ou moins forte de cohérence, filtrages de tous ordres. » Ceci s'ajoute au fait qu'il est tout à fait possible qu'il y ait décalage entre ce qui est verbalisé et les représentations ou processus réellement en jeu dans la situation faisant l'objet de l'entretien. D'autre part, la technique de l'entretien induit un traitement des données qui repose sur l'analyse de contenu, dont on sait qu'elle implique une part importante de subjectivité, de par l'interprétation et les inférences qu'elle exige (Abric, 1994b ; Billiez et Millet, 2004)¹⁷⁴. Finalement, Abric (1994b), en lien avec la notion de noyau central, souligne que, si l'entretien constitue un bon moyen d'accès à une partie des contenus d'une représentation, il ne donne que peu d'information sur l'organisation de ces contenus.

En bref, l'entretien constitue un moyen riche et presque incontournable d'accès aux représentations, particulièrement pour enrichir la dimension qualitative d'une recherche. Par contre, à cause des multiples limites qu'elle présente, cette technique gagne à être jumelée à des approches complémentaires qui permettront de valider les données qui en résultent (voir 1.2). Voyons maintenant un type d'entretien particulièrement pertinent pour la recherche sur les représentations en didactique, qui fait partie de ce que Van der Maren (1999) appelle *entrevues cliniques*.

¹⁷⁴ Billiez et Millet (2004) indiquent qu'il est possible de diminuer l'impact de ce problème en se dotant d'outils visant à différer l'interprétation et en prenant en compte des données qualitatives autant que quantitatives.

1.1.1.1 L'ENTRETIEN D'EXPLICITATION

Fruit des apports de l'entretien non directif de Carl Rogers et des techniques utilisées en psychothérapie, l'entretien d'explicitation est défini et formulé par Vermersch (1994), qui transpose ces apports vers le domaine de la formation. La spécificité de ce type d'entretien est de viser la verbalisation de l'action (Vermersch, 2006) dans le but de saisir l'articulation des connaissances impliquées dans l'action et qui en détermine le succès ou l'échec. D'un point de vue de recherche en didactique, elle peut donc mettre au jour les processus en jeu dans une situation didactique et ainsi contribuer à faire émerger les représentations qui les sous-tendent.

L'entretien d'explicitation a souvent lieu après la tâche effectuée et le rôle de l'intervieweur se résume à recentrer le sujet sur la tâche accomplie, par des relances, en plus de lui fournir un environnement le plus près possible de l'environnement dans lequel la tâche a été réalisée. Trois types de produits sont impliqués dans l'entretien d'explicitation, qui peuvent s'étayer ou s'invalider les uns les autres :

1. *les traces*, par exemple un brouillon ou des réponses fournies préalablement à un questionnaire, etc.;
2. *les observables*, qui sont l'ensemble des comportements de l'interviewé, souvent inconscients de sa part et qui fournissent une multitude d'informations sur les processus en cours;
3. *les verbalisations*, qui permettent d'éclairer les processus déployés pendant la tâche et d'expliquer la raison d'être des traces et des observables.

Les principales limites de l'entretien d'explicitation, outre celles inhérentes à tous types d'entretiens (voir 1.1.1), sont les suivantes :

- le caractère peu naturel de la verbalisation de l'action, qui peut pousser le sujet à se limiter à des jugements et opinions, ou encore à des commentaires relatifs aux circonstances entourant la tâche plutôt qu'à son exécution.
- le fait que seuls les processus cognitifs conscientisés peuvent être verbalisés. Par exemple, dans des entretiens portant sur des activités de production linguistique, les processus épilinguistiques ne pourront

faire l'objet de verbalisations par le sujet, puisque celui-ci n'en a pas même conscience¹⁷⁵.

- le rappel des faits, qui peut être plus ou moins exact, puisque l'entretien est la plupart du temps réalisé après la tâche.
- les pièges que l'intervieweur doit éviter : 1) transposer sur le sujet sa propre façon de faire, 2) ne pas remettre en question des descriptions qui lui semblent évidentes ou considérer l'absence de verbalisations comme un gage de compréhension et 3) formuler des questions en « pourquoi », situées hors de l'action puisqu'elles relèvent du domaine des connaissances et s'avèrent moins efficaces que les « comment » pour favoriser la description de l'action.

L'entretien d'explicitation semble assez présent dans les méthodologies de recherches réalisées en didactique du français, particulièrement dans les travaux s'intéressant aux pratiques. Mentionnons par exemple la recherche de Pereira (2004) sur les liens entre les rapports des enseignants à l'écriture et leur mode d'enseignement de l'écriture, ainsi que les travaux de Dreyfus (2004), qui s'intéressent aux représentations des enseignants du primaire sur l'enseignement du lexique, notamment à travers leur pratiques. Ce type d'entretien permet de valoriser « essentiellement la description d'événements didactiques au détriment de la simple explication de la façon de les concevoir. » (Pereira, 2004: 326)

1.1.2 LE QUESTIONNAIRE

La seconde méthode d'enquête, qui est d'ailleurs la plus répandue dans la recherche sur les représentations (Abric, 1994b), est le questionnaire, qui présente des avantages certains. Tout d'abord, il permet des collectes de données à grande échelle qui favorisent la prise en compte d'éléments quantitatifs dans l'analyse des données, augmentant ainsi son potentiel de généralisation. De plus, son caractère standardisé réduit les risques de subjectivité – puisqu'il n'implique pas la présence d'un enquêteur – et les variations interindividuelles dans les thèmes abordés, leur ordre et les modalités de réponse (Abric, 1994b). Ces avantages ont pour corolaire une forme un peu

¹⁷⁵ Voir la distinction *activité épilinguistique* / *activité métalinguistique* en résolution de problèmes de langue dont parle Berthoud (1982).

plus rigide que celle de l'entretien, le questionnaire ayant un ordre de présentation statique et imposant un format de réponse à l'enquêté.

Évidemment, il existe plusieurs types de questionnaires dont les formats de réponses peuvent varier : alternative, choix multiples, échelle de valeur ou de fréquence, classement d'éléments, etc. (Van der Maren, 1996) Les questionnaires sont constitués soit de questions fermées (choix des réponses réduit), soit de questions ouvertes (contenu et forme laissés au choix du répondant) ou peuvent regrouper les deux types de questions (Lessard-Hébert *et coll.*, 1996). D'après les travaux sur les représentations que nous avons consultés, il semble que les échelles de Likert soient parmi les formats de questions les plus répandus (utilisées entre autres par Brodeur *et coll.*, 2003 ; Lavoie, Lévesque et Chénard, 2004 ; Lefrançois et Montesinos-Gelet, 2004 ; Chartrand et Blaser, 2008); il s'agit d'échelles d'accord utilisées pour que l'interviewé exprime son degré d'approbation ou son opinion à l'égard d'une série d'items (par exemple : excellent, très bien, bon, passable, insuffisant, nul). Lüdi et Py (2002) mentionnent aussi le procédé du différentiel sémantique comme un bon moyen d'accès aux représentations¹⁷⁶; il s'agit de demander au répondant de se situer quelque part entre deux adjectifs de sens parfaitement opposés (agréable/désagréable, utile/inutile, etc.). Dans nombre de travaux (voir notamment Brodeur *et coll.*, 2003 ; Lafont-Terranova et Colin, 2004 ; Pereira, 2004), des questions plus ouvertes permettent quant à elles d'éclairer les réponses fournies aux questions fermées.

Voici les principaux problèmes associés à l'utilisation de questionnaires comme méthode d'accès aux représentations. Tout d'abord, le questionnaire limite l'expression du sujet (Abric, 1994b), particulièrement dans le cas de questions fermées; ce problème se manifeste aussi à travers la sélection des thèmes abordés par le concepteur du questionnaire. De plus, le fait de proposer des choix multiples au répondant implique que le chercheur impose en quelque

¹⁷⁶ Les auteurs parlent en fait d'échelle d'attitude, mais cette méthode permet aussi l'accès aux représentations, ce qui nous ramène à la difficulté de bien distinguer les composante du « rapport à ».

sorte ses propres schémas de pensée au répondant, qui peut se voir contraint de répondre d'une façon ne correspondant pas à ses propres modes de pensée (Grawitz, 2001). Ensuite, la formulation des questions peut aussi remettre en cause, si elle laisse place à l'interprétation, la pertinence des données recueillies; le chercheur doit s'assurer que « ses questions permettent une compréhension non équivoque » (Van der Maren, 1996: 333). Par ailleurs, l'absence d'interaction entre le chercheur et le répondant d'un questionnaire peut aussi diminuer l'implication de ce dernier dans sa tâche et le pousser à répondre n'importe quoi (Abric, 1994b); on peut imaginer que ce sera particulièrement le cas pour des questionnaires anonymes.

En bref, le questionnaire demeure un outil central dans la recherche sur les représentations et constitue le moyen le plus rentable d'y introduire une dimension quantitative. L'impact de la plupart des limites présentées plus haut pourra être réduit à l'étape de conception du questionnaire, par exemple en faisant valider la formulation des questions par des experts et des individus typiques de la population ciblée, en introduisant des questions qui se recourent de façon à pouvoir vérifier la stabilité des réponses fournies, etc. (Van der Maren, 1996) Néanmoins, l'utilisation de sources de données complémentaires demeure le meilleur moyen de valider des informations recueillies par questionnaire (voir 1.2); nous avons d'ailleurs rencontré très peu de recherches se limitant à la seule utilisation de questionnaires, du moins en didactique.

1.1.3 LES MÉTHODES ASSOCIATIVES

Les méthodes associatives semblent beaucoup moins répandues que les méthodes interrogatives dans les travaux sur les représentations en didactique; par contre, elles sont très présentes dans la littérature traitant des méthodologies de recueil des représentations sociales. Il existe plusieurs formats de méthodes associatives, mais nous ne présenterons que la principale, qui permettra de saisir l'essence de ce type de méthodologie : l'association libre.

L'association libre vise à mettre au jour les représentations d'un individu en lui permettant de produire spontanément, en dehors des contraintes imposées par

le langage dans un contexte de production normal (structure syntaxique, cohérence, etc.), les mots qui constituent pour lui l'univers sémantique de l'objet étudié (Abric, 1994b). Le principe de base de cette méthode est de fournir un terme inducteur au sujet et de lui demander d'écrire, sans réfléchir et assez rapidement, les mots qui lui viennent en tête; le même procédé peut être repris avec différents termes inducteurs liés à l'objet étudié. Abric (1994b: 66) souligne que cette technique est « majeure pour recueillir les éléments constitutifs du contenu [d'une] représentation » et mentionne qu'elle est utilisée dans de très nombreuses recherches. Bien entendu, cette démarche produit un ensemble pêle-mêle de mots, « difficilement interprétable *a priori* » (p.66), qui demande ensuite à être organisé, dans le but de dégager les types de liens sur lesquels sont fondées les associations et de saisir les éléments organisateurs des contenus révélés. Cette étape d'organisation peut être guidée par un entretien avec le sujet ayant produit les associations, mais des variantes de l'association libre peuvent produire des données plus organisées. Nous pensons par exemple à la *carte associative* d'Abric (1994b), inspirée par Jaoui (1979), ainsi qu'au *réseau d'associations* de De Rosa (2003).

Si les approches associatives sont plutôt typiques des travaux sur les représentations sociales menés en psychologie, nous croyons qu'elles pourraient permettre d'offrir un éclairage différent dans le cadre de travaux sur les représentations en didactique.

1.1.4 LES MÉTHODES D'OBSERVATION

Les représentations et les attitudes se révélant à travers les comportements, une autre méthode pour y avoir accès est précisément d'observer ces comportements. Van der Maren (1996) distingue deux types d'observations : l'*observation participante* et la *participation observante*. La première est typique des démarches anthropologiques et ethnologiques (Lessard-Hébert *et coll.*, 1996); le chercheur s'immerge complètement dans la situation qu'il désire observer en y prenant part. Cette technique permet un regard de l'intérieur qui augmente la pertinence et la finesse de l'observation; par contre, elle demande

une grande implication – qui se traduit souvent par un investissement de temps important – et réduit l’objectivité du chercheur, qui est à la fois acteur, sujet et observateur (Van der Maren, 1999). Le cas de la participation observante, quant à lui, est différent : le chercheur s’en tient à un rôle d’observateur en étudiant les pratiques et les comportements des acteurs de la situation qui l’intéresse. Cette moins grande implication augmente l’objectivité du chercheur, qui ne prend pas directement part à la situation observée. Par contre, son statut d’« intrus » peut avoir comme effet de modifier les comportements des acteurs de la situation, ce qui diminue la pertinence des observations (Van der Maren, 1996); bien sûr, cet effet négatif diminue avec le temps, puisque le chercheur s’intègre de plus en plus au milieu qu’il observe.

La subjectivité est inhérente à l’activité d’observation. Même dans le cas de la participation observante, le chercheur « déforme » la situation observée par le simple fait de sélectionner les éléments qui lui semblent intéressants, d’en ignorer d’autres, etc. *L’observation systématique*, qui utilise une grille d’observation comme interface entre la situation et le chercheur, permet de donner davantage d’objectivité (donc de crédibilité) à l’observation, notamment en augmentant la stabilité des dimensions prises en compte. Par contre, la sélection des catégories constituant la grille implique elle-même des choix qui relèvent d’une certaine subjectivité. Par ailleurs, l’utilisation d’une grille limite en quelque sorte l’observation puisque « l’observateur ne peut voir que les faits prévus par la grille, et [que] tous les faits qu’il observe ne peuvent être rapportés que dans les termes de la grille. »¹⁷⁷ (Van der Maren, 1996: 297)

La recherche sur les représentations peut bénéficier énormément de l’observation, puisque les pratiques et les comportements peuvent en être très révélateurs. Il nous semble par contre que l’utilisation de grilles fermées est peut-être moins susceptible de permettre la mise au jour des représentations, puisque celles-ci ne peuvent pas toujours être prévues par le chercheur. La

¹⁷⁷ Il faut tout de même mentionner qu’il existe des grilles ouvertes qui admettent des modifications en cours d’observation.

recherche sur les représentations en didactique a parfois recours à l'observation, la plupart du temps en conjonction avec d'autres méthodes de recueil (entretien ou questionnaire); nous pouvons par exemple mentionner le cas de Dreyfus (2004), qui a éclairé les informations recueillies lors d'entretiens sur les représentations des enseignants sur l'enseignement du lexique au primaire par l'observation de séances de classe portant sur le lexique, ou de Blaser (2007), qui, dans la cadre de sa thèse de doctorat sur les pratiques et conceptions d'enseignants de sciences et d'histoire relativement à la fonction épistémique de l'écrit¹⁷⁸, a complété des données quantitatives obtenues à l'aide de questionnaires par des séquences d'enseignement filmées dans des classes du secondaire.

1.1.5 LE RECOURS AUX DONNÉES INVOQUÉES

Une dernière source possible d'informations pour l'étude des représentations est le recours à des données déjà existantes, que Van der Maren (1996) appelle *données invoquées*. Ces données peuvent être de tous types : production écrite, examen déjà complété, journal de bord, portfolio, etc. Évidemment, la prise en compte de tels documents demande que l'on sache exactement dans quel contexte ils ont été produits pour s'assurer de la valeur des observations qui en seront tirées.

Certaines des recherches consultées ont eu recours à ce type de documents. Par exemple, dans une recherche traitant des représentations sur les pratiques multidisciplinaires, Le Manchec et Barré-de Miniac (2004) ont étudié des dossiers relatant des pratiques de classe mises en œuvre durant leur stage par des étudiants en formation initiale. Dans une étude portant sur l'évolution des représentations sur la langue et son enseignement, Lefrançois et Montesinos-Gelet (2004) ont contrasté les autoévaluations de futurs enseignants sur leur compétence en français écrit avec les résultats obtenus par ces derniers à deux cours sur le français écrit, données extérieures à leur recherche.

¹⁷⁸ La fonction épistémique de l'écrit est « le rôle que joue l'écrit dans l'appropriation et la construction de connaissance et d'habiletés disciplinaires. » (Chartrand *et coll.*, 2006: 277)

1.2 IMPORTANCE DE LA TRIANGULATION

Ce survol des différentes méthodes de recueil des représentations nous a permis de dégager les avantages et les limites associés à chacune. Compte tenu du fait qu'il n'existe aucune méthode parfaite, et vu la complexité de l'objet « représentation », l'importance de la triangulation, lors du recueil des données tout comme dans l'analyse, semble généralement admise par tous les chercheurs s'intéressant aux représentations (Abric, 1994b ; Van der Maren, 1996 ; Vermersch, 2006). La triangulation consiste à avoir recours à plusieurs méthodologies pour étudier un objet et permet de combiner les avantages des différentes approches utilisées et de réduire l'impact de leurs limites (Apostolidis, 2003). Il existe différents types de triangulation; Apostolidis (2003: 15) en mentionne cinq, dont trois sont plus spécifiquement applicables à l'étape de recueil des données, les deux autres relevant plutôt de l'analyse :

1. *triangulation des données* : utiliser différentes sources de données
2. *triangulation du chercheur* : recours à plusieurs chercheurs pour la cueillette et l'analyse des données (cf. accord interjuges)
3. *triangulation méthodologique* : recours à différentes méthodes et techniques pour étudier un phénomène

La triangulation s'avère très présente dans les recherches que nous avons consultées; en effet, très rares sont les travaux qui n'ont recours qu'à une seule source de données. Selon Abric (1994b), toute recherche s'intéressant aux représentations sociales peut difficilement faire l'économie d'une méthodologie impliquant des entretiens; ceux-ci devraient être accompagnés d'au moins une méthode associative.

CONCLUSION

En conclusion, les méthodes interrogatives constituent les méthodes les plus pertinentes d'accès aux représentations. Par contre, il ne faut pas oublier que les données recueillies par ces méthodes ne sont pas des représentations en soi, mais bien des indices qui peuvent permettre de les inférer (Abric, 1994b ; Wirthner, 2008); les phases de traitement et d'analyse des données seront donc

cruciales dans l'élucidation des représentations. Parce qu'elles sont spontanées et peut-être plus susceptibles de faire ressortir des éléments latents, les méthodes associatives présentent aussi un intérêt certain, même si elles semblent moins répandues dans les travaux menés en didactique. Finalement, la triangulation devrait faire partie intégrante de toute méthodologie de recherche s'intéressant aux représentations, par exemple par l'observation des pratiques des différents acteurs de la situation étudiée et le recours à des données invoquées. En combinant ces différentes méthodologies pour tirer le meilleur de chacune, le chercheur en éducation apportera validité et pertinence à ses travaux.

2. CONSTITUTION D'UN CORPUS POUR LA DESCRIPTION DU RAPPORT À L'ERREUR LEXICALE D'ENSEIGNANTS

Comme nous venons de le voir, aucune recherche s'intéressant aux représentations d'un sujet à propos d'un objet – ou plus largement, au rapport d'un sujet avec un objet – ne peut faire l'économie des méthodes interrogatives. Comme ce deuxième volet de notre recherche se veut très exploratoire et vise à défricher une question pour laquelle nous disposons de très peu d'informations, nous avons choisi de nous concentrer sur un nombre restreint de sujets et d'opter pour une perspective résolument qualitative; l'entrevue nous est donc apparue comme le choix méthodologique le plus pertinent et a constitué notre source principale de données.

Notre deuxième objectif n'avait aucune visée quantitative. Par conséquent, nous avons rapidement écarté le questionnaire comme méthode d'accès au rapport à l'erreur lexicale des enseignants, puisqu'il nous semblait plus à propos de retenir une méthodologie permettant une certaine flexibilité, de façon à laisser émerger leur rapport à l'erreur au gré d'échanges et à pouvoir approfondir certains points avec eux. Nous n'avons pas non plus jugé les méthodes associatives comme très pertinentes au regard de notre objet de recherche, d'autant plus que nous disposons de peu d'appareillage pour exploiter cette méthode d'accès aux représentations dans une perspective didactique. En ce qui concerne les

méthodes d'observation, nous avons une fois encore jugé qu'elles n'étaient pas appropriées pour notre recherche, puisque nous anticipions que le rapport des enseignants à l'erreur lexicale se manifesterait principalement à travers la tâche de correction, tâche résolument solitaire, et donc difficilement observable. En effet, si nous avons demandé à des enseignants de les observer durant la correction de productions écrites, nous aurions dû leur demander aussi de verbaliser à voix haute leurs réflexions et leurs pratiques, ce qui nous aurait conduit à réaliser plutôt un entretien d'explicitation qu'une réelle observation.

2.1 LA TRIANGULATION DE NOS DONNÉES D'ANALYSE

L'entretien a donc constitué notre principal moyen d'accès aux représentations. Dans le but d'assurer une certaine triangulation des données qui ont servi à notre analyse, nous avons tout de même diversifié les sources de données ainsi que les tâches demandées aux enseignants interviewés.

Durant les entretiens semi-directifs, en plus des questions visant plus précisément la mise en lumière des aspects conceptuels de leur rapport à l'erreur lexicale, nous avons fait verbaliser les enseignants à propos de leurs pratiques de classe relatives au vocabulaire, à la correction de textes et à l'erreur lexicale; sans constituer des observations à proprement parler, leurs réponses nous ont donné un certain accès à leurs pratiques¹⁷⁹. Certains sujets ont même appuyé certaines de leurs déclarations à l'aide de rédactions d'élèves qu'ils avaient corrigées ou en nous montrant du matériel utilisé en classe pour travailler le vocabulaire.

Nous avons aussi fait verbaliser les enseignants sur sept erreurs lexicales précises issues de notre corpus d'erreurs dans le but de voir comment ils réfléchissent devant de tels problèmes. Sans constituer un entretien d'explicitation à proprement parler, puisqu'il ne vise pas l'explicitation d'une tâche réellement réalisée, ce volet de l'entretien s'en rapprochait tout de même,

¹⁷⁹ Sans pouvoir affirmer qu'il s'agit ici des pratiques réelles, ces pratiques déclarées contribuent à la mise au jour du rapport des enseignants à l'enseignement du vocabulaire et à l'erreur lexicale.

puisqu'il visait l'explicitation d'une tâche hypothétique, mais bien concrète, très similaire à celle que vivent quotidiennement les enseignants : réfléchir à propos d'erreurs de vocabulaire d'élèves de troisième secondaire et prendre des décisions relativement à ces erreurs en lien avec l'évaluation de l'écrit (pénaliser ou non, classer à un endroit précis dans la grille de correction, expliquer le problème à l'élève).

Nous avons aussi eu recours à des données invoquées en utilisant les rédactions corrigées du corpus de textes de notre premier objectif de recherche – 224 sur les 300 productions recueillies – pour observer quelles erreurs lexicales les enseignants avaient relevées, à quels types d'annotations elles avaient donné lieu et de quelle façon les enseignants en avaient tenu compte dans l'évaluation. Ces corrections, fruits du travail d'enseignants de troisième secondaire différents de ceux que nous avons rencontrés, nous ont donc permis de comparer le contenu de nos entretiens à ce que nous retrouvons sur des rédactions réelles d'élèves corrigées par des enseignants avant qu'ils ne sachent que ces productions feraient partie de notre corpus de recherche, donc dans un contexte naturel de correction.

Évidemment, l'effort de triangulation que nous avons mis de l'avant présente certaines limites. Tout d'abord, les pratiques auxquelles nous avons eu accès demeurent les pratiques déclarées par les enseignants et nous ne sommes pas allé en classe (faute de temps et de moyens !) pour vérifier qu'elles sont effectivement mises en œuvre. Par ailleurs, bien que les erreurs lexicales que nous avons soumises aux enseignants provenaient de textes d'élèves, certains sujets nous ont admis qu'ils n'avaient pas toujours le temps de se poser autant de questions à propos d'erreurs de cet ordre et qu'ils n'auraient sans doute pas relevé certaines de ces erreurs; ainsi, leurs verbalisations à propos de ces cas précis, si elles demeurent authentiques, ne sont peut-être pas représentatives des réflexions qu'ils auraient eues réellement en les rencontrant lors d'une tâche de correction. En ce qui concerne l'analyse des rédactions corrigées par les enseignants, comme il s'agissait principalement de productions écrites de fin

d'année, elles ne donnent sans doute qu'une idée partielle de ce que l'on retrouve sur les copies des élèves en cours d'année, alors que ces derniers peuvent bénéficier des commentaires écrits de leur enseignant, qui, en conséquence, produit probablement plus d'annotations, ou des commentaires plus approfondis.

Compte tenu du caractère exploratoire de ce deuxième volet de notre travail, nous jugeons que ces limites sont pourtant acceptables et que nos choix méthodologiques permettent une triangulation satisfaisante en regard de notre objectif résolument exploratoire.

2.2 LE RECRUTEMENT DES SUJETS

Les enseignants que nous avons interviewés afin de décrire leur rapport à l'erreur lexicale sont des enseignants de français du secondaire. Comme ce deuxième volet de notre recherche s'intéresse à une question assez pointue à propos de laquelle nous disposons de très peu d'informations, nous avons choisi de limiter notre échantillon à huit enseignants, mais d'opter pour des entrevues assez longues qui nous ont permis d'explorer la question en profondeur et d'observer certaines tendances ou certaines divergences dans le rapport qu'ont les enseignants à l'erreur lexicale.

Comme nous voulions avoir accès aux représentations des enseignants concernant l'erreur lexicale et à leurs pratiques y touchant, il a fallu s'assurer que nos sujets aient été confrontés à ce type de problème assez souvent pour avoir eu l'occasion de se forger un rapport à l'erreur lexicale. Pour cette raison, nous avons retenu pour nos entrevues des enseignants ayant au moins trois ans d'expérience d'enseignement en classe de français. Pour assurer une certaine cohésion avec le corpus étudié dans le cadre de notre premier objectif de recherche, nous avons choisi des enseignants qui travaillent principalement avec des élèves ayant le français pour langue maternelle et qui enseignent en troisième secondaire.

Nous sommes entré en contact avec les enseignants par l'entremise de conseillères pédagogiques et à travers les directions d'écoles. Encore une fois, un court descriptif du projet de recherche a été transmis aux participants potentiels, toujours en spécifiant que la thèse s'intéressait aux problèmes de français écrit en général plutôt que spécifiquement à l'erreur lexicale ; nous avons ainsi évité de pousser les enseignants à amorcer une réflexion ciblée sur l'erreur lexicale avant l'entretien, de façon à contourner le piège contre lequel nous met en garde Charlot (2003), à savoir d'induire chez le sujet des représentations qui n'auraient pas existé comme un contenu explicite de conscience sans l'intervention du chercheur.

Les mêmes considérations éthiques que celles liées au premier objectif touchent le second, c'est-à-dire que tous les enseignants participant à l'étude ont dû signer un formulaire de consentement (voir annexe II). La confidentialité des participants a été respectée tant au moment de l'analyse des données que dans la rédaction de la thèse, c'est pourquoi les enseignants sont désignés par des numéros dans notre chapitre de résultats.

2.2 ÉCHANTILLON

Les huit sujets qui ont accepté de participer à notre étude sont donc des enseignants de français de troisième secondaire travaillant dans des écoles publiques de la région montréalaise (Centre, Ouest de l'île, Couronne nord). Ils possèdent de 3 à 20 années d'expérience comme enseignant au secondaire; la moyenne est de 11,7 années d'expérience (dont 6,5 en troisième secondaire). Sept des huit sujets sont des femmes.

2.3 LES ENTREVUES

Comme l'entretien semi-directif semble le plus approprié pour la mise au jour des représentations (Lüdi et Py, 2002), c'est celui que nous avons retenu. Ainsi, nous avons, préalablement aux entrevues, déterminé certains thèmes à aborder et construit une liste de questions, mais l'ordre dans lequel les échanges se sont

déroulés variait en fonction des réponses de l'enseignant et de son désir d'approfondir tel ou tel aspect des questions soulevées.

Chacun des huit enseignants participant à la recherche a été rencontré, une seule fois, entre les mois de mars et de mai 2010, lors d'une entrevue d'en moyenne quarante-cinq minutes, la plus courte étant de 30 minutes et la plus longue d'une heure et six minutes. Les entretiens, que nous avons menés personnellement, ont été enregistrés en format audio aux fins de l'analyse.

Les entrevues avec les enseignants comportaient deux parties. Le cœur de l'entrevue prenait la forme d'un échange entre le chercheur et l'enseignant, échange guidé par une liste de thèmes à aborder et de questions préparées (voir 2.3.1). À la suite de cette discussion, nous avons soumis sept erreurs lexicales issues de notre corpus du premier objectif, de façon à les entendre verbaliser sur des erreurs spécifiques (voir 2.3.2). Voyons plus en détail chacun des deux volets.

2.3.1 LE PREMIER VOLET DE L'ENTRETIEN

Dans le but de préparer les entretiens, nous avons rédigé une liste de questions, en tenant compte des trois dimensions du rapport à l'erreur lexicale des enseignants présentées au chapitre VI-3. Ces questions ont fait émerger quatre thèmes principaux, qui ont structuré nos entrevues : *i*) la maîtrise de la langue et du vocabulaire, *ii*) l'enseignement du vocabulaire, *iii*) la correction de productions écrites et l'erreur de vocabulaire, *iv*) le dictionnaire en classe de français. Le tableau de la page suivante contient les questions que nous avons rédigées ; la lettre entre parenthèses devant chacune des questions permet d'identifier la dimension du rapport à l'erreur lexicale à laquelle elle est liée (A = axiologique, C = conceptuelle, P = praxéologique) ; vous trouverez dans annexe 5 le document que nous avons utilisé lors des entrevues.

Tableau 22. Questions posées aux enseignants lors du premier volet de l'entrevue

<p>I. LA MAITRISE DE LA LANGUE ET DU VOCABULAIRE</p> <p>(C) Selon vous, que signifie « maîtriser une langue »?</p> <p>(A) Selon vous, est-il important d'avoir un vocabulaire étendu? Pourquoi?</p> <p>(C) Qu'est-ce que « quelqu'un qui a du vocabulaire »?</p> <p>(C) Comment décririez-vous le vocabulaire de vos élèves? Percevez-vous de grandes différences entre vos élèves relativement à leur vocabulaire? Comment vous expliquez-vous ces différences?</p> <p>(C) Lors de l'épreuve unique de la fin du secondaire, les élèves québécois obtiennent année après année de très bons résultats pour le critère vocabulaire. Ces résultats sont-ils selon vous représentatifs de leur niveau de maîtrise du vocabulaire? Si non, qu'est-ce qui explique de tels résultats?</p>
<p>II. L'ENSEIGNEMENT DU VOCABULAIRE</p> <p>(C) Quel rôle l'école a-t-elle à jouer dans l'apprentissage du vocabulaire?</p> <p>(P) Enseignez-vous le vocabulaire en classe? De quelle façon?</p> <p>(P) Quels sont les aspects du vocabulaire sur lesquels porte votre enseignement?</p> <p>(P) À quelle fréquence enseignez-vous le vocabulaire?</p>
<p>III. LA CORRECTION DE PRODUCTIONS ÉCRITES ET L'ERREUR DE VOCABULAIRE</p> <p>(P) Quel type de grille utilisez-vous lors de la correction des productions écrites de vos élèves?</p> <p>(P) Comment tenez-vous compte de la dimension lexicale des textes de vos élèves dans l'évaluation de leurs productions écrites? Enlevez-vous des points pour les erreurs de vocabulaire?</p> <p>(C) La plupart des grilles de correction proposent une évaluation globale du vocabulaire qui repose sur les critères de richesse, de variété et de précision du vocabulaire. Selon vous, que signifie « avoir un vocabulaire riche, varié et précis »? Quels sont les indices qui vous permettent de poser un jugement sur le vocabulaire utilisé par vos élèves? En d'autres mots, comment se manifestent la richesse, la variété et la précision du vocabulaire?</p> <p>(P) Trouvez-vous facile d'évaluer le vocabulaire dans une production écrite?</p> <p>(P) Lorsque vous corrigez des productions écrites, relevez-vous aussi les erreurs de vocabulaire? De quelle façon (correction/suggestion, commentaire, renvoi au dictionnaire, soulignement, etc.)?</p> <p>(C) Croyez-vous que cela aide vos élèves?</p> <p>(P) Acceptez-vous l'orthographe rectifiée lorsque vous corrigez un texte?</p> <p>(P) Faites-vous un suivi lorsque vous avez remis un texte corrigé aux élèves? Doivent-ils retravailler leur texte? Si oui, demandez-vous aux élèves de corriger les problèmes de vocabulaire présents dans leurs rédactions?</p> <p>(P) En classe, revenez-vous sur les problèmes de vocabulaire constatés dans les rédactions pour faire des activités lexicales ou des « capsules linguistiques »?</p>
<p>IV. LE DICTIONNAIRE EN CLASSE DE FRANÇAIS</p> <p>(P) Quel dictionnaire utilisez-vous lorsque vous corrigez? Avez-vous un dictionnaire favori?</p> <p>(P) Quels sont les dictionnaires disponibles dans votre classe?</p> <p>(P) Dans quels cas vos élèves utilisent-ils le dictionnaire? Pouvez-vous décrire leur façon de s'y référer? Selon vous, tirent-ils profit de leur utilisation du dictionnaire?</p> <p>(P) On parle beaucoup d'anglicismes à éviter au Québec. Êtes-vous sensible à l'emploi des anglicismes en rédaction? À quoi vous référez-vous pour établir qu'un mot est un anglicisme?</p>

Bien entendu, comme les entretiens étaient de type semi-directif, les questions n'ont pas toujours été abordées dans l'ordre ni nécessairement formulées exactement comme elles apparaissent dans le tableau. Par contre, un ordre précis a tout de même été respecté dans certains cas, afin de ne pas influencer les réponses des enseignants. Par exemple, pour le premier thème, nous nous sommes assuré de faire émerger d'abord les conceptions relatives à la maîtrise de la langue avant de mentionner que notre recherche s'intéresse plus précisément au lexique, ceci dans le but d'éviter que les sujets ajustent leurs réponses en fonction de leur interlocuteur.

2.3.2 LE SECOND VOLET DE L'ENTRETIEN

Dans la deuxième partie des entretiens, nous avons voulu faire verbaliser les enseignants à propos d'erreurs lexicales spécifiques. Nous avons donc choisi sept erreurs dans notre corpus du premier objectif; ce modeste échantillon d'erreurs se voulait représentatif des problèmes fréquents des élèves, mais aussi diversifié quant aux types de problèmes qu'il contient. Nous avons demandé aux enseignants de nous expliquer pourquoi ils auraient ou non relevé ces erreurs, de quelle façon ils en auraient tenu compte dans leur évaluation et à quel endroit ils les auraient classées dans leur grille de correction. Nous leur avons aussi demandé le type de rétroaction qu'ils fourniraient à un élève qui aurait commis une erreur semblable.

3. LE TRAITEMENT ET L'ANALYSE DES DONNÉES

Les entretiens ont été transcrits intégralement, avant d'être codés à l'aide du logiciel d'analyse qualitative *QDA Miner*. Ce logiciel nous a permis de faire un codage par thèmes pour ensuite regrouper facilement l'ensemble des commentaires des sujets portant sur un point précis (l'orthographe rectifiée par exemple); il permet aussi d'effectuer un double codage sur un extrait, ce qui s'est avéré très utile pour nous, puisqu'il arrivait fréquemment qu'un même commentaire de nos sujets véhicule certaines informations relatives à ses pratiques, mais révèle en même temps certaines de ses conceptions.

Nous avons utilisé une grille de codage ouverte, c'est-à-dire que les rubriques qui la constituent ont été déterminées au fur et à mesure de l'analyse (Van der Maren, 1996) en fonction des réponses fournies par les enseignants ; nous avons commencé par des rubriques assez larges, correspondant grosso modo aux quatre grands thèmes de notre questionnaire d'entrevue, que nous avons raffiné en cours d'analyse¹⁸⁰. La liste finale des rubriques utilisées apparaît dans l'annexe 6.

Une fois les commentaires codés, nous avons pu, en regroupant l'ensemble des commentaires des sujets relatifs à chaque thème, faire une analyse thématique de contenu (L'Écuyer, 1990), dont les principaux résultats seront présentés au chapitre X.

4. LES LIMITES ASSOCIÉES AU DEUXIÈME OBJECTIF

Il importe de souligner une fois encore le caractère exploratoire de notre deuxième objectif de recherche. Le choix d'un échantillon restreint entraîne plusieurs limites liées au potentiel de généralisation des observations qu'a permis l'analyse de nos entretiens.

En effet, les enseignants qui ont participé à nos entretiens représentent une sous-classe bien spécifique d'enseignants, soit ceux oeuvrant dans un milieu socioéconomique moyen et auprès d'élèves de troisième secondaire essentiellement francophones; ceci n'est bien sûr pas représentatif de l'ensemble des enseignants de français du Québec. Par exemple, nous pourrions poser l'hypothèse que les enseignants de français langue seconde, ou ceux oeuvrant dans des milieux à forte proportion d'élèves allophones, par la nature même de leur tâche, fournissent un traitement plus approfondi et systématique des problèmes lexicaux de leurs élèves.

¹⁸⁰ Le logiciel *QDA Miner* possède d'ailleurs une fonction qui permet de scinder aisément un ensemble de commentaires ayant reçu un même code en sous-ensembles plus précis; ainsi, il est possible de raffiner son codage en cours d'analyse sans avoir à tout revoir depuis le début lors de l'ajout de codes nouveaux.

Comme nous l'avons souligné plus haut, la triangulation des données proposée ici n'est pas tout à fait optimale, notamment en ce qui a trait à la description des pratiques d'enseignement du vocabulaire, puisque cet aspect de nos entrevues repose exclusivement sur des pratiques déclarées. Par contre, en ce qui a trait au traitement de l'erreur lexicale, thème central de ce second volet de notre thèse, le recours à des données invoquées (rédactions déjà corrigées du corpus de l'objectif 1) a permis de constater que les pratiques déclarées de nos sujets correspondaient à ce que nous avons pu observer sur les rédactions déjà corrigées; en d'autres mots, les pratiques déclarées de nos sujets pour l'objectif 2 semblent fidèles aux pratiques réelles des enseignants qui ont corrigé les textes de l'objectif 1.

DEUXIÈME OBJECTIF - RÉSULTATS

X. LES ENSEIGNANTS DE FRANÇAIS DU SECONDAIRE ET L'ERREUR LEXICALE

Le chapitre précédent nous a permis de dresser un portrait exhaustif des erreurs lexicales des élèves de troisième secondaire; celui qui suit nous permettra d'aborder la question de l'erreur lexicale en nous intéressant maintenant aux enseignants. Rappelons d'abord notre second objectif de recherche :

2. Décrire le rapport d'enseignants de français du secondaire à l'erreur lexicale en ce qui concerne

- ***leur conception de ce qu'est une erreur de nature lexicale;***
- ***leurs pratiques d'évaluation du vocabulaire en production écrite;***
- ***leurs modes de rétroaction par rapport à des erreurs lexicales.***

Cette description du rapport à l'erreur lexicale d'enseignants de français du secondaire fera l'objet de la troisième partie du chapitre. Mais tout d'abord, les deux premières sections feront le pont avec les données du chapitre précédent. Nous observerons en premier lieu de quelle façon les enseignants ont traité les erreurs lexicales du corpus lors de leur correction à partir des copies corrigées de notre corpus (section 1). Ensuite, à la section 2, nous présenterons les réflexions des huit enseignants que nous avons interrogés sur sept erreurs lexicales de notre corpus qui leur ont été soumises en cours d'entretien.

Une petite remarque terminologique s'impose avant de continuer. Comme ce chapitre vise à mettre en lumière le rapport des enseignants à l'erreur lexicale et que la terminologie associée à l'étude du lexique ne leur est pas familière, nous avons choisi d'utiliser ici leur métalangage. Ainsi, le terme MOT, que nous avons soigneusement évité tout au long de cette thèse, apparaîtra fréquemment dans les pages qui suivent. Par contre, nous ne croyons pas que l'emploi de certains termes moins précis mette en péril la clarté de notre propos et, chaque fois où cela aurait pu être le cas, nous avons utilisé la terminologie adéquate.

1. LES ERREURS LEXICALES ET LA CORRECTION DES ENSEIGNANTS

Lorsque nous avons effectué notre collecte de données, 224 des 300 rédactions recueillies avaient déjà été corrigées par les enseignants. Nous avons observé que tous les textes annotés dont nous disposions étaient corrigés de la même façon : quatre colonnes à la droite de chaque page permettent aux enseignants d'inscrire le nombre d'erreurs de chaque type (1. *accord* ou *grammaire*; 2. *usage* ou *orthographe*; 3. *syntaxe*; 4. *ponctuation*) présentes à chaque ligne. C'est en convertissant ces nombres d'erreurs en cotes que les enseignants attribuent à la fin une note pour les critères de langue de la rédaction. Le vocabulaire, quant à lui, est toujours évalué qualitativement, par une appréciation globale du vocabulaire utilisé dans la rédaction (voir 3.2.2.1), ce qui explique qu'il n'y ait pas de colonne précise pour cet aspect de l'évaluation.

1.1 LES ERREURS PÉNALISÉES, RELEVÉES ET NON RELEVÉES ET LE CLASSEMENT DES ERREURS

Ces 224 copies corrigées nous ont donc permis d'observer de quelle façon les enseignants traitent les erreurs lexicales de leurs élèves lors de la correction. Ce sous-corpus de textes (224 rédactions corrigées) comportait 1954 erreurs selon notre analyse, si nous incluons celles d'orthographe et de flexion. Le Tableau 23 présente la position des enseignants relativement à ces 1954 erreurs en indiquant, pour chaque classe d'erreurs, combien ont été pénalisées, relevées seulement ou non relevées par les enseignants ayant corrigé les rédactions.

Tableau 23. Traitement de l'erreur en fonction des types d'erreurs

	Pénalisées	Relevées	Non relevées	Total
F-Orth	757	7	195	959
F-Barb	71	9	66	146
F-Flex	66	2	15	83
F-Inex	5	5	8	18
F-Déf	2	2	6	10
CM-Genr	30	3	6	39
CM-Inv	24	0	20	44
CM-Class	4	0	0	4
CM-Rés	0	0	2	2
CS-ChPr	14	1	25	40
CS-Const	10	1	11	22
CS-AbsC	10	1	6	17
CS-Trans	5	0	7	12
CS-Intrans	2	0	3	5
CS-AbsPr	2	0	2	4
CL-Inex	20	12	53	85
CL-Cont	6	0	5	11
CL-FL	0	3	5	8
CL-Dér	0	0	3	3
S-Fict	12	18	43	73
S-Proc	10	21	55	86
S-Pléo	12	2	12	26
S-Act	5	2	18	25
S-Ang	1	2	31	34
S-Syn	1	5	20	26
CP-Regis	16	38	115	169
CP-Conn	0	1	1	2
CP-Dial	0	0	1	1
Total	1085	135	734	1954

La Figure 19 indique dans quelles proportions les erreurs de chaque grande classe d'erreurs ont été pénalisées, relevées seulement ou non relevées par les enseignants ayant corrigé les rédactions.

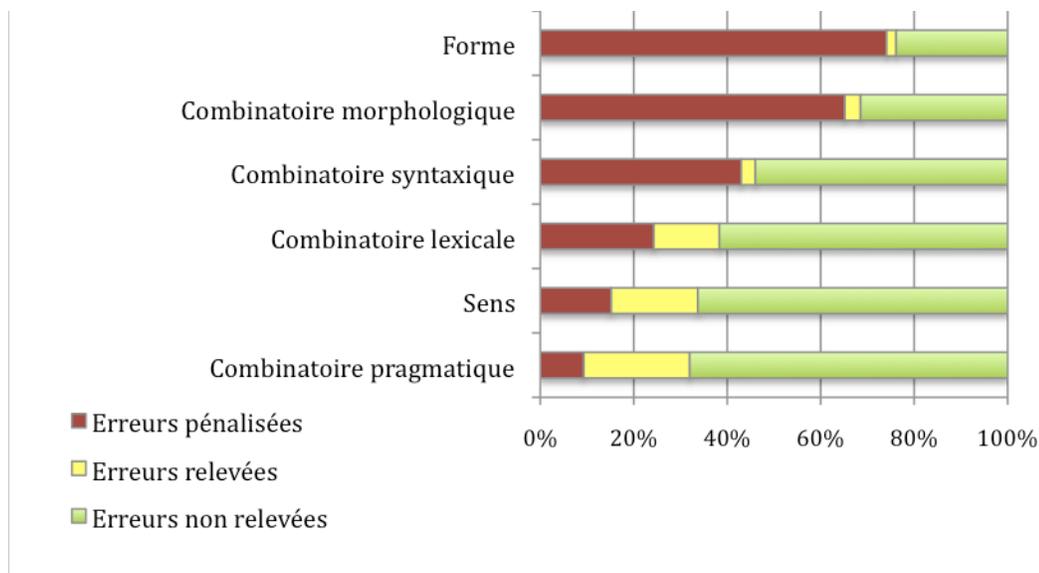


Figure 19. Prise en compte des erreurs par les enseignants dans leur correction

Ce qui frappe à première vue est le nombre important d'erreurs non relevées par les enseignants (37,6 %). Il ne faut pas s'alarmer devant cette statistique, puisque l'on se rappelle que nombre d'erreurs de notre corpus ont obtenu un score peu élevé sur notre échelle de gravité et ne méritaient par conséquent pas d'être sanctionnées par les enseignants, ni même peut-être d'être soulignées, considérant la quantité importante d'erreurs de tous types présentes dans les textes (syntaxe, accords grammaticaux, cohérence, etc.). On retrouve un certain nombre d'erreurs lexicales relevées mais non pénalisées, ce qui s'explique sans doute par le fait que le vocabulaire est généralement évalué de façon qualitative dans les grilles de correction utilisées par les enseignants; ainsi, ces erreurs relevées, même si elles n'ont pas été comptabilisées au même titre que les erreurs relevant des autres catégories mentionnées plus haut, ont donc sans doute eu un impact sur la note attribuée aux élèves pour le critère « vocabulaire ». Il faut aussi souligner que beaucoup des rédactions recueillies constituaient des examens de fin d'année scolaire et n'étaient par conséquent pas remises aux élèves; on peut donc supposer que les enseignants ont limité leurs annotations sur ces copies, d'où le faible pourcentage d'erreurs relevées mais non pénalisées. En effet, les enseignants auraient sans doute pointé

davantage d'erreurs « non comptabilisées » en cours d'année, alors que leurs élèves peuvent tirer parti de leurs annotations.

On constate que les classes d'erreurs lexicales sur lesquelles semblent se concentrer les enseignants sont les erreurs formelles (76,2 % d'erreurs relevées ou pénalisées) et celles liées aux propriétés morphosyntaxiques des UL (57,3 % d'erreurs relevées ou pénalisées).

Les erreurs d'orthographe sont les seules erreurs lexicales qui disposent d'une colonne permettant de les comptabiliser dans les grilles de correction utilisées par les enseignants; des 757 erreurs d'orthographe lexicales pénalisées par les enseignants, 72 % ont été classées dans cette colonne *orthographe d'usage*. Vingt-deux et sept dixièmes pour cent ont plutôt été considérées comme des erreurs d'orthographe grammaticale, ce que l'on peut comprendre, puisqu'elles mettent souvent en jeu des confusions entre des formes homophones (*[un] *pieux* [193], *[colègues de] *travaille* [210]), sont liées à l'accord de termes constituant une locution (**nul part* [29], **camp de jours* [66], **bien sure* [215], **Quoi qu'il en sois* [299]) ou se retrouvent dans la forme conjuguée d'un verbe (**mourrut* [30], **tombaierent* [161], **voullait* [194], **fesait* [201])¹⁸¹. Les erreurs liées à l'emploi du trait d'union (**c'est à dire* [234], **quelques unes* [259]) ou de la majuscule (**angleterre* [257], *[mon] *Curriculum Vitae* [272]) sont comptabilisées presque aussi souvent en ponctuation qu'en orthographe d'usage. Parmi les autres problèmes formels pénalisés, nous retrouvons 71 barbarismes, que les enseignants répartissent presque également entre les classes *orthographe d'usage* (**enflanbées* pour *enflammées* [116], **véiril* pour *viril* [131], **malifique* pour *maléfique* [200]) et *syntaxe*, ce deuxième classement étant principalement réservé aux locutions déformées (**se faire l'idée* pour *se faire à l'idée* [45], **couper les finances* pour *couper les vivres* [161], **comme exemple* pour *par exemple* [233]). Si quelques problèmes liés à la flexion ont été classés parmi les erreurs d'*orthographe d'usage* (**cela nous convaint* [51], **acquérie* pour *acquise* [233], **Marie preнна* pour *Marie prit* [256]), les trois quarts (76 %) se sont plutôt retrouvés

¹⁸¹ Nous avons tout de même rencontré plusieurs erreurs classées en grammaire qui constituent selon nous des problèmes purement orthographiques (ex. **interessantes* [292], **possibilité* [288], **plusieur* [209], **orteilles* [295]).

en *grammaire* (**GusGus apprenit* [169], **journals* [272], **ses parents croyurent* pour *ses parents crurent* [292]), ce qui nous apparaît beaucoup plus approprié¹⁸².

Comme nous nous y attendions, les erreurs liées aux propriétés morphologiques des UL (principalement *genre nominal* et *invariabilité*) sont comptabilisées par les enseignants parmi les erreurs de grammaire neuf fois sur dix; seuls trois problèmes liés à l'invariabilité ont plutôt été classés en orthographe d'usage. Les erreurs de régime sont quant à elles considérées comme des problèmes de syntaxe dans 95 % des cas. On peut aisément comprendre que les enseignants classent ainsi ces types d'erreurs, puisqu'elles se manifestent dans les accords et la syntaxe, mais nous verrons en analysant les entrevues s'ils font ressortir la source lexicale de ces erreurs lorsqu'ils les expliquent aux élèves; en effet, l'explication de telles erreurs demande la consultation d'un dictionnaire plutôt que d'une grammaire, au contraire des erreurs morphosyntaxiques « réelles » (ex. accord du verbe avec le sujet, accord de l'adjectif attribut, etc.).

Des 270 erreurs sémantiques présentes dans la partie du corpus corrigée par les enseignants, seulement 91 (33,7 %) ont été relevées, dont moins de la moitié ont été pénalisées (41). Étonnamment, les trois quarts de ces erreurs ont été classées parmi les problèmes syntaxiques, peut-être parce les problèmes sémantiques se manifestent souvent à travers la cooccurrence de deux UL (ex. **les vacances d'été firent enfin leur apparition* [285] [S-Proc], **le soleil [...] éblouissait ses yeux* [224] [S-Pléo]). Cette hypothèse vaut aussi pour les erreurs de collocations pénalisées dans ce sous-corpus, puisqu'elles ont aussi été classées en *syntaxe* dans presque tous les cas (24/26) (ex. **l'excursion venait à sa fin* [259], **Marie-Roxanne se retira de l'eau* [280]). Soulignons finalement que les quelques erreurs de registre pénalisées ont aussi été considérées comme des problèmes syntaxiques (ex. **à cause qu'il lui manquait de l'argent* [368], **Tout d'un coup, il commence à pleuvoir et à venté* [48]).

¹⁸² Rappelons qu'une proportion importante des erreurs de flexion sont liées aux formes du passé simple.

Le taux d'erreurs relevées ou pénalisées pour les erreurs sémantiques (33,7 %) ou liées à la combinatoire lexicale (38,2 %) ou pragmatique (32 %) sont sensiblement les mêmes et laissent une plus grande place aux erreurs relevées mais non pénalisées, sans doute parce que ces types d'erreurs font partie de l'évaluation qualitative du vocabulaire faite par les enseignants. On peut tout de même regretter l'importante proportion de problèmes sémantiques non relevés, qui dépasse les 65 %; en ne relevant pas les à-peu-près lexicaux de leurs élèves, les enseignants contribuent en quelque sorte à l'*approximatism* lexical qu'ils leur reprochent souvent. Le même commentaire pourrait s'appliquer aux problèmes de collocation, à la nuance près que l'on comprend mieux que les enseignants ne relèvent pas des problèmes concernant un phénomène linguistique dont ils ne sont peut-être pas eux-mêmes conscients (la collocation) et qu'ils auraient de la difficulté à justifier auprès de leurs élèves (voir 2.7). Pour eux, la cooccurrence relève sans doute de choix tout à fait libres, dont certains sont plus heureux que d'autres.

1.2 LE TRAITEMENT DES ERREURS ET LEUR GRAVITÉ

Nous pouvons nous demander si la gravité des erreurs a un impact sur le traitement qu'en font les enseignants lors de la correction de productions écrites. Le Tableau 24 semble démontrer que le lien entre la gravité des erreurs et le fait de les pénaliser ou non n'est pas évident, même si l'on constate que le taux d'erreurs pénalisées ou relevées est moindre pour les erreurs les moins graves (niveaux 1 et 2). Le niveau 4 montre la plus forte proportion d'erreurs pénalisées ou relevées, mais cela découle selon nous du fait que l'écrasante majorité d'erreurs d'orthographe se sont vues attribuer une cote de gravité de 4 et que ce type d'erreurs est le plus fréquemment relevé. Il est frappant de voir que parmi les erreurs les plus graves – principalement des problèmes sémantiques –, 61 % demeurent non relevées.

Tableau 24. Traitement des erreurs en fonction de la gravité

Niveau de gravité	Relevées et pénalisées (%)	Relevées et non pénalisées (%)	Non relevées (%)
1	5,6	22,2	72,2
2	20,6	10,3	69,1
3	22,7	18,3	59,0
4	69,8	2,7	27,5
5	32,8	16,4	50,9
6	16,7	22,2	61,1

Il apparaît donc que la gravité des erreurs, du moins telle que nous la concevons dans cette thèse, n'influence pas de façon très évidente le traitement qu'en font les enseignants dans la correction. C'est plutôt le type d'erreur lexicale qui semble déterminant : les erreurs lexicales pouvant aisément être classées parmi les catégories des grilles de correction utilisées par les enseignants (orthographe, syntaxe, grammaire) sont les plus susceptibles d'être relevées et pénalisées. Les erreurs plus « purement » lexicales, c'est-à-dire celles liées aux registres de langue, au sens et à la cooccurrence, sont quant à elles souvent ignorées dans la correction, sans doute parce qu'elles sont soumises à des normes moins claires, ce qui les rend plus difficiles à classer et à expliquer aux élèves.

Maintenant que nous avons observé de quelle façon les enseignants ayant corrigé les copies de notre corpus de textes traitent et classent les erreurs lexicales dans leur correction, cédon la parole à huit enseignants de français qui ont bien voulu nous rencontrer pour une entrevue sur la correction de productions écrites et le traitement de l'erreur lexicale en rédaction. Leurs propos nous permettront de voir si les observations que nous avons faites à partir des textes de notre corpus correspondent à leur façon de réfléchir sur l'erreur lexicale et de la traiter.

2. DIAGNOSTIC DE SEPT ERREURS LEXICALES PAR DES ENSEIGNANTS DE FRANÇAIS

En plus du questionnaire que nous avons utilisé pour faire émerger le rapport à l'erreur lexicale d'enseignants de français du secondaire, dont nous ferons la description au point 3, nous avons jugé qu'il serait intéressant d'entendre ces mêmes enseignants réfléchir à propos d'erreurs lexicales spécifiques. C'est pourquoi, en fin d'entrevue, nous avons soumis à chaque enseignant sept erreurs lexicales tirées de notre corpus. Nous avons retenu des erreurs de nature diverse, afin d'entendre nos sujets sur différentes classes d'erreurs lexicales.

Nous avons donc présenté sept phrases d'élèves aux enseignants, à l'intérieur desquelles un mot apparaissait en caractères gras. C'est précisément sur ce mot que les enseignants devaient poser un jugement. Ils devaient d'abord nous dire si la phrase était erronée selon eux en regard du mot mis en évidence. Si tel était le cas, ils devaient proposer une correction et mentionner si l'erreur aurait été pénalisée ou non, et de quelle façon ils l'auraient classée dans leur grille de correction. Finalement, ils étaient invités à verbaliser à propos de l'erreur pour tenter de l'expliquer ou pour suggérer une intervention à faire auprès de l'élève l'ayant commise. Voici donc un survol de leurs réactions à ces erreurs, présentées dans l'ordre des classes d'erreurs les plus fréquentes aux moins fréquentes (sémantiques, pragmatiques, lexicomorphologiques, lexicosyntaxiques, collocationnelles).

Nous tenons aussi à souligner que si certains des enseignants rencontrés paraissent être plus sensibles à la question lexicale, aucun n'a semblé avoir une compétence nettement moins marquée en cette matière. Ainsi, si certaines des réflexions présentées dans les sections qui suivent paraissent en décalage par rapport à la théorie linguistique exposée dans notre cadre théorique ou plus difficilement défendables d'un point de vue didactique, il faut garder en tête qu'elles n'émanent pas toutes d'un seul enseignant. Nous avons bien sûr relevé quelques lacunes dans la connaissance qu'ont les enseignants du fonctionnement du lexique et sommes parfois en désaccord avec leur façon de

traiter certaines erreurs lexicales, mais nous demeurons convaincu que les enseignants font au meilleur de leurs connaissances pour traiter cet aspect de la maîtrise linguistique pour lequel ils n'ont reçu que peu ou pas de formation.

2.1 UNE ERREUR DE SENS PROCHE

Aux pieds du volcan Asama trônait un village. [187]

L'emploi du verbe TRÔNER dans cette phrase nous apparaît problématique, puisque ce verbe véhicule le sens d' 'être placé bien en évidence, être comme sur un trône'; or, dans la phrase, le village se trouve au pied du volcan plutôt qu'au sommet (ce qui, au fond, n'est peut-être pas une si mauvaise chose pour ses habitants!). Cette erreur sémantique a aussi reçu la cote 4 pour sa gravité – elle n'induit pas de problème de compréhension – et nous l'avons expliquée par une connaissance approximative du verbe TRÔNER de la part de l'élève.

Les avis des enseignants sont très divergents relativement à cette erreur. Tout d'abord, la moitié des sujets n'auraient pas relevé cette erreur à première vue; dans deux cas, même après que nous avons expliqué le problème sémantique, les enseignants continuaient de trouver correct le choix du verbe TRÔNER, allant même jusqu'à dire que *ça crée un belle image* [E1]. Des quatre sujets trouvant problématique l'emploi de TRÔNER dans cette phrase, trois l'auraient pénalisé, deux parmi les erreurs de vocabulaire (V-) et un en syntaxe. Ce dernier choix de classer cette erreur parmi les problèmes de syntaxe nous semble difficilement justifiable, mais l'enseignant concerné, lorsque nous l'avons interrogé plus spécifiquement son classement, est demeuré sur sa position (*c'est une mauvais utilisation du verbe [...], à ce moment-là, je l'aurais mis comme une erreur de syntaxe* [E6]), malgré le fait que son explication de l'erreur soit purement sémantique.

Les trois enseignants qui ont tenté une explication de l'erreur ont bien fait ressortir son caractère sémantique (même celui le classant parmi les erreurs syntaxiques), en mentionnant que *trôner, ça d'air d'être en haut [...] surplomber* [E5], *c'est comme être surélevé* [E6]. Aucun des enseignants n'a proposé de correction pour cette erreur.

Nous estimons que les divergences observées dans le traitement qu'auraient fait les enseignants de cette erreur de sens proche est assez représentative de ce qui doit se passer en correction dans le cas d'erreurs sémantiques de ce type, les enseignants n'ayant pas tous la même acuité sémantique et les mêmes niveaux d'exigence par rapport à la justesse du vocabulaire employé par les élèves, sans compter que la multiplicité des erreurs de code dans les rédactions doit souvent détourner l'attention des correcteurs des considérations purement liées au sens.

2.2 UNE ERREUR LIÉE AU TYPAGE DES ACTANTS SÉMANTIQUES

Dyoll devait absolument passer. Il a trouvé un arbre très grand et l'a couper comme du beurre avec la hach magic. L'arbre s'est affalé sur le ruisseau. [8]

Nous avons présenté une seconde erreur sémantique aux enseignants. Le verbe S'AFFALER, qui signifie 'se laisser tomber lourdement', nécessite un premier actant sémantique qui correspond à un humain, comme le mentionne explicitement le logiciel *Antidote RX*. Le problème dans la phrase ci-dessus réside donc dans l'incompatibilité entre le verbe et le groupe nominal qui constitue son sujet, *l'arbre*, qui ne dénote pas un être humain. Cette erreur a reçu une cote de 5 pour la gravité et nous avons émis l'hypothèse que l'élève avait peut-être été influencé par le verbe S'AFFAISSE¹⁸³, sémantiquement et formellement lié à S'AFFALER, mais compatible avec un sujet ne dénotant pas nécessairement un humain.

Six enseignants sur les huit rencontrés pénaliseraient l'erreur présente dans cette phrase; parmi eux, deux disent qu'ils effectueraient une vérification dans le dictionnaire avant d'écrire qu'il s'agit d'une erreur. Des deux sujets qui choisiraient de ne pas pénaliser l'erreur, un justifie sa décision en disant que s'il doit lui-même consulter le dictionnaire pour vérifier s'il y a erreur ou non, il est injuste de pénaliser un élève de troisième secondaire; l'autre trouve que S'AFFALER est un beau verbe et, même s'il ne s'agit pas du meilleur choix lexical, il ne voudrait pas pénaliser un élève qui fait des efforts sur le plan du vocabulaire.

¹⁸³ Un de nos sujets a d'ailleurs posé la même hypothèse lors de l'entrevue.

Les justifications fournies par les enseignants pour expliquer cette erreur mettaient bien en lumière les contraintes sémantiques pesant sur le premier actant de S’AFFALER. En effet, cinq sujets ont mentionné explicitement que le sujet de S’AFFALER doit être un humain. Un autre s’est plutôt appuyé sur une composante de sens du verbe pour justifier qu’il n’admette qu’un sujet humain : le fait que S’AFFALER dénote un acte conscient.

C'est très cute! Ah! Ça c'est mon genre d'erreurs préféré! C'est vraiment cute! Ça, c'est sûr que je la lirais devant la classe pis que j'dirais que l'arbre a fait « frrr... » [geste imitant le fait de se laisser tomber] De lui-même, n'est-ce pas! Y'était fatigué! Ça, encore une fois, c'est du vocabulaire. J'aurais expliqué que l'arbre... Tsé, quand tu t'affales sur qqch, tu fais le choix conscient de t'affaler, parce que t'es épuisé, fatigué, blasé ou quoi que ce soit, mais l'arbre prend pas une décision, là. Y'é coupé, faque l'arbre s'est effondré sur le ruisseau ou s'est... Tu comprends? Mais c'est du vocabulaire encore. [E8]

Un autre sujet fait aussi référence au sens lorsqu’il tente une justification de l’erreur, mais son questionnement semble plutôt concerner le sens même du verbe et son régime, qui ne sont ni un ni l’autre en jeu.

AFFALER, ben ça tu vois, y'aurait fallu que je fasse la recherche de moi-même. Le mot AFFALER, c'est quoi la signification exacte? Parce que c'est embêtant. Si s'affaler, c'est comme tomber, est-ce que tu peux t'affaler... Tsé, tu t'affales, mais tu peux t'affaler sur qqch ? Tsé, y'aurait vraiment fallu que je fasse la recherche, pis [...] dépendamment de la réponse que j'aurais eue au verbe AFFALER, ben là, j'aurais pu dire si la phrase est bonne ou pas... [E6]

Trois des sujets mentionnent que cette erreur permet tout de même de créer *une belle image* [E1]; un souligne même que dans le contexte d’un texte fantastique dans lequel l’arbre serait vivant, il serait tout à fait correct d’utiliser ce verbe pour le personnifier. Finalement, un des enseignants se demande si ce verbe est un québécisme¹⁸⁴, alors qu’un autre croit qu’il s’agit peut-être d’un verbe appartenant au registre familier, ce qui correspond à la position du dictionnaire *Antidote RX*, mais pas des deux autres ouvrages retenus.

¹⁸⁴ L’enseignant n’ayant pas précisé son questionnement sur ce point, il est impossible de savoir s’il s’interrogeait sur le caractère familier de S’AFFALER ou réellement sur son appartenance au français québécois, mais nous avons l’intuition que c’est davantage le registre de langue qui était visé par ses propos.

Pour cette erreur, les enseignants s'appuient donc en majorité sur leurs intuitions de locuteurs du français et proposent une explication tout à fait pertinente. Pour ce qui est de la façon de vérifier ce type d'information dans le dictionnaire, un des sujets nous a dit avoir recours aux exemples fournis dans les dictionnaires, ce qui constitue selon nous un très bon réflexe, puisque les dictionnaires ne répertorient pas systématiquement et explicitement ces caractéristiques.

2.3 UNE ERREUR DE REGISTRE DE LANGUE

James n'était pas trop au convaincu, mais il essaya pareil. [48]

L'adjectif PAREIL utilisé comme adverbe, véhiculant le sens de 'quand même', appartient au registre familier du français québécois. Cette erreur très fréquente a reçu une cote de 3 sur notre échelle de gravité et a été expliquée par son intégration à la langue orale.

Sept enseignants sur les huit rencontrés auraient pénalisé cette erreur : tous ont suggéré de remplacer PAREIL par QUAND MÊME, TOUT DE MÊME OU MALGRÉ TOUT. Hormis l'enseignant qui jugeait la phrase acceptable, un autre a mentionné qu'il n'aurait pas relevé l'erreur par écrit sur la copie de l'élève, mais qu'elle aurait influencé son jugement global pour le critère « vocabulaire ».

Six des sept enseignants qui auraient pénalisé l'erreur ont fait référence à la langue orale ou au registre familier.

J'vais leur écrire le remplacement. Le mot de remplacement. Parce que... justement, y l'entendent. M'as y aller pareil! Tsé, faque y faut qu'on leur montre à changer ça. Mais j'ferai pas un cours pour leur dire que c't'inacceptable, pis je reviendrai pas en grand groupe. Mais individuellement, celui qui va l'utiliser, j'vais y noter. [E1]

Un seul de ces enseignants n'a pas précisé la nature de son questionnement par rapport à PAREIL, mais a dit qu'il vérifierait dans *Le Multidictionnaire* pour voir si on peut l'utiliser dans ce contexte. Malheureusement, sa recherche resterait vaine, puisque l'emploi adverbial de PAREIL au sens de 'quand même' n'apparaît pas dans ce dictionnaire.

Si la plupart des sujets comptent cette erreur parmi les erreurs de vocabulaire, un préfère considérer qu'il s'agit d'un problème syntaxique. Un autre aurait envisagé les deux possibilités, mais aurait plutôt opté pour un classement parmi les problèmes de vocabulaire.

J'mets plus d'erreurs de vocabulaire que d'erreurs de syntaxe. C'parce que j'me dis que la phrase, même si l'adverbe était changé... Il essaya quand même, il essaya tout de même, bon ok... Il essaya sans relâche, ou bon... dis ça autrement, là. C'est le vocabulaire, c'est juste le mot qui est pas correct. Parce que le reste de la phrase, y'est bon. [...] J'trouve que ça pénalise moins l'élève en vocabulaire qu'en syntaxe. Parce que le vocabulaire, j'peux mettre 20 [erreurs], sont pas comptabilisées. [E2]

Si plusieurs ont spécifié qu'ils noteraient cette erreur sur la copie de l'élève à l'aide d'une annotation *voc* ou *voc -*, un seul enseignant a dit qu'il écrirait aussi la mention *niveau de langue* pour guider l'élève dans la compréhension de son erreur.

Le traitement que font les enseignants des erreurs de registre de langue semble donc assez uniforme. Bien entendu, on peut par contre imaginer que le nombre d'erreurs de ce type varie énormément d'un enseignant à l'autre selon son « purisme » et son degré de connaissance des différents registres de langue. Le dernier commentaire du paragraphe précédent nous pousse à dire qu'il serait pertinent de réserver une annotation spéciale pour les erreurs de registres de langue, puisque des annotations vagues du type *voc* doivent souvent laisser perplexes les élèves et ne les poussent pas à trouver un terme du registre standard équivalent à celui utilisé.

2.4 UNE ERREUR LIÉE À L'INVARIABILITÉ

*Un premier job d'été peut s'avérer un très bon tremplin pour tes autres **travails** dans le futur. [39]*

Nous avons choisi de soumettre ce cas difficile aux enseignants parce que nous trouvions l'erreur de l'élève intéressante et étions curieux de voir comment ils allaient l'analyser. Ce problème a déjà été discuté à la section 2.2.4.1 du chapitre précédent; rappelons seulement que l'UL de registre standard TRAVAIL au sens

d' 'emploi' existe au singulier, mais qu'elle ne semble pouvoir être utilisée au pluriel. Cette erreur obtient une cote de 3 pour la gravité.

Sept des huit enseignants pénaliseraient cette erreur parmi les erreurs de vocabulaire; un des sujets mettrait même une erreur d'accord en plus de l'erreur de vocabulaire, puisque la forme *travails n'existe pas¹⁸⁵. Près de la moitié des sujets ont d'abord classé l'erreur en grammaire, avant de relire la phrase et de se rendre compte du problème lié au sens de TRAVAIL dans cette phrase. Tous les enseignants ont suggéré de remplacer *travails par emplois, et celui qui ne pénaliserait pas l'erreur justifie sa position en disant qu'il s'agit de vocabulaire trop poussé et que les élèves vont dire *un travail* plutôt qu'*un emploi*; l'utilisation de TRAVAIL au pluriel ne semblait donc pas poser problème à cet enseignant.

Ce problème était familier à l'un des sujets, qui l'avait rencontré alors qu'il donnait une formation en choix de carrière.

Ça, ça marche pas! Euh... On a travaillé beaucoup, justement... le cours que j'ai en choix de carrière. On a travaillé le CV, pis euh... Pis j'leur ai expliqué que tu peux « aller travailler », que tu peux « faire un travail », mais que t'as pas eu « des travaux », pis encore moins... ben, « des travaux ». Mais... Ah non ! Ça marche pas. C'est plus « des emplois ». [E7]

Les autres enseignants réfléchissaient à voix haute sur cette erreur et semblaient trouver qu'il s'agissait d'un cas compliqué.

Euh, « tes autres travaux ». Ok, mais... TRAVAIL dans le sens de 'travail'. Attends, « un travail », mais « des travaux », mais là... « tes autres travaux ». A'm'embête, celle-là. C'est de la façon que c'est... « job d'été ». Parce que dans le sens qu'y veut dire, c'est dans le sens de 'job'. Mais tsé, « un travail » dans le sens que le matin tu vas au travail, est-ce qu'« un travail » dans ce sens-là devient « travaux » au pluriel? [...] Ça, c'est le genre d'erreur que j'aurais été voir mes coéquipiers en français. [E6]

Oh!... « travail » dans le sens de 'job'... « bon tremplin pour tes autres... emplois d'été ». Emplois dans le futur. Elle est compliquée, celle-là! C'est un mauvais... Y'aurait une erreur c'est sûr... « tes autres travaux »... Attends un petit peu... « emplois » c'est plus beau, mais est-ce que, à la limite... Hey! Elle est bonne, celle-là ! Je l'aime pas! [...] C'est comme... voyons! Est-ce qu'y se met au pluriel avec un « s »? Ou ça devient « travaux »? [E8]

¹⁸⁵ En fait, il existe un vocable TRAVAIL signifiant 'dispositif utilisé pour immobiliser les animaux de grande taille afin de les soigner, de les ferrer', dont le pluriel correspond à *travails*, mais ce vocable est un homonyme du vocable TRAVAIL qui nous intéresse ici.

Plus de la moitié des enseignants ont explicitement fait référence au sens de TRAVAIL pour expliquer que, dans cette acception, ils trouvaient étrange de l'employer au pluriel, tant à la forme **travail* qu'à la forme **travaux*. Un des sujets condamne cet emploi parce que TRAVAIL au sens d' 'emploi' est, selon lui, familier. Un autre considère que l'élève ayant produit l'erreur manque énormément de vocabulaire, puisqu'il utilise aussi JOB dans la même phrase. Un des enseignants a aussi dit qu'il vérifierait dans le dictionnaire pour être certain qu'il s'agit d'une erreur, mais qu'il suggérerait tout de même à l'élève d'utiliser EMPLOI.

Nous avons trouvé très intéressant d'entendre les enseignants réfléchir à propos d'un cas compliqué et faire explicitement des liens entre le sens d'une UL et ses propriétés morphosyntaxiques. Nous trouvons aussi encourageant de constater que la plupart auraient classé cette erreur parmi les problèmes de vocabulaire.

2.5 UNE ERREUR LIÉE AU GENRE NOMINAL

Dans une immeuble très luxueuse, les deux hommes discutent tranquillement. [126]

Cette erreur découle de la méconnaissance du genre du nom IMMEUBLE et a reçu une cote de 4 sur notre échelle de gravité. Elle est sans contredit celle qui a le plus fait l'unanimité parmi les enseignants rencontrés : tous s'entendent pour dire que c'est une erreur qui découle de la méconnaissance du genre nominal d'IMMEUBLE et ne la pénaliseraient qu'une seule fois *parce que c'est la même* [E1]. Six sujets sur huit classeraient cette erreur dans la catégorie des erreurs d'orthographe grammaticale (ou « grammaire »); un seul a plutôt parlé d'orthographe d'usage, mais c'est sans doute une erreur d'inattention de sa part, puisqu'il faisait aussi référence aux accords. Un seul enseignant, après avoir bien analysé l'erreur, a émis la possibilité de classer cette erreur plutôt parmi les problèmes de vocabulaire.

Euh... probablement que ça ici, j'aurais mis une erreur de vocabulaire pis une erreur d'accord. Mais... c'est vite de même là. Non, parce que dans l'fond, si c'était « une... » Non, c'est ça... Il l'a bien accordé, mais c'est pas d'même. J'aurais été embêté, là... [E7]

Il est intéressant que l'enseignant en question avait travaillé spécifiquement les noms au genre « difficile » avec ses élèves.

Ah non! Ça non plus, ça passe pas. C'est « un immeuble », faque ça peut pas être « une immeuble luxueuse ». Pis ça, j'en ai fait à un moment donné avec les élèves. J'leur ai même fait faire une dictée là-dessus. « une avion ou un avion, une ou un escalier, un ou une ascenceur »? Tsé, des mots comme ça... qui font un petit cillement dans mes oreilles quand j'les entends, là. Comme « aréoport ». « autobus ». C'est des mots que j'pense que c'est important qu'y sachent si c'est masculin ou féminin, là. Souvent y l'savent pas, tsé, « une grande escalier », ça s'dit bien! Oui, mais, essaie d'y penser. D'habitude sont capables de l'trouver pour la plupart. J'sais qu'on a eu ben d'la misère avec OPINION, là. Même moi, j'trouve que « un opinion », ça s'dit bien, faque c'fatiquant! [E7]

Certains commentaires des enseignants laissent sous-entendre qu'ils ne font pas vraiment de différence entre l'accord en genre, qui découle d'une propriété intrinsèque de l'UL, et l'accord en nombre, voire l'accord du verbe avec le sujet.

*Ici, c'est pas le sens, mais y connaît pas son genre et son nombre pour IMMEUBLE. [E3]
L'erreur des accords, [...] c'est constant. C'est ça que j'te disais, au début. Mais... là c'est du féminin pis du masculin, mais souvent c'est le pluriel aussi. Pis les accords de verbes. [E1]*

Nous croyons que les problèmes liés au genre nominal sont particulièrement fréquents avec les noms qui commencent par une voyelle, ce qu'a corroboré un de nos sujets; cet enseignant souligne aussi que le fait que le nom IMMEUBLE finisse par un « e » explique sans doute aussi la méprise de l'élève.

Parce que pour eux, un mot, quand ça finit en « e », c'est féminin. C'est tout! [...] Surtout quand ça commence par une voyelle. Les liaisons, là. [E2]

Si les enseignants comptabilisent généralement ce type d'erreurs parmi les erreurs d'orthographe – comme nous pouvions nous y attendre –, ils sont tout de même conscients de la nature lexicale de la connaissance qui fait défaut à l'élève. Ainsi, la moitié des sujets ont affirmé que dans de tels cas, ils laissent une trace sur la copie de l'élève pour l'orienter lorsqu'il prendra connaissance de ces erreurs, que ce soit par une annotation du type *genre* ou *masculin*, par de petits cercles reliant les terminaisons des différents items variables du GN en jeu ou par une consigne du type *cherche ce mot dans le dictionnaire*.

2.6 UNE ERREUR DE PRÉPOSITION RÉGIE

L'horloger se dirigea à la caverne à grande vitesse. [88]

Cette phrase contient une erreur de préposition régie, puisque le verbe SE DIRIGER contrôle la préposition VERS, comme l'indiquent les trois dictionnaires retenus dans la recherche. Nous lui avons attribué une cote de gravité de 4 sur 6 – essentiellement parce qu'elle n'entrave pas la compréhension – et l'avons expliquée par une analogie de construction avec les verbes ALLER [à] et RETOURNER [à].

Six enseignants sur huit ont considéré l'emploi de la préposition À comme erroné dans cette phrase, alors que deux le jugeaient correct. Deux sujets auraient pénalisé l'erreur et l'auraient classée parmi les problèmes de syntaxe, classement qui correspond à ce que nous avons observé dans les rédactions corrigées de notre corpus. Parmi ceux qui n'auraient pas pénalisé l'erreur, deux ont souligné la difficulté de savoir s'il faut pénaliser ou non de telles maladresses.

J'aurais pas mis une erreur. J'aurais peut-être suggéré, j'aurais peut-être écrit, mais j'aurais pas mis une erreur. Parce que... j'comprends, là, quand même. C'est peut-être pas riche comme façon de l'écrire. C'est peut-être pas parfait, mais... C'est quand même « diriger », là. J'trouve que c'pas si mal. Mais j'viens que je l'sais pas aussi... En quelle année que t'es supposé écrire un texte comme ça ou comme ça? [E7]

Les six enseignants qui voyaient une erreur dans cette phrase ont tous proposé VERS comme correction et ont mentionné qu'ils auraient écrit cette suggestion sur la feuille de l'élève. Les explications qu'ils nous ont fournies permettent de supposer qu'ils sont conscients que le choix de la préposition dépend du verbe SE DIRIGER; un des sujets suggérerait même à ses élèves d'aller consulter *Le Multidictionnaire* pour cette erreur. Un autre des sujets, par contre, tente une analyse syntaxique et arrive à la conclusion qu'à *la caverne* n'est pas un complément indirect du verbe SE DIRIGER et s'en remet à une justification de l'erreur plutôt sémantique.

Tu m'en poses une question! Bon, c'est au niveau d'la préposition. Ça indique un lieu... c'est pas un complément indirect. J'aurais pas su quoi lui répondre. Disons que pour moi, c'est l'sens qui va là parce que c't'une direction, là... [...] on peut dire on se

dirige à la maison pourquoi on se dirigerait pas à la caverne? Si c'est à la caverne, ça indique un lieu où il habite là avec sa famille, oui, j'l'aurais accepté. Ça dépend du contexte... [E3]

En bref, si la plupart des enseignants s'entendent pour dire que cette phrase contient une erreur de syntaxe, ils ne sont pas aussi unanimes lorsque vient le temps de déterminer si cette erreur doit être comptabilisée ou non.

2.7 UNE ERREUR DE COLLOCATION

Les publicitaires portent une très mauvaise influence sur les jeunes dans la société. [88]

Cette erreur de collocation, comme nous l'avons vu dans le chapitre précédent (voir VIII-2.2.5.1), correspond à la FL Op_{r12} (*influence*) = *exercer, avoir [une ~ sur]*. Nous lui avons attribué une cote de 3 sur l'échelle de gravité et avons pour l'expliquer la seule hypothèse que l'élève qui l'a commise s'est probablement fait dire d'éviter dans ses textes les verbes AVOIR, ÊTRE, FAIRE, etc.

Les enseignants rencontrés sont unanimes quant à la correction à proposer : tous ont suggéré de remplacer PORTER par AVOIR [une influence]. Deux enseignants ont aussi mentionné qu'il serait peut-être plus élégant d'éviter la collocation et d'employer directement le verbe INFLUENCER (ex. *Les publicitaires influencent négativement les jeunes dans la société.*). Tous les sujets auraient pénalisé cette erreur sauf un, soit parmi les erreurs de vocabulaire – ce qui n'aurait pas nécessairement fait baisser la note de vocabulaire d'un cran, mais aurait influencé le jugement de l'enseignant au moment de déterminer la note pour ce critère –, soit parmi les erreurs de syntaxe.

Oui ça aussi, c'est des erreurs de syntaxe, c'est comme ça que je le... tsé tu comprends c'est pas le bon mot, mais tsé on va avoir, on va compter les erreurs de vocabulaire on dirait que je vas l'inclure, mais dans les grilles que j'ai là, je le passerais encore une fois pour une erreur de syntaxe, parce que c'est ça, c'est pas le bon verbe qui est utilisé. [E6]

Si la position des sujets relativement à cette erreur est plutôt unanime, l'explication de ce problème de collocation n'est, de l'aveu même des enseignants, pas simple.

Ça, c'est difficile à expliquer, j'suis ben d'accord avec toi. On dirait qu'on le sait comme instinctivement, parce que ça fait longtemps que... mais expliquer ça... Y'en a qui sont faciles, mais celle-là, j'avoue que... [E8]

Bonne question... comment j'aurais pu faire... J'trouve ça difficile. C'est ça que j'te disais tantôt, quand c'est pas instinctif... Pis j'trouve que ça, c'est difficile à corriger, parce que c'est pas... y'a pas vraiment de... y'a pas une règle. C'est pas comme « les souliers », c't'un « s ». C'est plus flou... Tu peux pas vraiment « porter une mauvaise influence »... C'est clair qu'y va m'poser la question! J'saurais probablement pas quoi dire. J'vas lui suggérer de l'dire autrement, mais... [E7]

Le réflexe des enseignants, devant cette erreur, semble de fournir une explication sémantique pour expliquer le caractère boiteux de la combinaison du verbe et du nom. Ainsi, quatre de nos sujets iraient lire avec l'élève la définition du verbe PORTER pour *démontrer qu'y a pas lieu de l'utiliser* [E3]. Cette technique risque bien sûr de mener à une impasse, puisque dans le cadre d'une collocation de type Oper, le collocatif est sémantiquement vide; devant le nombre important de définitions associées à PORTER, et confronté à des exemples comme *porter intérêt* ou *porter un jugement*, l'enseignant manquera sans doute d'arguments pour appuyer son intuition que *ça s'applique pas* [E6], pour expliquer que *tu peux porter qqchose, mais [...] tu vas pas porter une [influence]... tu vas pas aller lui porter sur le dos* [E4]. Deux enseignants seulement ont suggéré d'aller plutôt voir l'article d'INFLUENCE, ce qui correspond à une démarche beaucoup plus concluante; en effet, les trois dictionnaires consultés mentionnent la collocation *avoir de l'influence* dans les exemples d'utilisation d'INFLUENCE. Deux enseignants proposent quant à eux une intervention qui repose sur les intuitions linguistiques du jeune, soit de lire la phrase à voix haute pour faire réaliser le problème à l'élève.

Quand j'le lis à haute voix, je l'sais souvent y vont voir que ça s'dit pas. Mais quand y font juste le lire, y l'voient pas. [E4]

Bien qu'ils aient un peu de difficulté à expliquer ce genre d'erreur, les enseignants semblent avoir une connaissance intuitive du phénomène de la collocation. En effet, plus de la moitié des sujets ont mentionné que le problème découle de la combinaison entre le verbe PORTER et le nom INFLUENCE, et certains ont spécifié qu'il n'y pas de règle pour l'expliquer, que l'on sait instinctivement quelles combinaisons sont correctes.

... c'est des erreurs qui font souvent. Y comprennent pas qu'un but, faut « atteindre un but ». Tu peux pas « réussir un but ». Objectif... Même pas... tu peux-tu « réussir un objectif »? Tsé, c'est comme... C'est pas un bon mariage. [E2]

Nous croyons qu'il serait utile de sensibiliser davantage les enseignants à la collocation, afin qu'ils soient capables d'identifier facilement la base et le collocatif et qu'ils orientent plus naturellement les élèves vers l'article de la base dans le dictionnaire. Une meilleure connaissance de ce phénomène les aiderait aussi à exploiter le dictionnaire comme outil d'aide à l'écriture en tirant parti des exemples d'utilisation fournis.

CONCLUSION

La présente section nous a permis de constater les divergences dans le traitement qu'auraient fait les enseignants d'erreurs lexicales spécifiques. Ces divergences résident autant dans les jugements que posent les enseignants sur les erreurs que dans leur façon de les classer dans leur grille de correction et, dans une moindre mesure, dans leur manière de les expliquer.

Prenons maintenant une certaine distance par rapport à notre corpus d'erreurs pour nous intéresser plutôt au rapport à l'erreur lexicale des enseignants, tel qu'il nous a été révélé à travers les entretiens que nous avons menés.

3. LE RAPPORT À L'ERREUR LEXICALE D'ENSEIGNANTS DE FRANÇAIS DU SECONDAIRE

Nous présenterons dans les pages qui suivent différents aspects du rapport à l'erreur lexicale des enseignants tel que nous l'avons analysé dans notre corpus de huit entrevues menées auprès d'enseignants de français de 3^e secondaire.

Nous nous intéresserons d'abord à la dimension conceptuelle du rapport à l'erreur lexicale des enseignants en dressant un portrait de leurs représentations de la maîtrise de la langue et de l'erreur lexicale ainsi que de l'opinion qu'ils se font du vocabulaire de leurs élèves. Nous éclairerons ensuite la facette praxéologique de leur rapport à l'erreur en présentant leurs pratiques d'enseignement du vocabulaire avant d'aborder plus spécifiquement la question de la correction des productions écrites et du traitement de l'erreur lexicale en correction. Nous terminerons avec quelques remarques sur la place du dictionnaire en classe de français. Il est à noter que la dimension axiologique du rapport à l'erreur lexicale – c'est-à-dire la valeur que les enseignants attribuent à l'erreur lexicale ainsi qu'à la maîtrise et à l'enseignement du vocabulaire – ne fera pas en tant que telle l'objet d'une section du présent chapitre; par contre, certains résultats relatifs à cette facette du rapport à l'erreur émergeront au fil des prochaines pages. De plus, comme les facettes du rapport à un objet sont difficilement dissociables, certaines conceptions des enseignants seront mises en lumière à travers l'explication de leurs pratiques.

Rappelons d'entrée de jeu que les résultats que nous présentons ici proviennent d'entrevues semi-dirigées que nous avons menées auprès d'enseignants de troisième secondaire de mars à mai 2010. Nos sujets travaillent tous dans des écoles publiques de la région montréalaise (Centre, Ouest de l'île, Couronne Nord) et possèdent de 3 à 20 années d'expérience comme enseignant au secondaire.

À cause du caractère éminemment qualitatif de nos données et du nombre limité de sujets, nous éviterons d'employer des pourcentages et des figures dans les

pages qui suivent. Le caractère très exploratoire de notre deuxième objectif de recherche nous pousse en effet à être prudent quant au potentiel de généralisation de nos résultats. Néanmoins, les conceptions convergentes et les pratiques communes des enseignants rencontrés nous permettent de supposer que de mener des entretiens similaires auprès d'un échantillon plus large conduirait rapidement à une saturation des résultats.

3.1 LA DIMENSION CONCEPTUELLE DU RAPPORT À L'ERREUR LEXICALE

La dimension conceptuelle du rapport à l'erreur lexicale des enseignants concerne les représentations qu'entretiennent ces derniers relativement à différentes facettes de cette question. Nos entretiens ont permis de faire émerger des conceptions liées à trois objets, que nous aborderons tour à tour : la maîtrise de la langue, l'erreur lexicale et le vocabulaire des élèves.

3.1.1 UNE CONCEPTION ORTHOGRAPHIQUE DE LA MAITRISE DE LA LANGUE

Lorsqu'on leur a demandé en quoi consiste la maîtrise d'une langue, la quasi-totalité des enseignants (7/8) ont fait référence à la capacité d'écrire sans fautes, ce qui s'explique sans doute par la place prédominante qu'occupe l'enseignement de la grammaire dans les cours de français. Trois enseignants sur huit ont aussi mentionné la capacité d'un bon locuteur à transmettre son message et à transposer ses idées avec clarté et cohérence, ce qui correspond aussi au nombre d'enseignants ayant pointé le vocabulaire comme une composante essentielle de la maîtrise d'une langue : les élèves doivent utiliser *les mots justes* ^[E2] et être *capable[s]* de trouver *différents mots de vocabulaire pour s'exprimer, pas juste les mots communs qu'y vont utiliser en parlant* ^[E7]. Seuls deux enseignants ont souligné l'habileté à comprendre ce qu'on lit et ce qu'on entend comme une manifestation de la maîtrise d'une langue; l'ensemble des autres commentaires étaient orientés vers la production. Un des sujets a aussi mentionné que la maîtrise d'une langue passe par la dimension culturelle, c'est-à-dire la connaissance de son histoire et de sa littérature.

Nous constatons donc que les sujets interviewés ont une conception de la maîtrise de la langue très pragmatique, intimement liée à leur tâche d'enseignants : celle d'apprendre aux élèves à écrire sans fautes. Le vocabulaire semble occuper une place périphérique dans leur conception de la maîtrise de la langue, bien que la connaissance d'un certain vocabulaire semble en être une composante implicite permettant *d'être capable de dire c'qu'on a à dire clairement* [E1].

3.1.2 UNE CONCEPTION DE L'ERREUR LEXICALE CENTRÉE SUR LE SENS ET LA RÉPÉTITION

Lorsqu'on leur demande d'expliquer ce qu'est une erreur de vocabulaire, les enseignants font spontanément référence à deux types de problèmes : les erreurs sémantiques (*mauvais mot pour une réalité donnée* [E3], *mot qui est hors contexte* [E4], *qui veut dire d'autre chose* [E6]) et la répétition. Dans des proportions beaucoup plus faibles – soit deux enseignants sur huit –, ils mentionnent aussi l'emploi d'anglicismes, de termes familiers (ou *langage oral* [E5]) ainsi que le recours à des périphrases, c'est-à-dire *un groupe de mots plutôt que le bon mot* [E3]. Un des sujets fait aussi référence à l'appropriation du vocabulaire au type de texte :

Pis que ça soit adapté au conte. Moi, un élève qui commence un paragraphe de conte en disant « premièrement », c'est pas adapté au conte. J'veux avoir « Deux jours plus tard », etc. [E4]

La moitié des enseignants rencontrés ont mentionné la relation entre certaines erreurs lexicales et des problèmes syntaxiques (*Si y a pas utilisé le bon mot, ça va faire une erreur de syntaxe des fois dans la phrase aussi.* [E6]), notamment l'emploi des prépositions et les anglicismes; un enseignant a aussi noté l'incidence de certaines erreurs lexicales sur la cohérence du texte, entre autres dans le cas d'utilisation de marqueurs de relation erronés. Ces problèmes à cheval entre le lexique et d'autres dimensions du texte rendent plus difficile le classement des erreurs.

C'est pas facile, j'trouve. Parce que des fois on a envie d'mettre l'erreur en structure de phrase, pis là on relit comme y faut pis on s'dit : « Ben non, c'est pas la phrase qu'y est mal structurée... Qu'est-ce qu'y a, tsé? » Des fois, faut arrêter pis s'questionner. [...] On trouve ça l'fun parce qu'on s'questionne pis on aime ça... On

aime ça, s'questionner, hein, les profs? Mais euh [...] des fois, une erreur de vocabulaire, faut vraiment s'questionner à savoir où est-ce qu'on le met. [E1]

La répétition et l'erreur sémantique sont donc au cœur de la conception qu'ont les enseignants de l'erreur lexicale; de façon plus marginale, ils font référence à l'emploi d'anglicismes et de termes familiers. Leur conception de l'erreur lexicale est donc, tout comme celle qu'ils ont de la maîtrise linguistique, intimement liée aux grilles de correction qu'ils utilisent : les types d'erreurs qu'ils ont mentionnés sont ceux qu'ils classent parmi les erreurs de vocabulaire lorsqu'ils corrigent, ceux qui appellent une annotation *voc*. Néanmoins, les enseignants semblent tout à fait conscients de la source lexicale de certaines erreurs relevant d'autres dimensions (syntaxe, cohérence), mais ne considèrent pas ces dernières comme des erreurs de vocabulaire à proprement parler parce qu'ils les classent généralement ailleurs dans leur grille de correction.

Si nous comparons la conception qu'ont les enseignants de l'erreur lexicale à la définition que nous en proposons dans cette thèse, nous constatons qu'elle y correspond en grande partie, mais qu'elle couvre un spectre moins large de problèmes. En lien avec les grandes divisions de notre typologie, nous pourrions dire que les sujets interviewés considèrent les erreurs sémantiques, lexicopragmatiques et (dans une moindre mesure) lexicosyntaxiques parmi les erreurs lexicales. Comme on pouvait s'y attendre, les erreurs formelles sont traitées dans une catégorie à part. Les erreurs lexicomorphologiques, quant à elles, n'ont jamais été mentionnées comme relevant du vocabulaire et aucun enseignant n'a fait mention des problèmes de collocation dans sa définition de l'erreur lexicale.

3.1.3 UNE PERCEPTION PLUTÔT POSITIVE DU VOCABULAIRE DES ÉLÈVES

Contrairement à ce que nous attendions, un seul des enseignants interviewés trouve que ses élèves manquent de vocabulaire, ce qui explique à son avis le nombre de répétitions dans les rédactions. La plupart des autres sujets s'entendent pour dire que les élèves possèdent un vocabulaire *moyen* [E2/3/6] ou *bon* [E1/7], voire *beau* [E4], même s'ils commettent beaucoup d'erreurs d'orthographe

lexicale. Un des enseignants mentionne par contre qu'il faut que les élèves soient évalués et qu'ils aient accès à des ouvrages de référence si on veut les voir utiliser un bon vocabulaire, autrement *c'pas fort et y'a plus d'abréviations que d'autre chose* [E1]. Un autre sujet affirme que les élèves connaissent des beaux mots, mais ne les utilisent pas.

J'ai travaillé beaucoup le champ lexical avec eux autres, parce que je trouvais qu'y avaient pas beaucoup de vocabulaire, justement. Mais y'en ont, mais ils l'utilisent pas. Quand on fait les chansons, la poésie et tout ça, on part d'un thème et là, on sortait plein de mots, y me sortaient des mots que je faisais : « Wow ! Pourquoi vous les utilisez pas, ces mots-là? » Faque c'est la facilité, rapidement... « Ça me tente pas de faire ce travail-là, faque j'me dépêche. » Y'ont pas le gout de s'investir et de faire plus, je pense qu'y se rendent pas compte qu'y pourraient être fiers d'eux et feraient vraiment des beaux textes avec des beaux mots. [E7]

Bien entendu, tous les élèves n'ont pas le même niveau de vocabulaire. Si la moyenne des élèves ont un vocabulaire moyen, certains présentent de graves lacunes – environ deux ou trois élèves par groupes selon E3 –, alors que d'autres – cette fois un ou deux élèves par classe, toujours selon E3 – utilisent un vocabulaire enrichi et *prennent un plaisir fou à intégrer des nouveaux mots* [E1], allant parfois même jusqu'à en mettre trop, surtout dans le texte littéraire [E2]. Un des enseignants, qui met beaucoup l'accent sur le vocabulaire dans ses cours, note que les filles ont souvent un meilleur vocabulaire et perçoit surtout une amélioration chez ses élèves moyens.

Du début de l'année, quand j'les prends euh... purs et... c'est vraiment presque à sec, là, pis qu'on travaille, on travaille... oui, y'a d'amélioration. [...] Ben, les élèves forts sont toujours forts et les élèves faibles sont toujours faibles. Malheureusement. [...] Là où tu vas voir la grosse différence, c'est les élèves réguliers, là, moyens, qui vont être capables de s'améliorer... par acharnement, [...] par volonté. Eux, là, vont s'améliorer. Mais sinon, si tu l'pousses pas un peu... [E2]

Un point sur lequel tous les enseignants rencontrés semblent d'accord est la capacité des élèves à bien différencier les registres de langue : E3 remarque qu'ils ont un très bon vocabulaire lorsqu'ils font des présentations orales et souligne qu'ils sont bien conscients des différences entre l'oral et l'écrit.

Contrairement à y'a trois quatre ans, peut-être, où on voyait des jeunes qui écrivaient comme y parlaient. Cette année, ces deux dernières années, j'ai pas eu ça. J'ai pas eu d'abréviations Internet ou quoi que ce soit, là. On l'voyait, là je l'vois pas.

[E4]

Nous avons aussi posé des questions plus spécifiques sur certains types d'erreurs lexicales : soit la présence de « termes québécois »¹⁸⁶ ou d'anglicismes. Dans les deux cas, les enseignants ont dit en voir très peu dans les rédactions des élèves.

Évidemment, même si les enseignants pensent globalement que leurs élèves ont un bon vocabulaire, ils s'entendent pour dire qu'il y a aussi place à l'amélioration. Pour eux, l'école a bien sûr un rôle à jouer dans l'accroissement du vocabulaire des élèves, mais c'est surtout la lecture qui leur permettra d'élargir et d'approfondir leur vocabulaire. On souligne aussi l'impact du milieu familial sur le vocabulaire des élèves, notamment le rôle primordial que jouent les parents en échangeant beaucoup (ou non) avec leurs enfants. Un des sujets, qui avoue pourtant n'y consacrer que peu de temps, affirme que *l'enseignement pis l'évaluation du vocabulaire, c'est important* [E3]; il enfonce même le clou en disant que *quelqu'un qui a du vocabulaire, ça va directement avec l'intelligence*.

3.2 LA DIMENSION PRAXÉOLOGIQUE DU RAPPORT À L'ERREUR LEXICALE

La dimension praxéologique du rapport à un objet concerne les pratiques mises en œuvre en lien avec cet objet. Dans le cadre de cette thèse, nous avons choisi de dresser un portrait des pratiques d'enseignement du vocabulaire des enseignants, de décrire leur façon d'évaluer le vocabulaire en production écrite et de traiter l'erreur lexicale lors de la correction et de définir la place qu'occupe le dictionnaire en classe de français au secondaire.

3.2.1 LES PRATIQUES D'ENSEIGNEMENT DU VOCABULAIRE

Les enseignants rencontrés admettent ne travailler que très peu le vocabulaire en classe, et toujours *par la bande* [E5], *à travers autre chose* [E3]. Le travail fait sur le

¹⁸⁶ Par cette formulation, nous voulions voir si les enseignants assimilaient les termes québécois à des termes familiers, ce qui n'est pas vraiment ressorti des réponses que nous avons obtenues. En effet, les enseignants semblaient nous dire qu'ils relèvent une erreur parce qu'un mot appartient au registre familier ou à la langue orale (et non pas parce que c'est un terme québécois), comme en témoigne l'extrait suivant : « *Tsé, des québécoismes, ça dépend quel niveau de langage aussi. Y'en a que je vais passer par-dessus, mais si y'en a beaucoup, j'vas enlever des points dans le vocabulaire. Y'en a que le mot est trop, tsé, populaire en tant que tel, ça j'vas le souligner.* » [E6]

vocabulaire est souvent subordonné à une tâche d'écriture et consiste la plupart du temps en la construction d'une banque de mots autour d'un thème comme le conte, le merveilleux ou la science-fiction, souvent en se concentrant sur la recherche de synonymes pour éviter les répétitions. Malgré le fait qu'ils disent que l'enseignement du vocabulaire est l'affaire du cours de français, l'écrasante majorité des enseignants interviewés n'y accordent que peu d'importance, pour deux raisons principales : le manque de temps et le manque d'outils pour enseigner le vocabulaire.

En effet, plusieurs enseignants ont affirmé qu'ils aimeraient travailler davantage le vocabulaire, mais que des contraintes de temps les obligent à établir certaines priorités didactiques, dont le vocabulaire fait rarement partie, ce qu'on peut comprendre lorsqu'on considère le nombre effarant d'erreurs d'orthographe dans les rédactions des élèves.

C'est la job du cours de français [d'enseigner le vocabulaire], mais je peux pas dire que j'accorde beaucoup de temps à ça. Mais c'est sûr qu'on met toujours des priorités. [...] Mais non, moi, je pense qu'il faut accorder du temps au vocabulaire. Ça serait mon idéal. [E5]

R'garde... je l'enseigne... c'est insatisfaisant... on fait des choix euh... On choisit nos bagarres, on choisit nos luttes... [...] J'pourrais demander [les mots] à l'intérieur de phrases, j'pourrais composer des textes. J'pourrais faire plein d'affaires, je l'sais... si on avait trois fois plus de temps. [E3]

La moitié des enseignants rencontrés ont aussi mentionné le manque d'outils pour travailler le vocabulaire. Seuls deux des sujets ont dit utiliser parfois des exercices de vocabulaire issus de matériel didactique; quelques autres ont déclaré utiliser plutôt leur propre matériel.

Pis j'ai pas non plus les outils adéquats pour faire plus d'exercices de vocabulaire. [...] C'est difficile aussi de travailler avec ça, parce qu'on n'a pas les outils. Souvent dans les cahiers d'exercices que les maisons d'édition nous fournissent, c'est très euh... très sommaire, hein, comme activités. Pour le vocabulaire, y'a rien. C'est toute du matériel maison du coup. [E2]

Le temps consacré au vocabulaire en classe est donc très limité. Mais que font les enseignants en lien avec le vocabulaire lorsqu'ils décident de le travailler? Voici un survol des différentes pratiques qui nous ont été rapportées par nos sujets.

3.2.1.1 LES PRATIQUES D'ENSEIGNEMENT DU VOCABULAIRE EN LIEN AVEC L'ÉCRITURE

Ce qu'on remarque d'abord, comme nous l'avons mentionné plus haut, c'est que beaucoup d'activités de vocabulaire sont au service d'une tâche d'écriture. En effet, beaucoup d'enseignants ont dit, préalablement à une production écrite, travailler avec les élèves à construire une banque de mots utiles, à développer un champ lexical autour d'un thème. La plupart du temps, ce travail est fait de façon collective – sur le principe d'un remue-méninge où chaque élève peut proposer des termes –, mais il peut arriver que l'enseignant demande ensuite aux élèves d'enrichir de façon individuelle la liste commune. Il ne semble pas que ce type d'activités soit menées de façon très structurée, les élèves fournissant un peu au hasard les mots qu'évoque pour eux le thème proposé, sans avoir à expliquer les liens entre ces mots ou à les organiser en les regroupant en sous-ensembles (ex. synonymes, antonymes, etc.). Le seul lien lexical qui semble exploité de façon plus explicite est la synonymie, pour aider les élèves à éviter les répétitions lors de l'écriture. En lien avec la synonymie, un des enseignants rencontrés fait aussi des petits *quiz* informels au cours desquels il demande aux élèves de trouver *le plus de synonymes d'un terme* ou *cinq formulations pour décrire votre état d'âme* [E2], par exemple si vous êtes triste. Ce même enseignant tente aussi de sensibiliser les jeunes à la polysémie de termes synonymes.

Ben, par exemple si j leur dis, OK, un synonyme à PROFOND? Y vont dire CREUX. Oui, bon, c'est bien, tsé... mais si je dis « j'ai des sentiments profonds » pis « j'ai des sentiments creux », est-ce que ça veut dire la même chose? Non. Ben, tsé, là on travaille [...] vraiment au niveau du sens. [E2]

Un des enseignants dit aussi utiliser la dictée pour présenter la structure propre à chaque type de texte, mais aussi pour introduire un nouveau vocabulaire.

Quand on fait le texte explicatif, ben ma dictée va durer toute l'étape, mais tsé, j'donne l'intro, j'donne un paragraphe, j'donne un autre paragraphe, pis j leur donne un exemple, dans l'fond, de texte... J'dis « Tu vois? Ça, j'aime ben ça comme mot, comme vocabulaire, comme formulation, là. C't'impresionnant. » Y ont un modèle en dictée. [E4]

Lors de tâches d'écriture, l'imposition de certaines contraintes par les enseignants amène aussi les élèves à viser un vocabulaire plus riche et varié. Un enseignant exige que les élèves résumant l'idée principale de leur texte au début

et à la fin de chaque paragraphe, ce qui les oblige à varier leur vocabulaire pour exprimer une même idée à plusieurs reprises. Un autre interdit l'utilisation de certains mots *passerpartouts*.

Moi, au début d'année... bon, j'leur laisse un mois, y'écrivent comme y pensent. Pis après ça, j'leur interdis le verbe FAIRE. Après ça, j'leur interdis le verbe AVOIR. J'leur interdis GENTIL, BEAU, GROS, GRAND, mais j'leur fournis des banques de synonymes. C'est ça, les feuilles de vocabulaire avec lesquelles y vont pouvoir travailler. [...] Sont pas très habiles avec ça, mais... à force de travailler, de travailler, ben... [E2]

Plusieurs sujets ont aussi affirmé travailler avec un guide ou fascicule qu'ils fournissent aux élèves; ce petit aide-mémoire contient des règles de grammaire, des procédures de révision ou stratégies d'autocorrection, des modèles de conjugaison, des trucs pour distinguer les principaux homophones, mais aussi des listes de synonymes pour des mots fréquents, de marqueurs de relation ou de termes liés à un thème, le conte par exemple. Il s'agit donc d'un outil que les élèves peuvent enrichir tout au long de l'année avec des mots de vocabulaire qu'ils peuvent ensuite réinvestir lors de tâches de rédaction.

3.2.1.2 LES PRATIQUES D'ENSEIGNEMENT DU VOCABULAIRE EN LIEN AVEC LA LECTURE

La lecture constitue bien sûr une source importante d'accroissement du vocabulaire et certains enseignants tentent d'en tirer profit pour enseigner le vocabulaire, notamment en encourageant les élèves à noter les mots nouveaux qu'ils rencontrent ou encore ceux qu'ils trouvent beaux dans le but de les réinvestir plus tard en écriture.

Un des enseignants propose une activité de vocabulaire plus structurée en lien avec la lecture :

Pour chaque œuvre qu'y lisent, je ressors tous les mots qui je crois leur posent problème qu'y seront, peut-être pas tous, mais pour plusieurs, obligés d'chercher dans l'dictionnaire. Mettons pour une œuvre de 60 pages, Novecento de Baricco, m'a sortir à peu près, j'sais pas, moi, 120 mots. Michel Tremblay, Un ange cornu avec des ailes de tôle : 200 quelque mots. Et là, j'leur demande à chaque semaine, y doivent chercher dans l'dictionnaire les définitions d'ces mots-là, mais... les définitions qui respectent le contexte qui correspond au sens dans l'œuvre. Bon. Et à chaque semaine ou à chaque deux semaines, en même temps que mes évaluations de conjugaison, j'vais dicter une définition, et ils doivent m'orthographier correctement le bon mot en lien avec la définition. Pis, j'fais pas vingt mots une semaine, pis vingt mots l'autre semaine... ça s'accumule : 20 mots, pis là 40, pis 60... je reviens tout le temps piger dans les mots des semaines précédentes. [E3]

D'autres enseignants se servent de la lecture pour aborder certaines notions en lien avec le lexique. Par exemple, un des sujets utilise le conte pour sensibiliser les élèves aux différents registres de langue; en écoutant des conteurs et en lisant des contes, les élèves réalisent que la narration, plutôt familière à l'oral, doit être faite dans une langue standard à l'écrit, mais qu'il est correct d'utiliser une langue plus proche de l'oral dans les dialogues. Deux de nos sujets s'appuient sur la lecture de textes publicitaires ou poétiques pour aborder les figures de style avec leurs élèves. Un demande aussi à ses élèves de repérer les marqueurs de relations dans des textes dans le but de comprendre comment ils fonctionnent. Il est intéressant de souligner que les activités présentées dans ce paragraphe ne visent pas directement l'accroissement du stock lexical des élèves, mais visent plutôt à faire apprendre des notions liées au lexique, ce qui s'éloigne de l'enseignement plus traditionnel du vocabulaire pour tendre vers une véritable didactique du lexique.

3.2.1.2 LES ACTIVITÉS « DÉCROCHÉES » EN LIEN AVEC LE VOCABULAIRE

Malgré le fait que les enseignants disent avoir peu de pratiques ciblées sur le vocabulaire, certaines activités purement lexicales nous ont été rapportées.

Plusieurs enseignants ont dit travailler l'analyse morphologique en faisant des exercices sur les préfixes et suffixes fréquents, souvent dans le but de fournir aux élèves des outils supplémentaires pour comprendre le sens de mots nouveaux rencontrés en lecture. Des exercices visant spécifiquement les organisateurs ont aussi été mentionnés, tout comme d'autres portant sur les noms au genre difficile (ex. OPINION) ou les anglicismes. C'est principalement pour ce type de notions que les enseignants ont recours aux exercices proposés dans le matériel didactique, même si beaucoup préfèrent encore utiliser du matériel qu'ils ont eux-mêmes fabriqué ou accumulé au cours des années; un des sujets a dit avoir recours à certains exercices proposés sur le site Internet du *Centre collégial de développement de matériel didactique* (www.ccdmd.qc.ca).

Certaines des pratiques qui nous ont été rapportées sont orientées davantage vers l'acquisition de stratégies en lien avec le développement de la compétence

lexicale. Quelques enseignants ont dit mettre beaucoup l'accent sur la notion de synonymie lorsqu'ils abordent les procédés de reprise. Un des sujets dit travailler explicitement l'utilisation du contexte pour inférer le sens d'un mot inconnu : il s'agit aussi de faire comprendre aux élèves que tous les mots n'ont pas la même importance dans un texte et de les habituer à repérer lesquels sont essentiels à la compréhension et méritent qu'on s'y attarde.

Je l'ai fait avec eux autres, essayer de comprendre avec le contexte. Y'en a certains qui ont tendance à chercher tous les mots qu'y comprennent pas. Ben, moi non plus, j'comprends pas tout le temps tous les mots. Tu peux revenir si tu te rends compte, après ça, que c'tait ben important que tu comprennes ce mot-là ou que tu saches exactement c'tait quoi, là. [...] Y'en a qui s'en sacrent, mais y'en a qui disent : « Oh, mais là, j'comprendrai rien... j'comprends pas trois mots! ». [...] Faut leur montrer qu'on peut le trouver d'une autre façon. Mais pas chacun des mots parce que sinon, ça vient lourd, lire les textes... Y veulent pu lire, j'les comprends! [E7]

Certaines pratiques visent plus spécifiquement l'utilisation du dictionnaire, parce, comme le mentionne un des sujets : *y faut tu fasses des activités de recherche dans le dictionnaire pour les familiariser avec le dictionnaire, sinon personne va en chercher un [lors de tâches d'écriture] [E6]*. Une des activités les plus répandues semble être le *concours de dictionnaire*, au cours duquel les élèves disposent d'un temps limité pour trouver un certain nombre de mots dans le dictionnaire et noter le numéro de la page; ce type d'exercice, plus fréquent au début du secondaire aux dires de certains, s'accompagne parfois d'une consigne supplémentaire : il s'agit alors de demander aux élèves de créer une phrase exemple pour chacun des mots cherchés, afin de vérifier plus finement leur compréhension du sens. Un des enseignants rencontrés a aussi dit tenter de sensibiliser les élèves à la richesse d'informations fournies par le dictionnaire, notamment en ce qui concerne les collocations.

[Mais l'élève qui a écrit « j'ai réussi mon but », j'imagine qu'y est venu vous voir. Mais est-ce qu'il aurait pu trouver la bonne formulation dans l'dictionnaire ?] Oui. Parce que tu vas voir le mot BUT, pis y vont t'donner des verbes qui vont aller avec. Quand t'utilises correctement l'dictionnaire... Au secondaire, faut que ce soit acquis. [...] Y'a des profs qui mettent beaucoup beaucoup l'emphase là-dessus, pis d'autres non. [...] C'est pour ça que j'leur dis : « Le Petit Robert, c'est ton nouveau meilleur ami ». Parce que tu cherches un mot, t'as l'verbe, on t'donne même un petit bout de phrase. Tu peux r'prendre ça dans ton texte. [E2]

Les entrevues que nous avons menées nous ont donc permis de faire un modeste inventaire des pratiques liées à l'enseignement du vocabulaire en classe de français au secondaire. Nous constatons que les activités qui ciblent spécifiquement le vocabulaire sont plutôt rares, et que celui-ci est plutôt abordé par l'intermédiaire d'autres tâches, comme l'écriture ou la lecture. Les activités typiquement lexicales semblent souvent perçues par les enseignants comme des exercices d'enrichissement plus ou moins essentiels.

Dans le cahier d'exercices qu'on avait acheté cette année, y avait des petits exercices là-dessus. Je donnais ça à la fin d'un cours [...] pis les élèves aiment ça faire ça habituellement, là. [E6]

3.2.2 L'ERREUR LEXICALE ET LA CORRECTION DE PRODUCTIONS ÉCRITES

Un des buts de notre thèse était d'observer de quelle façon les enseignants notent et traitent l'erreur de vocabulaire dans le cadre des productions écrites de leurs élèves, lieu privilégié de l'évaluation du vocabulaire en classe de français. C'est sur cette question que porte la présente section. Nous nous pencherons tout d'abord sur les grilles de correction utilisées par les enseignants pour l'évaluation de l'écrit et la place qu'y occupe le vocabulaire. Nous verrons ensuite ce que pensent nos sujets de la dimension « vocabulaire » de la tâche d'évaluation, pour ensuite nous intéresser à un aspect plus concret : l'annotation de l'erreur lexicale et le suivi après la correction.

3.2.2.1 DES GRILLES DE CORRECTION QUI LAISSENT PEU DE PLACE AU VOCABULAIRE

La grille de correction du ministère de l'Éducation¹⁸⁷ utilisée pour la correction du texte argumentatif dans le cadre de l'épreuve d'écriture de 5^e secondaire sert de modèle pour la majorité des grilles dont se servent les enseignants. C'est pourquoi nous jugeons pertinent de présenter rapidement cet outil.

La grille du MELS grille se divise en deux grands volets, valant chacun 50 % de la note finale : « cohérence de l'argumentation » et « respect du code linguistique ». Le second volet, qui nous intéresse plus particulièrement ici, contient trois critères : « utilisation des mots », qui représente 5 % de la note; « construction

¹⁸⁷ http://www.mels.gouv.qc.ca/dfgj/de/pdf/flm5_grille03.pdf

de phrases et ponctuation », qui vaut 25 %; « orthographe », qui regroupe les erreurs d'orthographe d'usage, d'accord et de conjugaison, pour un total de 20 %. Contrairement aux deux autres critères linguistiques, le vocabulaire est évalué de façon qualitative par une appréciation globale du vocabulaire utilisé par l'élève dans son texte. Voici l'échelle descriptive proposée par le ministère pour ce critère.

		NIVEAU DE COMPÉTENCE				
		E	D	C	B	A
			Compétence insuffisante pour un finissant du secondaire	Compétence acceptable pour un finissant du secondaire	Compétence assurée pour un finissant du secondaire	Compétence marquée pour un finissant du secondaire
4. Utilisation des mots (5 %)	isante aire		Utilise des mots ou des expressions dont plusieurs sont imprécis, incorrects ou appartiennent à un registre de langue familier	Utilise des mots ainsi que des expressions généralement précis, corrects et appartenant à un registre de langue standard et parfois familier	Utilise des mots ainsi que des expressions précis, corrects et appartenant à un registre de langue standard, à l'exception de rares erreurs dans les termes courants	Utilise des mots ainsi que des expressions précis, corrects et appartenant à un registre de langue standard, à l'exception de rares erreurs principalement dans les termes peu courants

Figure 20. Les niveaux de compétence pour le critère vocabulaire dans la grille de correction du MELS (2003: 15)

Relativement au vocabulaire, le MELS apporte aussi quelques précisions supplémentaires. On mentionne notamment que le vocabulaire est évalué au critère 4 sous un angle normatif, par exemple à travers les expressions fautives et les problèmes liés au registre de langue, mais qu'on en tient aussi compte sous l'angle de la communication au critère 2 (*Utilise divers moyens [dont le vocabulaire] pour susciter l'intérêt du destinataire*) ainsi qu'au critère 3 (*substituts variés, liens variés*) (MELS, 2003: 10).

Tous les enseignants rencontrés utilisent pour la correction de productions écrites une grille s'apparentant à celle proposée par le ministère de l'Éducation, grille qui peut changer en fonction du type de texte produit (narratif ou explicatif). Trois des huit sujets ont construit eux-mêmes leur grille en s'inspirant de celle du MELS, alors que celle utilisée par un autre est le fruit du travail de tous les enseignants de troisième secondaire de son école; les quatre autres sujets corrigent à l'aide d'une grille fournie par leur commission scolaire. Dans tous les cas, la grille de correction comporte, comme celle du MELS, deux grandes parties : une s'intéressant au fond (contenu, structure du texte et cohérence) et l'autre au code (syntaxe, ponctuation, orthographe grammaticale, usage), la seconde valant généralement un peu moins de la moitié de la note

(souvent 40 %). L'évaluation des critères s'intéressant au fond est généralement faite de façon globale, à l'aide d'échelles descriptives, alors que les critères liés au code font l'objet d'une évaluation quantitative, soit basée directement sur le nombre d'erreurs, soit par un calcul plus complexe visant à établir le pourcentage d'erreurs¹⁸⁸.

Dans un peu plus de la moitié des grilles utilisées par les enseignants interviewés, le vocabulaire correspond à un critère spécifique de l'évaluation, généralement parmi les critères liés au code, mais parfois du côté du fond. Dans trois des grilles que nous avons pu observer, le vocabulaire ne fait pas l'objet d'une évaluation spécifique et entre plutôt dans un critère plus global appelé « contenu », toujours évalué qualitativement à l'aide d'une échelle descriptive (ex. *l'élève utilise un français parfois riche et très bien adapté au conte*). Dans le cas où le vocabulaire correspond à un critère spécifique, il représente généralement 5 % de la note finale¹⁸⁹ et est généralement évalué par un jugement global, souvent à l'aide d'une échelle descriptive allant de 1 à 5 ou de A à E.

3.2.2.2 L'ÉVALUATION DU VOCABULAIRE : UNE TÂCHE FLOUE ALLÉGÉE PAR L'EXPÉRIENCE

Les enseignants que nous avons rencontrés admettent que l'évaluation du vocabulaire en production écrite n'est pas toujours facile. Certains soulèvent le caractère un peu flou des critères sur lesquels repose l'appréciation du vocabulaire et des échelles descriptives utilisées :

C't'un peu flou, là. « Utilise des mots précis, appropriés et variés... euh, des expressions, des mots appropriés la plupart du temps... euh généralement »... euh... Ici c'est clair, là : « des mots incorrects... plusieurs sont incorrects ou répétés »... Mais tsé, c'est quoi la différence entre « généralement » pis « la plupart du temps », ou « tout le temps », là? Euh... j'y vas avec mon jugement. [E7]

¹⁸⁸ Par exemple en divisant le nombre d'erreurs d'orthographe grammaticale par le nombre de mots, puis en convertissant ce pourcentage en lettre selon un barème du type 0 à 1 % d'erreurs = A, 1 à 2 % = B, etc.

¹⁸⁹ Un seul de nos sujets accorde 5 points sur 65 au vocabulaire, soit près de 8 %.

D'autres pointent la subjectivité liée à l'évaluation du vocabulaire d'un texte, qui peut varier d'un correcteur à l'autre davantage que les autres dimensions de la correction.

Ça dépend toujours aussi de ton enseignement pis de tes exigences. Parce que moi, [...] j'peux considérer qu'un jeune a utilisé un très beau vocabulaire, pis mon collègue à côté va trouver que c'est ben ben ben ordinaire. [E2]

Les correcteurs ne savent pas trop comment évaluer le vocabulaire, indique E2 à propos des correcteurs engagés par le ministère pour corriger l'épreuve unique de 5^e secondaire. Certains sujets avouent d'ailleurs se trouver dans la même situation :

Mais... sur le plan du vocabulaire euh... ben je sais pas mais en tout cas, c'est toujours flou là, la correction du vocabulaire pour moi. Tsé, [...] faut-tu que tu comptes des mots qui sont ultra recherchés dans le texte, tsé combien qu'y en a mis ? Ça, ça mérite un 4, un 5 sur 5? [E5]

Les enseignants disent ne pas avoir été formés pour évaluer le vocabulaire, ce qui fait qu'ils se sentent un peu dépourvus face à cette tâche, particulièrement en début de carrière. Avec l'expérience, ils s'en remettent à leur intuition.

Au début, oui, [je trouvais difficile d'évaluer le vocabulaire] parce que je savais pas, mais là, à un moment donné, tu te fixes un barème personnel, pis tu y vas à partir de ça. Mais là non, je trouve plus ça difficile d'évaluer le vocabulaire parce que je me suis vraiment fixé des choses qui pour moi étaient avoir un bon vocabulaire ou ne pas en avoir un. [E6]

Lors des entrevues, nous avons senti que certains enseignants étaient un peu déstabilisés par nos questions précises sur leur façon d'évaluer le vocabulaire, parce qu'il semble justement que ce volet de l'évaluation repose en grande partie sur leur impression générale et ne fait pas souvent l'objet d'une réflexion de leur part, ce que résume assez bien E3 :

... pis j'me suis jamais arrêté à dire : « Bon, comment j'pourrais évaluer le vocabulaire autrement? » [...] Non, moi je pense que c'est le critère sur lequel on accorde le moins de temps et le moins de concentration quand on corrige.

Globalement, l'évaluation du vocabulaire semble être pour les enseignants une tâche nécessaire de la correction de productions écrites pour laquelle ils ne disposent que de peu de balises ; ils s'en remettent donc à leur expérience, ce qui explique que cette tâche perçue comme difficile en début de carrière s'allège au fil des années.

3.2.2.3 L'ÉVALUATION DU VOCABULAIRE : UN JUGEMENT GLOBAL AYANT PEU D'IMPACT SUR LE RÉSULTAT FINAL

Nous avons demandé aux enseignants de quelle façon ils se servaient des échelles descriptives pour évaluer le vocabulaire d'une rédaction. Il ressort tout d'abord des entretiens que l'évaluation du vocabulaire est toujours soumise à une appréciation globale, c'est-à-dire que les enseignants n'établissent pas de correspondance directe entre le nombre de problèmes de vocabulaire relevés (répétition, mot mal utilisé, terme familier) et une note sur l'échelle du critère « vocabulaire » de la grille de correction (par exemple, un élève ayant commis deux erreurs de vocabulaire obtient une cote de B pour ce critère); bien sûr, le nombre d'erreurs relevées fournit tout de même un point de repère à l'enseignant lorsqu'il pose son jugement global pour ce critère.

La plupart des enseignants rencontrés utilisent une note « par défaut » pour le vocabulaire, généralement 3 sur 5 ou B dans une échelle de A à E¹⁹⁰. Cette note correspond selon eux au niveau de vocabulaire moyen d'un élève de troisième secondaire; lorsque le vocabulaire utilisé est trop répétitif, imprécis ou présente un certain nombre d'erreurs, ils ajustent à la baisse la note attribuée pour le vocabulaire. Lorsqu'ils jugent au contraire que l'élève emploie un vocabulaire plus riche que celui auquel on pourrait s'attendre de la part d'un jeune de cet âge, qu'ils perçoivent un effort, qu'ils retrouvent des mots de la *banque de mots* [E5] développée avec les élèves autour d'un thème ou qu'ils sont *impressionnés* [E4], ils augmentent la note.

Et j'vas donner – c'est très rare – 5 sur 5 à un élève quand j'le lis et que les mots de vocabulaire qu'il utilise, j'suis obligé d'aller les chercher dans l'dictionnaire pour bien comprendre ce qu'y veut dire, pis j'me rends compte que sa maîtrise de la langue dépasse presque c'que moi j'arriverais à faire dans le même type de texte... Là, j'vas donner 5 sur 5, en expliquant que ça veut pas dire qu'y ont un vocabulaire parfait... C'est pas parce que t'as la note parfaite... t'as un vocabulaire que je considère parfait pour un élève de niveau secondaire 3. [E3]

Deux des huit enseignants rencontrés utilisent au contraire une approche soustractive, c'est-à-dire que les élèves obtiennent par défaut la note parfaite,

¹⁹⁰ Un seul de nos sujets attribue par défaut la cote C pour le vocabulaire, mais ses élèves sont des élèves à risque et des doubleurs, ce qui fait qu'ils utilisent peut-être effectivement un vocabulaire plus simple que celui de leurs camarades des classes régulières.

qui sera ensuite revue à la baisse selon le niveau de répétition et le type de vocabulaire utilisé.

Ben écoute. Moi j'y vais très simplement, hein. Quand j'vois qu'y a un mot qu'y est utilisé trop souvent, souvent j'vas aller l'encadrer, pis j'vas y faire des lignes pour y dire, r'garde, là, lui... [...] Euh... j'leur montre qu'y l'utilisent trop souvent. Pis quand qu'y a un mot qu'y est mal utilisé, j'vas écrire « voc - ». Mais quand qu'y a des beaux mots, j'mets rien : j'vais vers plus le négatif. Faque quand j'arrive dans ma grille pis que je r'garde mes, mes critères, pis là j'me dis « vocabulaire », pis tu regardes vite vite pis parce que j'me suis faite des carrés, ben j'vas voir plus rapidement. Pis j'vois les « voc » aussi rapidement. Faque tout d'suite je l'sais que ici, y'aura pas A... [E1]

Trois de nos sujets ont mentionné faire preuve de clémence face à certaines erreurs lexicales lorsqu'ils ont l'impression que l'élève qui les a commises a fait des efforts pour utiliser un bon vocabulaire, c'est-à-dire qu'ils ne revoient pas la note attribuée au vocabulaire à la baisse simplement sur la base de ces maladresses; par contre, l'élève en question n'obtiendra jamais une note parfaite.

Pis même si y font des erreurs, ça montre une certaine... Tsé, j'leur dis « Tu peux t'lancer à l'eau ». Parce que c'est sûr que si t'hésites entre deux mots et que tu prends toujours le plus facile, ben euh... t'es pas pénalisé, tsé. Mais quand t'essaies, quand tu joues dangereusement... ou quand t'essaies d'utiliser du vocabulaire un petit peu plus recherché, ben j'trouve que ça démontre un effort supplémentaire. [...] si l'enfant fait une erreur de vocabulaire, mais que je vois, que je sens que dans son texte y'a un effort pour varier son vocabulaire – faut pas que ça soit toujours boiteux, là -, mais si y'a une erreur ou deux, je l'pénaliserais pas, parce qu'au moins, y va compenser... Tsé, y va avoir comme le défaut de ses qualités. Y va avoir pris des chances. Pis ça, j'trouve qu'y faut encourager les jeunes à prendre des chances. Prendre des chances, parce que sinon euh... on a tout le temps des formulations, euh... toutes cuites d'avance. [E2]

Tous les enseignants interviewés disent ne jamais attribuer la note la plus basse possible à un élève pour critère « vocabulaire », sauf un, qui avoue donner parfois un D à ceux qui se sont pas forcés partout [E7]. Généralement, un nombre important de répétitions ou un vocabulaire du niveau primaire [E3] fera baisser d'un cran seulement la note de vocabulaire « par défaut » (par exemple, C au lieu de B). Deux enseignants ont aussi admis que la note attribuée pour le vocabulaire pouvait aussi être influencée par ce qu'ils connaissent du niveau de vocabulaire de l'élève évalué; ainsi, un élève ayant beaucoup amélioré son vocabulaire pourrait recevoir un B pour un texte qui aurait valu un C à l'un de ses camarades.

En ce qui concerne l'impact du résultat obtenu au critère « vocabulaire » sur la note globale de la rédaction, il semble plutôt minime. Nous savons déjà qu'il représente un faible pourcentage de cette note, mais les enseignants avouent en plus n'y accorder que peu de temps et semblent le considérer comme un critère de deuxième importance¹⁹¹, comme le démontre le dialogue suivant entre E8 et une collègue qui est intervenue lors de l'entrevue.

COLLÈGUE : *Est-ce que le vocabulaire affecte ta note, toi? Mettons y'a E en vocabulaire?*

E8 : *Ça sera pas mon critère principal.*

COLLÈGUE : *Moi, l'vocabulaire, là, pour vrai... Mettons qu'y a B, là, pis qu'y a eu D dans toute, sa note c'est D quand même.*

E8 : *Oui, oui! C'est clair que j'y mettrai pas C parce qu'y a eu B en vocabulaire.*

COLLÈGUE : *Oui, oui, j'comprends c'que tu veux dire. Mais tsé, d'ailleurs, dans les anciennes grilles, t'avais 20 % pour grammaire orthographe, 20 % pour syntaxe ponctuation, 10 % pour le vocabulaire. Faque un élève... Tsé dans l'fond, on parle de vocabulaire, mais c'pas une note... c'est pas la note qui va affecter tout le reste, tsé. En tout cas!*

E8 : *Pis d'ailleurs, ici, c'est bien écrit « Au moment de porter un jugement global, l'enseignant accorde une importance particulière aux critères 4 et 5 ».¹⁹² Faque le vocabulaire, dans l'fond... C'est pas que c'est pas important, là, mais...*

Il arrive par contre que l'erreur de vocabulaire affecte le résultat global de la rédaction par l'intermédiaire d'autres critères. Nous avons en effet observé que plusieurs erreurs de vocabulaire sont comptabilisées par les enseignants dans d'autres classes d'erreurs, particulièrement la syntaxe, et dans une moindre mesure, la cohérence; un enseignant dit aussi que le vocabulaire influence la note du critère « style » de sa grille de correction. Les erreurs de vocabulaire qui sont le plus souvent considérées comme telles sont les problèmes de répétitions, les problèmes sémantiques et les erreurs de registres de langue; ce sont

¹⁹¹ Un seul de nos sujets a affirmé accorder beaucoup d'importance au vocabulaire dans la correction, surtout pour le texte narratif : « *Dans l'prochain qu'on fait, qui est un texte narratif, ben va falloir qu'y utilisent les éléments du merveilleux, du fantastique ou de la science-fiction. Donc va falloir qu'y adapte son vocabulaire. [...] Faque là, moi, j'insiste beaucoup beaucoup sur le vocabulaire. Parce qu'y ont des ressources, pis parce qu'y faut qu'y s'pratiquent aussi, pis parce qu'on est rendus au mois de mai. Ça fait longtemps qu'y travaillent.* » [E2]

¹⁹² Les critères 4 et 5 sont ceux liés à l'orthographe (grammaticale et lexicale), à la syntaxe et à la ponctuation.

généralement ces problèmes que les enseignants prennent en compte lorsqu'ils portent un jugement sur le vocabulaire d'un texte.

Nous avons soulevé plus haut la difficulté pour les enseignants de classer les erreurs lexicales dans leur grille de correction, notamment parce qu'elles entraînent souvent des problèmes syntaxiques. Le fait que le vocabulaire soit évalué de façon qualitative semble aussi poser problème à certains enseignants, qui préfèrent pouvoir compter le nombre d'erreurs pour donner une note et utilisent donc davantage les grandes classes de leur grille de correction, même pour les problèmes de répétition, les erreurs sémantiques et l'emploi de termes familiers.

J'vas enlever, moi, aussi au niveau d'la syntaxe si c'est trop répétitif, là. [À propos d'un problème sémantique :] Ouais. J'l'aurais mis dans syntaxe : mauvais mot utilisé. Moi, je l'coche [dans la colonne syntaxe]. Regarde, honnêtement ça va plus vite. J'vas marquer « vocabulaire » à côté, mais de toute façon, que j'l'enlève là (...) j'vas y'enlever pareil. [E4]

Y'en a que le mot est trop, tsé, populaire en tant que tel ; ça, j'vas le souligner, j'vas le compter aussi dans l'orthographe. [E6]

3.2.2.4 DES ANNOTATIONS VAGUES ET PEU SYSTÉMATIQUES

Les informations que nous ont données les enseignants quant à leur façon d'annoter les erreurs lexicales correspondent en tout point à ce que nous avons pu observer sur les rédactions de notre corpus qui étaient déjà corrigées (voir 1.1).

La répétition est le type de problème le plus souvent relevé par écrit parmi les erreurs lexicales. Les enseignants entourent, soulignent (par des lignes droites ou des *petites vagues* [E4]), surlignent ou encadrent les mots utilisés de façon abusive et, dans presque tous les cas, relient les répétitions par des lignes ou des flèches de façon à former *une toile* [E3] ou *un réseau* [E5], ceci dans le but de conscientiser l'élève aux nombreuses répétitions, mais aussi de justifier la note attribuée pour le vocabulaire.

Les mots mal utilisés – c'est-à-dire les erreurs sémantiques, mais aussi certains problèmes de collocation et de registre – sont quant à eux soulignés, barrés ou

entourés d'un *petit nuage* ^[E2/8], marques généralement accompagnées d'une annotation (« *V* », « *voc* », « *vocabulaire* » ou « ? ») dans la marge ou au-dessus du mot problématique. Dans certains cas, beaucoup plus rares, l'annotation pourra être plus précise, par exemple « *forme orale* » ou « *familier* »; plus rarement encore, un court commentaire expliquant à l'élève pourquoi le terme utilisé est problématique sera fourni. Deux enseignants ont dit marquer tant les bons coups lexicaux de leurs élèves que leurs erreurs à l'aide des annotations « *V+* » et « *V-* ».

Seul un des huit sujets a affirmé fournir systématiquement une suggestion de correction au-dessus d'un mot mal utilisé; les autres enseignants rencontrés fournissent parfois une proposition de correction, mais pas à tout coup. D'ailleurs, comme les enseignants ne jugent pas que toutes les erreurs méritent d'être relevées, il est fréquent qu'ils ne fournissent aucune annotation pour une erreur lexicale qu'ils considèrent sans importance, voire qu'ils ne la relèvent même pas.

Nous constatons donc que les annotations les plus fréquemment utilisées par les enseignants pour noter les erreurs lexicales autres que la répétition (soulignement, cercle, « *V* », « *voc* », « *vocabulaire* », « *V+* », « *V-* ») demeurent plutôt vagues et ne sont sans doute pas très aidantes pour les élèves, puisqu'elles ne les renseignent que très peu sur la nature du problème relevé par l'enseignant; pour employer la typologie des annotations proposée par Roberge (2006), nous pourrions affirmer qu'en lien avec le vocabulaire, les enseignants privilégient les commentaires non développés et la correction de l'erreur. Devant un mot souligné accompagné d'un « *voc.* », l'élève doit-il comprendre que le terme employé est inexistant, qu'il appartient au langage familier, qu'il s'agit d'un anglicisme, qu'il ne correspond pas au sens visé? Certains enseignants admettent d'ailleurs qu'ils utilisent ces annotations davantage pour eux que pour leurs élèves.

Bon, j'ai mis un petit « v », vocabulaire! [Donc tu soulignes, pis tu mets « v » ?] Ouin, ben ça c'est plus pour moi, là. Tsé les lettres, j'pense pas qu'y... les comprennent. Si y'utilisent un mot dans l'mauvais sens, je l'souligne, pis j'mets un point

d'interrogation. L'élève des fois y va venir me voir : « Qu'est-ce qu'y a? ». « Qu'est-ce que tu voulais dire? » Tsé... [...] mais ça, c'est pour moi. Après ça, j'les compile, là. Des fois, je corrige tout en même temps pis je regarde les fautes à la fin. J'essaie d'être plus impartiale en faisant « v, v, v, a, c, v ». En l'faisant après coup, j'me souviens pu c'est qui. [E7]

Les enseignants rencontrés semblent d'ailleurs très conscients de l'utilité limitée d'annotations du type « *voc.* » ou « *V-* », qui de leur propre aveu n'aident pas les élèves à améliorer leur vocabulaire.¹⁹³ Ils croient par contre que le fait de suggérer un terme de remplacement aux élèves peut favoriser l'accroissement de leur vocabulaire, dans la mesure où ces derniers font l'effort de lire leurs commentaires, ce qui est plutôt rare selon eux.

À améliorer [leur vocabulaire], non. Ben... les mots que je propose, peut-être... Ceux qui vont s'donner la peine de bien observer ma correction pis d'en tenir compte. Ça, c'est deux, trois, quatre par groupe de trente. [Les autres,] ça passe dans le beurre. [E3]

[Est-ce que les annotations, par exemple suggérer un mot, peuvent les aider? Est-ce que tu les vois réutiliser ce mot après?] Je pense pas. Ben là, c'est dur de le réévaluer dans un autre contexte, mais j'ai l'impression que souvent j'écris des commentaires pis que ça sert pas à grand-chose. Peut-être parce que je suis pessimiste un peu... J'imagine que ça doit servir à quelques élèves. J'ose espérer que oui. [E5]

Une courte remarque s'impose à propos des rectifications orthographiques, puisque nous avons demandé aux enseignants s'ils en tenaient compte dans leur correction. La très grande majorité des enseignants ne semblent pas très au fait des rectifications et sont plutôt perplexes devant le manque de politiques claires quant à la façon d'en tenir compte. Aucun de nos sujets n'enseigne l'orthographe rectifiée et plusieurs ont affirmé n'accepter que l'orthographe qui apparaît dans les dictionnaires dont ils disposent en classe, qui souvent datent d'une dizaine d'années et ne font pas mention des nouvelles graphies. Pour un des sujets, c'est à l'élève de prouver qu'il a appliqué la nouvelle orthographe parce que la plupart du temps, *il le sait pas* [E6]. De façon générale, les enseignants nous ont paru indifférents face à cette évolution de la langue, lorsqu'ils ne la désapprouvent pas carrément.

¹⁹³ Un de nos sujets met même en doute la pertinence des codes de correction (S = syntaxe, P = ponctuation, G = accord. U = usage), pourtant très répandus, puisqu'il n'a pas l'impression qu'ils ont une très grande signification pour les élèves.

J'trouve ça triste, par exemple, les rectifications de l'orthographe. J'trouve que c'est comme prendre les jeunes pour des enclumes un peu. C'est trop... j'suis d'accord que le français évolue, j'suis tsé... C'est correct là, mais j'trouve que c'est prendre les jeunes pour pires qu'y sont. C'est pas leur faire confiance. Y'éprouvent déjà tellement de difficultés, j'trouve... pis c'est pas là qu'y font le plus d'erreurs. Tsé, NENUPHAR, y l'emploient pas. Mais, si ça avait été quelque chose de plus euh... tsé, on décide de plus accorder les participes passés avec AVOIR... pourquoi pas ça à place? Ça, ça aiderait vraiment et ça montrerait vraiment l'évolution. [...] Mais NENUFAR, les MILLEPATTES, personne utilise ça! [E2]

3.2.2.5 PEU DE SUIVI POUR L'ERREUR LEXICALE

Faute de temps, rares sont les enseignants qui demandent systématiquement aux élèves de retravailler leur texte après la remise d'une rédaction corrigée. En effet, seulement deux des enseignants interviewés demandent à ceux-ci de leur fournir une version corrigée de leur texte accompagnée d'une explication de leurs erreurs après la plupart des rédactions; un autre leur fait faire un travail similaire, mais par groupe de trois, de façon à ce que les élèves valident entre eux leurs corrections. Cependant, même lorsqu'il y a un suivi après la correction, celui-ci se concentre généralement sur les erreurs d'orthographe d'usage et les problèmes d'accord; seul un des enseignants demande aux élèves d'intégrer à leur texte ses suggestions lexicales et de varier les mots répétés.

Les cinq autres sujets rencontrés n'exigent pas de travail à posteriori sur le texte de la part des élèves, mais font tout de même un retour collectif sur les principaux problèmes rencontrés par les élèves. La plupart du temps, leurs commentaires visent la structure du texte ou certaines règles grammaticales, mais il peut arriver qu'ils concernent des éléments lexicaux, notamment orthographiques.

Ben, quand y'a une erreur qui revient fréquemment... à moment donné, on avait « tout à coup » avec des traits d'union, ou euh... des erreurs fréquentes d'orthographe... Ben là, j'leur dis « Écoutez, cette erreur-là revient souvent (...) Arrêtez, c't'une erreur qui est fréquente. » Si une erreur revient pour plusieurs élèves, j'vas la noter pis j'vas en parler. [E4]

Après le retour en grand groupe, les enseignants allouent ensuite une certaine période aux élèves pour qu'ils regardent leur rédaction corrigée et lisent les commentaires qui y figurent. Ils jugent qu'une minorité des élèves tirent profit

de cette période, puisque la plupart se contentent de regarder la note finale sans chercher à comprendre leurs erreurs. Ils mentionnent tout de même qu'un petit nombre d'élèves, *les plus motivés de la gang* [E5], viennent les voir pour tenter de comprendre leurs erreurs, notamment celles de vocabulaire. Un des enseignants souligne aussi que souvent, des élèves remettent en doute sa correction relativement à leurs choix lexicaux.

Oui, y en a qui viennent me poser des questions [...] Oui, pis souvent y s'obstinent, y croient pas qu'y ont faite une erreur avec le mot, tsé. Ils l'ont trouvé dans le dictionnaire des synonymes, pis y trouvaient que c'était vraiment un bon mot merveilleux qui était nouveau. Mais non, ça fonctionne pas! [E5]

Un autre des sujets pointe aussi le fait que si les élèves viennent poser des questions relatives à la correction du vocabulaire de leur texte, c'est souvent davantage pour comprendre la note qui leur a été attribuée pour ce critère qu'avec une réelle intention d'améliorer leur vocabulaire.

« Pourquoi vous m'avez mis « V- » là? J'ai perdu combien de points pour mon vocabulaire? Pourquoi j'ai juste 3 pis elle a l'a 4? » [E3]

Nous pouvons donc conclure que les erreurs lexicales font l'objet de peu de suivi de la part des enseignants. Ceci s'explique sans doute par le fait que les problèmes lexicaux varient énormément d'un élève à l'autre et que les enseignants n'ont pas le temps d'accompagner chaque élève individuellement dans la recherche du mot juste.

3.2.3 LE DICTIONNAIRE EN CLASSE DE FRANÇAIS

Outil lexical par excellence, le dictionnaire demeure notre meilleure fenêtre sur la norme lexicale et le plus puissant remède contre l'erreur lexicale. Nous avons donc voulu savoir de quelle façon ce type d'ouvrage est utilisé dans les classes de français au secondaire, tant par les enseignants que par les élèves.

3.2.3.1 POUR LES ENSEIGNANTS : UN FIDÈLE COMPAGNON... POURTANT MÉCONNU

Tous les enseignants interviewés ont dit utiliser un dictionnaire lorsqu'ils corrigent, principalement pour vérifier l'orthographe de certains mots (surtout après de longues heures de correction!); ils s'y réfèrent de temps à autre pour

s'assurer du sens d'un mot ou pour d'autres types d'informations, les prépositions régies par un verbe par exemple.

Nous avons été frappé par le fait que les enseignants ne semblent pas vraiment avoir de préférence quant au dictionnaire qu'ils utilisent. Hormis deux sujets, qui nous ont dit préférer *Le Petit Robert* parce qu'il est plus complet (que le *Le Petit Larousse illustré* ou *Le Multidictionnaire*), les autres choisissent simplement le dictionnaire qu'ils ont à la maison ou celui qui est le plus proche.

Celui que j'ai sous la main. Les deux sont à 'même place, mais j'vas prendre celui qui est le plus près. [...] Des fois y trainent! [E1]

Bien qu'ils n'éprouvent pas d'attachement particulier à un ouvrage ou à un autre, quelques enseignants ont tout de même mentionné que certains dictionnaires sont plus utiles pour certains types d'informations : *Le Petit Robert* est plus complet en ce qui a trait aux définitions, *Le Multidictionnaire* est intéressant pour les difficultés lexicales liées au français d'ici, par exemple les anglicismes et les registres de langue, ou pour le régime des verbes, puisqu'on y indique souvent de façon explicite quelle préposition utiliser avec un verbe particulier. Un enseignant a d'ailleurs affirmé tenter de sensibiliser les élèves aux différences entre les ouvrages.

Souvent, en lecture, y vont aller chercher le Multi, pis là j'leur dis : « Ça sert à rien. Y t'aidera pas. » Parce que souvent le Multi au niveau des descriptions... pis y'a des mots qui manquent aussi. Faque ça, j'leur dis que ça les aidera pas. [E1]

Le nombre et la variété des dictionnaires disponibles dans les classes semblent varier énormément d'une école à l'autre. Alors que certains enseignants se réjouissent de la panoplie d'ouvrages à leur disposition, d'autres déplorent le piètre état de leurs dictionnaires.

En classe on a toutes les dictionnaires, toutes les sortes, on a des Robert, des Larousse, des Multi, on en a comme 10 de chaque, moi je trouve ça merveilleux. [E6]

[Et en classe, les élèves travaillent avec quel dictionnaire?] *Ah mon Dieu! Des vieilles affaires! Comme le bleu, là... Le Petit Larousse illustré, y date des années... c'est ça qu'on a. [Pas de Robert? Pas de Multi non plus?] Ah, non, non! On n'a pas de dictionnaires de synonymes [...] J'ai des Bescherelle que j'ai amenés moi-même.* [E8]

Le Petit Larousse illustré et *Le Petit Robert* sont les dictionnaires que l'on retrouve le plus fréquemment dans les classes, même si *Le Multidictionnaire* semble gagner en popularité et est souvent présent en deux ou trois exemplaires. Environ la moitié des enseignants ont dit disposer aussi de quelques dictionnaires des synonymes dans leur classe (en plus de l'incontournable *Bescherelle*, bien sûr!).

3.2.3.2 POUR LES ÉLÈVES : UN OUTIL SOUS-EXPLOITÉ, DIFFICILE D'ACCÈS... ET DÉPASSÉ!

Selon les enseignants rencontrés, les élèves utilisent le dictionnaire lors de tâches d'écriture presque essentiellement pour vérifier l'orthographe de certains mots ou, de façon beaucoup plus marginale, pour varier leur vocabulaire en trouvant des synonymes; en lecture, c'est plutôt pour vérifier le sens de mots inconnus qu'ils s'y réfèrent. Si quelques enseignants ont mentionné que leurs élèves allaient spontanément chercher un dictionnaire lorsqu'ils se questionnent à propos d'un mot, d'autres disent que les élèves évitent de l'utiliser et vont plutôt tenter d'obtenir la réponse à leur question auprès de leur enseignant, quitte à laisser tomber dans le cas où ce dernier les renvoie au dictionnaire.

Oui, [les élèves utilisent souvent le dictionnaire], surtout cette année. C'est ce qui m'impressionne beaucoup. Même lors des examens de lecture. Quand y lisent des textes, y vont aller chercher l'dictionnaire parce qu'y comprennent pas les mots. C'est merveilleux! Wow! J'les regarde, j'les trouve beaux! [E8]

Y vont [vérifier le sens d'un mot], par exemple, quand on a eu à faire les notes de lecture, pis que là y'avaient un recueil de textes à lire, mais c'est ben plus facile d'aller voir le prof! Pis comme j'veux qu'y l'sachent, ben souvent j'aime mieux leur dire... Y'en a que j'vas dire « Ben, cherche-le! » pis qu'y vont dire « Ben tant pis! » À ce moment-là, j'ai tendance à leur expliquer, à leur dire, ce mot-là veut dire telle chose [...] Parce que sinon, ils l'sauront pas pis y vont juste continuer. [E7]

Nos sujets semblent dire que consulter le dictionnaire demeure une activité fastidieuse pour beaucoup d'élèves et qu'il faut toujours insister pour qu'ils aillent le chercher [E5]; certains vont même jusqu'à dire qu'ils en ont peur! En effet, les élèves ont souvent de la difficulté à trouver le mot qu'ils veulent lorsque ce dernier ne correspond pas à l'orthographe qu'ils avaient en tête.

Y'en a qui ont de la misère des fois, par exemple. Je pense à « harceler » qu'elle a écrit avec un « a », bien c'est clair qu'elle l'a pas trouvé dans le dictionnaire, qu'elle l'a pas trouvé dans le Bescherelle. Elle a dû se demander : « Bon, y'existe pas ce mot-

là ». Mais ça, y'ont de la difficulté et souvent ils viennent me demander. « Ben mettons que c'est le son « o », est-ce que c'est obligé d'être un « o »? Comment est-ce que tu pourrais l'écrire autrement? Ben, cherche là-dedans! » Mais ça les décourage. [E7]

Certains élèves disent même aux enseignants que le dictionnaire en format papier est un outil désuet, que *c'est plus la mode* [E2]; beaucoup aimeraient pouvoir utiliser un dictionnaire électronique de type *Lexibook*, de façon à pouvoir trouver facilement le mot visé même avec l'orthographe incorrecte, ce que bien sûr ne permet pas un dictionnaire papier.

Un enseignant soulève aussi les problèmes qu'ont les élèves à bien distinguer le sens approprié lorsqu'ils consultent un article de dictionnaire.

Ils se découragent souvent, quand ils voient que la définition est longue et qu'y a plusieurs définitions pis... c'est difficile, justement, de les faire travailler là-dessus, de voir laquelle serait la meilleure, selon le contexte. Tsé, souvent y vont lire la première pis dire : « C'est ça qu'ça veut dire ». [E5]

En ce qui concerne les critères sur lesquels s'appuient les élèves pour choisir un dictionnaire (lorsque la classe en dispose de plusieurs), ils demeurent assez simples : les élèves choisissent le plus beau! Un des enseignants mentionne que *Le Multidictionnaire* est bien apprécié des élèves parce qu'il est plus accessible.

C'est beaucoup le Multi, parce que c'est facile pour eux. Ils comprennent, c'est pas trop... [Y a moins d'mots ?] Oui, c'est ça. Pis les pages, c'pas des petits papiers de toilette! [E2]

Dans d'autres classes, *Le Multidictionnaire* est assez peu utilisé, parce qu'il demeure *méconnu* [E4]; les élèves préfèrent souvent *Le Petit Larousse illustré*, qui leur paraît souvent plus attrayant que l'austère *Petit Robert*, avec ses longues définitions et son absence d'images et de couleurs.

Globalement, les enseignants semblent unanimes pour dire que le dictionnaire est un outil indispensable en classe de français. Par contre, certains de nos sujets ne paraissent pas très au fait des différences entre les ouvrages, qu'ils ne mettent par conséquent pas de l'avant auprès de leurs élèves. Quant aux élèves, l'utilisation qu'ils font de ce type d'ouvrages vise principalement les dimensions

orthographiques et sémantiques; ils ne semblent pas conscients de la richesse d'information qu'ils pourraient y trouver.

CONCLUSION

Nous constatons donc que les enseignants ont une conception de la maîtrise de la langue centrée autour de la capacité à écrire sans fautes et que le vocabulaire y occupe une place périphérique, peut-être parce qu'ils ont une perception assez positive du vocabulaire de leurs élèves et que cette dimension de la maîtrise du code ne leur vient pas spontanément en tête lorsqu'ils pensent à leur mission d'enseignants de français. Si leur conception de l'erreur lexicale s'articule principalement autour des problèmes sémantiques et de la répétition – qui traduit selon eux un certain manque (ou un manque certain?) de vocabulaire –, ils ont tout de même souligné l'interaction entre le lexique et la syntaxe dans certains cas.

Malgré le fait qu'ils considèrent comme important l'enseignement du vocabulaire en classe de français, les enseignants s'y consacrent peu, les principales raisons évoquées étant le manque de temps et de matériel pertinent; en ce qui concerne le rapport des enseignants à l'enseignement du vocabulaire, donc, nous pourrions dire que l'axe axiologique n'est pas en harmonie avec l'axe praxéologique.

Si l'on s'intéresse plus précisément à l'évaluation du vocabulaire en production écrite, il ressort de nos entrevues que les grilles utilisées par les enseignants accordent peu de poids à la dimension lexicale du texte. Cette dimension de l'évaluation est par ailleurs perçue comme difficile – à cause du manque de balises permettant de juger de la qualité du vocabulaire d'un texte –, mais s'allège avec l'expérience, alors que les enseignants sont plus à même d'utiliser leur intuition pour qualifier le niveau de vocabulaire d'un élève.

Les annotations des enseignants concernant les erreurs lexicales des élèves sont souvent vagues et peu systématiques. Qui plus est, la dimension lexicale est généralement exclue de tout travail à posteriori sur les textes d'élèves, qui se

concentrent généralement sur les erreurs d'accord et de syntaxe lorsqu'ils doivent retravailler leurs productions.

Pour ce qui est du dictionnaire, outil omniprésent dans les classes, il semble méconnu des enseignants, qui le sous-exploitent, et demeure difficile de consultation pour les élèves, qui s'y perdent souvent à cause de l'orthographe et peinent à en tirer vraiment parti, puisqu'ils en connaissent plus ou moins la structure et font souvent preuve de paresse au moment de le consulter.

CONCLUSION

Nous voici donc arrivé au terme de notre vaste exploration de la question de l'erreur lexicale au secondaire. Ce dernier chapitre nous permettra de rappeler les principaux résultats qui découlent de notre étude et de faire certains liens entre nos deux objectifs de recherche. Nous rappellerons aussi les limites de notre étude et verrons de quelle façon des recherches futures pourraient permettre d'approfondir certaines des dimensions explorées. Nous proposerons finalement, à la lumière des nombreuses données analysées dans cette thèse, des pistes tant théoriques que didactiques qui pourraient, selon nous, améliorer la recherche sur l'analyse d'erreurs lexicales et le traitement des problèmes lexicaux par les enseignants du secondaire.

1. RAPPEL DES PRINCIPAUX RÉSULTATS DE LA THÈSE

Rappelons tout d'abord les deux objectifs de cette thèse avant de résumer les principaux résultats relatifs à chacun.

- 1. Analyser les erreurs lexicales d'élèves francophones de troisième secondaire en production écrite;*
- 2. Décrire le rapport d'enseignants de français du secondaire à l'erreur lexicale en ce qui concerne*
 - leur conception de ce qu'est une erreur de nature lexicale;*
 - leurs pratiques d'évaluation du vocabulaire en production écrite;*
 - leurs modes de rétroaction par rapport à des erreurs lexicales.*

1.1 LES ERREURS LEXICALES EN PRODUCTION ÉCRITE EN 3^e SECONDAIRE

C'est par l'analyse d'un corpus de 300 textes narratifs et descriptifs/explicatifs écrits par des élèves francophones de 3^e secondaire que nous avons pu faire une description des erreurs lexicales commises par ces derniers. Notre analyse d'erreurs comportait trois niveaux : 1) une description linguistique des erreurs à l'aide d'une typologie des erreurs lexicales (voir VII-31), 2) une évaluation de la gravité de chaque erreur à l'aide d'une échelle de gravité (voir VII-3.2) et 3) une explication des erreurs dans les cas pour lesquels nous pouvions poser une hypothèse. La correction des textes a révélé 2720 erreurs lexicales, dont un nombre important d'erreurs d'orthographe lexicale et de flexion (1576), que nous avons préféré mettre de côté pour concentrer notre analyse sur un sous-

corpus de 1144 erreurs plus intéressantes dans une perspective de description des erreurs lexicales.

Le premier niveau d'analyse nous a permis d'observer que les erreurs lexicales les plus fréquentes sont les problèmes sémantiques, dans une proportion de 30 %; la plupart de ces erreurs correspondent à des cas où l'UL utilisée présente un lien de sens avec l'UL visée. Les problèmes d'ordre formel, particulièrement les barbarismes, constituent la deuxième famille d'erreurs lexicales observée (21 %). Environ une erreur sur cinq (21 %) est en lien avec une propriété morphosyntaxique – principalement genre nominal, invariabilité et régime – de l'UL erronée, alors que 17 % des problèmes découlent de l'inadéquation pragmatique de l'UL utilisée et du contexte dans lequel elle apparaît, essentiellement par l'emploi d'un vocabulaire familier. Un peu plus d'une erreur sur dix (11 %) se rapporte à la combinatoire lexicale; il s'agit la plupart du temps de collocations douteuses. Cette analyse linguistique révèle que la moitié des erreurs lexicales du corpus (49 %) concernent la combinatoire restreinte des UL, c'est-à-dire les propriétés qui, au-delà de la forme et du sens, régissent leur comportement morphosyntaxique et leur façon de se combiner avec d'autres UL (voir II-3.6).

Le deuxième axe de notre analyse portait sur la gravité des problèmes observés, que nous avons évaluée sur la base de trois critères : l'acceptabilité linguistique de l'erreur selon trois dictionnaires, l'impact de celle-ci sur la compréhension du message et la perception de son intégration à l'usage. Il ressort de notre analyse que la plupart des erreurs du corpus (68 %) ont été considérées comme étant d'une gravité moyenne, les proportions d'erreurs peu graves (15,5 %) ou plutôt sérieuses (16,5 %) étant très similaires. Les cas d'emploi d'une UL appartenant au registre familier sont ceux qui ont été le plus souvent associés à un faible niveau de gravité, alors que les problèmes sémantiques représentent la grande majorité des erreurs plus sérieuses, leur gravité étant généralement proportionnelle à l'écart sémantique entre l'UL utilisée et le sens visé. Les erreurs liées à la combinatoire morphosyntaxique se concentrent au centre du

spectre de gravité. Ce deuxième axe d'analyse nous a permis de constater qu'il n'est pas toujours aisé d'expliquer des erreurs lexicales à l'aide d'ouvrages de référence, puisque les dictionnaires n'indiquent pas l'ensemble des informations nécessaires à la maîtrise d'une UL ou présentent des divergences dans leurs descriptions. Il ressort aussi de cette analyse que, contrairement à ce que l'on aurait pu attendre selon la littérature consultée dans le cadre conceptuel, à peine 4 % des erreurs lexicales de notre corpus induisent un réel problème de compréhension, la quasi-totalité de ces erreurs étant de nature sémantique; bien sûr, cela a sans doute à voir avec le fait que nous incluons dans notre conception de l'erreur lexicale des problèmes morphosyntaxiques et pragmatiques, qui entravent rarement la compréhension. Pour ce qui est de la perception de l'intégration des erreurs à la langue, les deux tiers des erreurs (64,7 %) ont été perçues comme très déviantes par rapport à l'usage, les anglicismes et les UL de registre familier étant sans contredit celles qui y paraissent le plus profondément intégrées.

Le troisième axe d'analyse visait l'explication des erreurs lexicales observées. Nous n'avons fourni d'hypothèses sur la source des erreurs que lorsque des scénarios d'explication nous venaient naturellement. Plusieurs hypothèses pouvaient être avancées pour un même problème, ce qui a produit un total de 896 hypothèses concernant 793 erreurs. Le scénario d'explication le plus fréquemment proposé (dans 41,4 % des cas) est l'intégration de l'erreur à l'usage; les élèves commettent donc très souvent des erreurs lexicales parce qu'ils reproduisent ce qu'ils entendent dans la langue de tous les jours. La proximité formelle et sémantique entre l'UL erronée et l'UL visée permettent aussi d'expliquer bon nombre d'erreurs (22 % des hypothèses posées) et agissent très souvent en synergie; c'est aussi souvent sur la base d'une proximité de sens ou de forme que les élèves transposent le régime d'une UL à une autre, ou lui accolent un collocatif erroné. Évidemment, un certain nombre d'erreurs peuvent simplement être expliquées par une lacune dans le vocabulaire de l'élève ou une maîtrise partielle de son sens (9,7 % de nos hypothèses). Certains

procédés syntaxiques, comme la coordination ou la pronominalisation, sont propices à l'apparition d'erreurs lexicosyntaxiques, tout comme l'application abusive de règles de dérivation morphologique peut être à l'origine d'erreurs formelles. Évidemment, dans certains cas, c'est plutôt la distraction qui nous semblait l'hypothèse la plus probable. De façon globale, nous retenons de ce volet de l'analyse l'influence de la langue orale et l'analogie comme principales sources d'erreurs lexicales.

1.2 LE RAPPORT À L'ERREUR LEXICALE D'ENSEIGNANTS DE FRANÇAIS DU SECONDAIRE

Le corpus de textes de notre premier objectif de recherche comportant 224 rédactions corrigées et annotées au moment du recueil de données, nous avons pu observer comment les enseignants de notre échantillon traitent l'erreur lexicale au moment de la correction.

Il ressort de ces observations que les enseignants relèvent principalement les erreurs d'orthographe d'usage et les erreurs lexicales de type morphosyntaxique lors de la correction; contrairement aux autres types d'erreurs lexicales, celles-ci sont comptabilisées quantitativement et sont généralement associées aux classes d'erreurs « usage », « syntaxe » et « grammaire » des grilles de correction. Moins de 40 % des erreurs plus purement lexicales de notre corpus (problèmes de collocations, erreurs sémantiques et emploi d'UL appartenant au registre familier) ont été relevées par les enseignants, et très peu de ces erreurs ont été pénalisées, sans doute parce que ces problèmes font généralement partie de l'appréciation qualitative qui permet d'évaluer le vocabulaire d'une production écrite. Nous avons remarqué que la répétition est très souvent relevée sur les copies des élèves, mais que les erreurs des trois types mentionnés plus haut, lorsqu'elles sont relevées, sont rarement accompagnées d'annotations plus précises que « *voc* ».

Une série de huit entretiens avec des enseignants de français de 3^e secondaire nous a permis de décrire le rapport de ces enseignants à l'erreur lexicale. Les enseignants rencontrés semblent avoir une conception de la langue très liée à

l'orthographe (grammaticale et lexicale), dans laquelle le vocabulaire n'occupe qu'une place périphérique. Leur conception de l'erreur lexicale s'articule autour du sens et de la répétition, qui témoigne selon eux d'un manque de vocabulaire. Plusieurs sujets soulignent les interactions entre lexique et syntaxe ainsi que la difficulté à classer les erreurs de vocabulaire. Une des prémisses de notre travail s'appuyait sur les commentaires des enseignants (rapportés par plusieurs auteurs, dont Dreyfus, 2004) concernant le manque de vocabulaire des élèves. Or, contrairement à ce que nous attendions, les enseignants rencontrés ont une perception plutôt positive du vocabulaire de leurs élèves, lorsque ceux-ci sont en situation d'évaluation.

En ce qui concerne l'enseignement du vocabulaire en classe de français, les enseignants avouent n'y consacrer que bien peu de temps et toujours le travailler par le biais d'autres tâches, la plupart du temps en préparation à une tâche d'écriture (souvent dans le but avoué d'éviter les répétitions) ou alors à travers des lectures. Les quelques activités portant spécifiquement sur le vocabulaire concernent la plupart du temps les préfixes et les suffixes, les organisateurs textuels ou les anglicismes, mais sont souvent considérées comme des activités d'enrichissement. Un petit nombre d'activités plus marginales, visant davantage le développement d'une compétence lexicale qu'un accroissement direct du vocabulaire des élèves, nous ont aussi été rapportées : exploitation de la synonymie comme procédé de reprise, utilisation du contexte pour l'inférence du sens, recherche dans le dictionnaire. Les enseignants expliquent le manque d'attention accordé à l'enseignement lexical par des contraintes de temps, mais aussi par un manque criant d'outils pertinents pour travailler le lexique.

Pour ce qui est de la correction de productions écrites, les grilles utilisées par les enseignants accordent peu de poids au vocabulaire, qui a toujours un impact minime sur le résultat global. Par ailleurs, l'évaluation du vocabulaire représente pour les enseignants une tâche floue et éminemment subjective, qui repose davantage sur leur intuition d'enseignant que sur des critères précis, le seul

observable unanimement mentionné étant la répétition. Les annotations laissées sur les copies d'élèves par les enseignants relativement aux erreurs de vocabulaire sont vagues (« *voc* », cercles, vaguelettes) et peu systématiques. D'ailleurs, l'erreur lexicale fait rarement l'objet d'un suivi après la remise de la rédaction annotée, même lorsque les élèves doivent retravailler leur texte.

Le dictionnaire, outil lexical par excellence, est présent dans toutes les classes, mais de façon plutôt inégale : si certains enseignants disposent d'une variété intéressante de dictionnaires et d'un nombre suffisant d'ouvrages en regard du nombre d'élèves, ce n'est pas le cas pour d'autres, qui disposent de vieux dictionnaires en piteux état. Nous avons été frappé de constater que la plupart des enseignants rencontrés n'ont pas vraiment de préférence quant au dictionnaire qu'ils utilisent, ce qui témoigne selon nous d'une méconnaissance des particularités de chaque ouvrage. Sans grande surprise, nous nous sommes vu confirmer que la consultation du dictionnaire par les élèves demeure fastidieuse; c'est une tâche qu'ils essaient souvent d'éviter. Lorsqu'ils se résignent à s'y astreindre, les jeunes choisissent le dictionnaire le plus beau (ou le plus proche) et l'utilisent presque exclusivement pour vérifier l'orthographe des « mots » ou, en lecture, pour comprendre le sens d'un terme inconnu.

Lorsque nous avons demandé aux enseignants, en fin d'entrevue, de réagir à sept erreurs lexicales issues de notre corpus, nous avons pu observer des traitements assez différents : certains auraient pénalisé des problèmes que d'autres n'auraient pas relevés et la façon de classer les erreurs variait aussi beaucoup d'un sujet à l'autre. Ce dernier volet de l'entrevue nous a aussi permis de constater que les enseignants possèdent certaines intuitions relatives à des phénomènes lexicaux qu'ils n'ont jamais étudiés spécifiquement (notamment la collocation), mais sont parfois un peu dépourvus pour expliquer des erreurs ou pour justifier leur classement des problèmes rencontrés. Nous avons d'ailleurs constaté les mêmes intuitions et les mêmes hésitations dans notre mémoire de maîtrise (Anctil, 2005), lorsque nous avons demandé à des étudiants universitaires issus de diverses facultés (sciences biomédicales, pharmacie,

communication, architecture, adaptation scolaire, science politique, histoire de l'art, anthropologie, études françaises) d'expliquer des erreurs lexicales ciblées présentes dans des phrases. Nous pouvons donc supposer que les enseignants de français du secondaire ne sont pas mieux outillés que toute autre personne possédant un baccalauréat pour comprendre et expliquer des erreurs de nature lexicale, ce qui est sans doute un symptôme de la formation lacunaire que reçoivent les maîtres en formation en matière de lexique (Tremblay, 2009).

2. MISE EN PARALLÈLE DES DEUX OBJECTIFS DE RECHERCHE

Nous avons choisi d'utiliser l'erreur lexicale comme pierre angulaire pour éclairer la question de l'évaluation du vocabulaire au secondaire; cet objet d'étude a été abordé dans notre premier objectif de recherche à travers les erreurs des élèves et, dans le second, par le biais d'entrevues réalisées avec des enseignants. La mise en parallèle des résultats de ces deux volets de notre travail permet de faire émerger certaines dissonances entre les façons d'appréhender l'erreur lexicale du point de vue du pôle « élève » et du pôle « enseignant » du triangle didactique.

2.1 L'ANALYSE DES ERREURS

Tout d'abord, l'erreur sémantique, le type le plus fréquent d'erreurs lexicales observé dans notre corpus, constitue l'un des types d'erreurs les moins relevées par les enseignants. Qui plus est, ces problèmes sont souvent jugés parmi les plus graves; pourtant, le sens demeure assez peu travaillé en classe selon les pratiques déclarées dont nous ont fait part les enseignants, et, toujours selon leurs dires, les élèves peinent à utiliser le dictionnaire pour vérifier le sens des mots parce qu'ils se perdent dans les nombreuses acceptions. Nous avons néanmoins constaté dans nos entrevues que les enseignants expliquent avec assez d'aisance les erreurs sémantiques et qu'ils sauraient sans doute mener d'intéressantes discussions sémantiques avec leurs élèves s'ils prenaient leurs erreurs comme point de départ.

Les erreurs en lien avec les propriétés morphosyntaxiques des UL (régime, genre nominal, invariabilité) représentent quant à elles, avec les problèmes orthographiques, les erreurs lexicales les plus souvent relevées par les enseignants. Par contre, le traitement qu'ils en font ne permet pas de mettre en lumière la source lexicale de ces problèmes, puisqu'ils sont classés parmi les erreurs syntaxiques, au même titre qu'une erreur de construction de phrase, ou parmi celles d'orthographe grammaticale, comme les erreurs d'accord sujet-verbe ou d'accord du participe passé. Ce type de classement, selon nous, ne fournit pas de piste intéressante aux élèves, qui doivent eux-mêmes distinguer les erreurs morphosyntaxiques ayant une source lexicale de celles qui relèvent réellement des règles de grammaire; une telle tâche est beaucoup trop difficile pour un apprenant dont la compréhension du système linguistique est en construction. Par ailleurs, cette façon de classer ce type d'erreurs lexicales a aussi des répercussions sur l'évaluation, puisqu'elles se trouvent ainsi à être comptabilisées et influencent quantitativement le résultat final, contrairement aux erreurs sémantiques, pragmatiques et collocationnelles. Ces erreurs lexicales ont donc davantage de poids dans la note finale de l'élève, alors qu'elles sont la plupart du temps considérées comme moins graves que les erreurs de nature sémantique.

Nous avons fait ressortir, au chapitre X (section 1.2), que les erreurs que nous avons considérées comme les plus sérieuses, à la lumière de notre calcul de gravité, ne correspondent pas nécessairement aux erreurs les plus pénalisées, ni même les plus relevées, par les enseignants. Nous avons donc émis l'hypothèse que pour eux, c'est plutôt le type d'erreur qui influence le traitement qu'ils en font, les erreurs les plus fréquemment pénalisées étant celles qui correspondent aux classes d'erreurs des grilles de correction (syntaxe, grammaire, orthographe lexicale). Or, leurs réactions aux sept erreurs que nous leur avons soumises en cours d'entrevue nous amènent à nuancer cette hypothèse. Il est vrai que dans les 224 copies corrigées, les enseignants ont eu tendance à relever et à pénaliser principalement les erreurs formelles et morphosyntaxiques, mais devant des

problèmes précis, leurs décisions étaient moins catégoriques. En effet, si tous nos sujets auraient pénalisé l'erreur de genre nominal (présentée en X-2.5), et que tous sauf un auraient réservé le même traitement à l'erreur d'invariabilité (X-2.4), nous constatons que le problème de registre de langue (X-2.3), l'erreur de collocation (X-2.7) et celle liée au typage des actants (X-2.2) auraient aussi été pénalisées par la majorité des enseignants (6 sur 8 ou plus). Or, ces trois dernières erreurs relèvent de classes généralement peu relevées ou pénalisées lors de la correction. À l'inverse, l'erreur de préposition régie (X-2.6), qui correspond pourtant à une classe généralement pénalisée, ne l'aurait été que par le quart des sujets. Nous pouvons donc supposer que le traitement que les enseignants réservent à une erreur particulière ne dépend pas seulement de son type, ni de sa gravité telle que nous la concevons dans cette thèse. D'autres critères entrent en jeu dans la norme évaluative des enseignants. Sans doute la conception qu'ils ont de ce que devrait être le vocabulaire d'un élève de 3^e secondaire a-t-elle une influence sur leur décision – c'est d'ailleurs ce que révèlent nos entretiens –, tout comme leur désir de mettre l'accent plus précisément sur les contenus travaillés en classe dans leur correction. Peut-être aussi hésitent-ils à pénaliser ou à pointer des erreurs qu'ils considèrent comme difficiles à expliquer, l'erreur de préposition régie par exemple. Pouvoir observer les enseignants durant une tâche de correction réelle nous aurait sans doute permis de mettre au jour les facteurs qui influencent véritablement leur traitement des erreurs lexicales.

En ce qui concerne les sources d'erreurs possibles, c'est l'influence de la langue orale qui a été l'hypothèse la plus souvent avancée. Or, l'utilisation d'UL de registre familier constitue le type d'erreurs le moins relevé et un seul enseignant a mentionné travailler en classe les registres de langue, et ce, principalement pour faire observer les différences de registres entre un conte écrit ou raconté à l'oral. Aucun de nos sujets n'a abordé la question des marques d'usage dans les dictionnaires, marques pourtant précieuses lorsque vient le temps de statuer sur le caractère familier d'une UL. La proximité sémantique ou formelle entre des UL

est aussi à l'origine de nombreuses erreurs. Encore une fois, les pratiques liées à l'enseignement du vocabulaire ne semblent pas très appropriées pour sensibiliser les élèves à ce type de confusion. La synonymie constitue le type de lien lexical le plus cité par les enseignants, mais les pratiques s'y rattachant visent la plupart du temps la construction d'une banque de mots dans le but d'éviter la répétition; un seul enseignant a dit travailler sur les nuances de sens entre des synonymes. Quant à la proximité formelle, qui correspond au phénomène de paronymie, elle n'a été évoquée par aucun sujet, ce qui nous laisse supposer que le travail sur des UL proches d'un point de vue formel doit se limiter à celui sur les traditionnels homophones.

2.2 LES CONCEPTIONS DES ENSEIGNANTS À PROPOS DU VOCABULAIRE

Les conceptions des enseignants mises en lumière dans le second volet de notre recherche permettent aussi quelques constats. Tout d'abord, le cadre théorique que nous avons retenu, la TST, place le sens au centre de sa description de la langue et accorde par conséquent une place centrale au lexique. Il est intéressant de constater que les enseignants ont quant à eux une vision très morphosyntaxique de la langue dans laquelle le vocabulaire occupe une place périphérique, vision sans doute très conditionnée par le programme de formation, les grilles d'évaluation des productions écrites, et le combat quotidien qu'ils doivent mener contre les erreurs d'orthographe. Cette façon d'appréhender la langue explique peut-être en partie pourquoi l'erreur lexicale est reléguée au second plan (voire au dernier plan!) dans la correction.

Comme nous l'avons mentionné plus haut, la perception plutôt positive qu'ont les enseignants interviewés du vocabulaire de leurs élèves tranche avec l'idée selon laquelle les élèves manquent de vocabulaire, idée corroborée par plusieurs auteurs. Sans aller jusqu'à affirmer que les élèves manquent réellement de vocabulaire – nous ne disposons pas des calculs de richesse et de variété lexicales qui nous permettraient une telle affirmation –, nous aimerions pointer de nouveau la sous-utilisation de locutions par les élèves, constat qui nous a frappé lors de l'analyse du corpus. Les enseignants n'ont fait aucun commentaire

à cet effet. Au cours des dernières années, nous avons pourtant rencontré nombre d'enseignants de primaire, dans le cadre d'ateliers que nous animions sur l'enseignement des expressions, qui pointaient eux aussi la pauvreté du vocabulaire de leurs élèves relativement aux expressions, certains mentionnant même que leurs élèves ne connaissent pas beaucoup d'expressions courantes qu'ils utilisent en classe.

2.3 LES PRATIQUES D'ENSEIGNEMENT DU VOCABULAIRE

Le constat principal lorsque nous observons les pratiques d'enseignement du vocabulaire que nous ont révélées les enseignants, outre le fait qu'elles sont peu nombreuses, est qu'une véritable didactique du lexique n'a pas encore fait son entrée dans les classes du secondaire. Le vocabulaire n'est jamais travaillé de façon structurée, les activités lexicales sont presque toujours subordonnées à des tâches d'écriture ou de lecture et visent la plupart du temps l'accroissement de la taille du vocabulaire des élèves par l'acquisition de mots nouveaux (banques de synonymes, définition de mots inconnus). Rares sont les activités portant plutôt sur l'acquisition de savoirs, d'habiletés ou de stratégies visant à renforcer la compétence lexicale des élèves dans le but de structurer leur vocabulaire et de leur fournir les outils nécessaires à l'acquisition efficace de termes nouveaux.

De plus, les activités dont nous ont fait part les enseignants ne sont pas directement en lien avec les erreurs lexicales observées dans notre corpus : le sens est rarement travaillé, tout comme l'emploi du dictionnaire, qui pourrait pourtant devenir un allié important des élèves pour éviter certaines erreurs. Cette ressource est à notre avis sous-estimée.

À la décharge des enseignants, il faut tout de même rappeler qu'ils sont souvent bousculés par des programmes surchargés, qu'ils n'ont pas reçu eux-mêmes de formation adéquate en matière de lexique et que les approches prônées par les didacticiens du lexique ne sont pas encore largement répandues et demeurent souvent inconnues des praticiens.

2.4 LES ENSEIGNANTS ET L'ERREUR DE VOCABULAIRE

Comme nous l'avons mentionné au chapitre X (section 3.1.2), il y a un certain recoupement entre notre conception de l'erreur lexicale et celle des enseignants, qui mentionnent souvent les problèmes de sens et, dans des proportions moindres, les erreurs de registres et les erreurs lexicosyntaxiques.

Par contre, ce qui nous frappe, c'est que l'erreur lexicale, pierre angulaire de notre thèse, n'occupe pas une place centrale dans l'évaluation du vocabulaire. En effet, la répétition, dont nous n'avons pas tenu compte dans notre travail, est citée par l'ensemble des enseignants comme un critère déterminant dans l'évaluation du vocabulaire d'un élève. L'erreur lexicale, d'après ce que nous constatons, sert surtout à moduler l'appréciation globale du vocabulaire que l'enseignant donne sur la base de son expérience; ainsi, si un élève utilise un vocabulaire que l'enseignant juge bon pour son niveau, mais qu'il commet deux ou trois erreurs sémantiques, il n'obtiendra pas le score parfait au critère « vocabulaire ». Ce constat nous rappelle qu'il serait réducteur de centrer l'évaluation du vocabulaire en production écrite autour de l'erreur lexicale.

En ce qui a trait aux explications fournies par les enseignants à propos d'erreurs lexicales, nous avons constaté qu'elles reposent en grande partie sur des intuitions et que certaines erreurs, notamment celles de collocation, semblaient plus difficiles à expliquer. Or, le cadre conceptuel étoffé que nous avons dû déployer pour être en mesure de bien décrire les erreurs lexicales de notre corpus démontre à quel point l'analyse de problèmes lexicaux demande la maîtrise de plusieurs notions : polysémie, sens, régime, collocation, registres de langue, locution, homonymie, paronymie, etc. Pour que les enseignants en viennent à fournir des explications à leurs élèves qui ne reposent pas que sur l'intuition et à proposer des interventions didactiques davantage en phase avec les erreurs observées, il faudrait, comme le mentionne Tremblay (2009), leur assurer une solide formation en sémantique et en lexicologie.

3. LIMITES DE LA THÈSE ET RECHERCHES FUTURES

Une recherche qui vise à défricher une question aussi nouvelle et vaste que celle de l'erreur lexicale au secondaire ne peut aspirer à en couvrir tous les aspects. Les choix méthodologiques que nous avons opérés (choix d'un corpus, méthodes d'analyse, moyens d'accès aux représentations, nombre de sujets) ont pour corolaire certaines limites, qui ont déjà été présentées à la fin de nos chapitres de résultats (VIII et X). Nous rappellerons ici les principales et verrons de quelle façon elles ouvrent la porte à des recherches futures.

3.1 ANALYSE D'UN CORPUS D'ERREURS D'ÉLÈVES ALLOPHONES

Nous n'avons retenu dans notre corpus que des textes produits par des élèves ayant été scolarisés en français depuis la maternelle et le parlant à la maison; la très grande majorité de ces élèves n'utilisent que le français dans la vie de tous les jours. Par conséquent, les erreurs analysées ont été produites par des élèves francophones, pour la plupart unilingues. Bien entendu, ce public n'est pas nécessairement représentatif de celle que l'on retrouve dans les classes du secondaire, particulièrement dans la région de Montréal.

Une recherche future pourrait analyser les erreurs lexicales commises par les élèves allophones présents dans les classes régulières du secondaire pour décrire en quoi elles diffèrent de celles des francophones et vérifier si les pistes didactiques proposées dans la présente thèse seraient pertinentes pour ce public. Bien entendu, pour qu'une telle recherche produise des résultats intéressants, il faudrait à notre avis qu'elle porte sur un vaste corpus qui distinguerait différents groupes selon la langue maternelle et le degré de maîtrise du français. En effet, il se pourrait fort bien que le profil linguistique des apprenants influence les erreurs commises, comme nous l'avons vu lors du survol historique des différents courants d'analyse d'erreurs (IV-2.3).

3.2 ANALYSE D'UN CORPUS PLUS VASTE PAR PLUSIEURS CODEURS

L'ensemble du codage a ici été effectué en entier par l'auteur de la thèse. Un contre-codage a permis de vérifier la validité du codage dans une certaine mesure, mais toutes les données produites ont tout de même été analysées par une seule personne, avec la part de subjectivité que cela implique.

Avec davantage de moyens, il aurait été intéressant de faire analyser un corpus plus vaste par plusieurs codeurs. Ainsi, nous aurions pu opérer un contre-codage systématique pour tous les niveaux d'analyse, dont celui de description linguistique à l'aide de notre typologie des erreurs lexicales. Le fait de disposer d'un nombre plus élevé de rédactions nous aurait aussi permis de prendre en compte d'autres paramètres dans notre analyse, pour vérifier par exemple l'impact du type de texte, de la consigne d'écriture ou du degré de compétence en français écrit sur les erreurs commises. Un corpus plus vaste permettrait aussi d'envisager une comparaison des erreurs à différents stades des études secondaires, par une analyse de textes produits par des élèves d'âge différent.

Dans l'idéal, il serait aussi intéressant de pouvoir analyser l'ensemble des erreurs linguistiques du corpus retenu. Ceci nous permettrait d'abord de quantifier l'importance des erreurs lexicales relativement aux autres types d'erreurs linguistiques. Nous avons pu observer qu'elles sont beaucoup moins fréquentes que les erreurs d'orthographe grammaticale, mais il nous est impossible d'établir dans quelle proportion. Un codage de ce type nous permettrait aussi de voir quel pourcentage des erreurs syntaxiques et grammaticales est en fait attribuable à un manque de connaissances lexicales (régime, genre nominal, invariabilité, etc.), ce qui nous apporterait sans doute un argument supplémentaire quant à l'importance de distinguer ces classes d'erreurs dans la correction de productions écrites.

3.3 PRISE EN COMPTE DE LA RICHESSE ET DE LA VARIÉTÉ DU VOCABULAIRE

Comme nous l'avons mentionné plus haut, les enseignants ne considèrent pas que l'erreur lexicale dans leur appréciation du vocabulaire utilisé dans un texte.

Nous croyons impératif que nos prochains travaux sur l'évaluation du vocabulaire prennent en considération des composantes de la richesse lexicale des textes d'élèves autres que la seule présence d'erreurs – par exemple la sophistication, la variation et la densité lexicales (Read, 2000) –, puisque ces paramètres peuvent influencer les enseignants relativement à l'évaluation du vocabulaire.

Plusieurs chercheurs ont déjà proposé des méthodes de calcul de richesse lexicale (voir notamment Arnaud, 1984 ; Reid, 1990 ; Engber, 1995 ; Read, 2000). Certains outils ont aussi été développés pour calculer de façon automatique la richesse lexicale d'un texte; nous pensons notamment au *VocabProfil* développé par Tom Cobb¹⁹⁴, dont une version expérimentale existe pour le français. Il serait intéressant d'opposer des calculs « scientifiques » de richesse lexicale de textes à l'appréciation du vocabulaire que font des enseignants de ces mêmes textes. Daller et Phelan (2007) ont mené une telle étude auprès d'enseignants d'anglais langue étrangère et il ressort de leur analyse que le jugement des enseignants est surtout influencé par la présence de vocabulaire avancé ou de mots rares dans le texte de l'élève. Ainsi, ce serait la sophistication lexicale qui aurait le plus d'impact sur l'appréciation du vocabulaire par les enseignants; les auteurs posent l'hypothèse que cela s'explique peut-être par le fait que la présence de mots rares est un moyen économique de juger le vocabulaire et d'alléger la tâche d'évaluation. Selon les propos des enseignants que nous avons rencontrés, c'est surtout la variété du vocabulaire (révélée selon eux par l'absence de répétitions) qu'ils considèrent lors de la correction. Il serait donc intéressant de mener une étude similaire à celle de Daller et Phelan (2007), mais auprès d'enseignants de français langue première.

3.4 NOUVELLE APPROCHE DE LA GRAVITÉ DES ERREURS

Comme nous l'avons maintes fois souligné, l'évaluation de la gravité des erreurs demeure très subjective. Nous avons tenté de diminuer cette part de subjectivité

¹⁹⁴ <http://www.lexutor.ca/vp/>

en opérationnalisant le concept de gravité en trois critères, mais, comme nous l'avons expliqué à la section VIII-3.4, ce choix méthodologique s'est avéré plus ou moins heureux.

À postériori, nous croyons que le problème pourrait être abordé de façon différente. Plutôt que d'opérationnaliser dès le départ le concept de gravité en critères, nous pensons maintenant qu'il serait sans doute plus judicieux de poser d'abord un jugement de gravité global sur chaque problème, puis de tenter par la suite de comprendre les éléments ayant influencé ce jugement global.

Si nous avons à retravailler sur la gravité des erreurs, notre méthodologie serait différente. Tout d'abord, nous ferions appel à plusieurs sujets pour poser des jugements de gravité globaux (par exemple à l'aide d'une échelle de Likert) de façon à avoir plusieurs appréciations pour chaque erreur; la plus grande faiblesse méthodologique de ce second niveau de notre analyse d'erreurs était sans contredit d'être le seul locuteur à poser des jugements sur les problèmes relevés, à part pour la petite quantité d'erreurs soumises au contre-codage.¹⁹⁵ Il serait aussi pertinent que les sujets posant les jugements de gravité soient des enseignants, de façon à obtenir un portrait plus représentatif de ce qui se passe au moment de la correction. Nous leur demanderions aussi de justifier leurs jugements de gravité, ce qui ferait émerger les critères sur lesquels ils reposent et nous permettrait éventuellement d'opérationnaliser de façon plus juste le concept de gravité, en considérant le poids relatif de chacun de ces facteurs.

3.5 IMPLICATION DES ENSEIGNANTS ET DES ÉLÈVES POUR L'EXPLICATION DES ERREURS

Pour notre troisième niveau d'analyse, nous sommes encore une fois le seul à avoir avancé des hypothèses pour expliquer les erreurs lexicales de notre corpus (sauf pour les erreurs du contre-codage, voir VIII-4.1). Sans doute aurait-il été intéressant de faire appel aux enseignants pour fournir des scénarios d'explication, eux qui connaissent bien leurs élèves et sont habitués à rencontrer

¹⁹⁵ Rappelons tout de même que le contre-codage a produit un accord interjuges assez concluant (voir IX-3.1).

ces types d'erreurs. Peut-être une telle approche aurait-elle permis de faire apparaître de nouvelles hypothèses, ou du moins de mettre en lumière davantage de sources possibles pour certaines erreurs.

Nous pourrions aussi approfondir la dimension de l'explication des erreurs en parlant directement avec les élèves qui en sont les auteurs. Ainsi, par leurs verbalisations, nous pourrions établir avec plus de certitude la source des erreurs en tranchant entre différentes hypothèses, ou simplement en découvrant des sources que nous n'avions pas envisagées.

3.6 OBSERVATION DES PRATIQUES RÉELLES DES ENSEIGNANTS LIÉES AU VOCABULAIRE

Nous l'avons souligné à plusieurs reprises, mais réaffirmons une fois de plus le caractère exploratoire du second volet de notre recherche : la description du rapport à l'erreur lexicale d'enseignants de secondaire. Les limites liées au nombre restreint de sujets (8) et à l'absence de méthodes d'observation dans notre méthodologie exigent une certaine prudence quant au potentiel de généralisation de nos observations.

Une recherche visant à approfondir la question devrait à notre avis être élargie au rapport des enseignants au vocabulaire et à son enseignement. Travailler avec un échantillon diversifié d'enseignants de français du secondaire (voire du primaire) permettrait de dresser un portrait beaucoup plus complet des pratiques d'enseignement du vocabulaire. Des observations en classe nous assureraient que les pratiques déclarées en entrevues constituent effectivement les pratiques réelles des enseignants; pouvoir observer la façon dont les élèves participent à ces activités ajouterait aussi un éclairage supplémentaire à notre description et permettrait peut-être de faire ressortir les pratiques effectives.

En ce qui concerne l'évaluation du vocabulaire en production écrite, il serait intéressant d'observer des enseignants en cours de tâche de correction, à travers des entretiens d'explicitation. Ainsi, ils auraient à s'exprimer à propos non pas d'un éventail limité d'erreurs lexicales imposées, mais des problèmes réels de leurs propres élèves. Une telle méthodologie permettrait aussi de voir quelles

erreurs attirent l'attention des enseignants en cours de correction, ce que l'étude de copies déjà annotées n'éclairait que partiellement; impossible en effet de savoir si un item linguistique, lorsqu'il n'est associé à aucun commentaire, suscite ou non une réflexion chez le correcteur. Ainsi, nous pourrions mieux comprendre pourquoi les enseignants ne relèvent pas certaines erreurs, notamment sémantiques. Est-ce parce qu'ils jugent qu'il n'est pas pertinent de relever ces problèmes en fonction du niveau de leurs élèves, parce qu'ils ne veulent pas pénaliser des erreurs qu'ils auraient de la difficulté à expliquer ou simplement parce qu'ils ne les perçoivent pas comme des erreurs?

Des observations sur la façon dont les enseignants et les élèves consultent le dictionnaire, ainsi qu'un approfondissement de leurs représentations relatives à ce type d'ouvrages, pourraient sans doute donner des pistes utiles qui orienteraient la façon d'aborder cet outil en classe de français.

4. APPORTS DE LA THÈSE, PISTES ET RECOMMANDATIONS

Le volume important de données analysées dans cette thèse nous a permis une foule d'observations à la lumière desquelles nous pouvons aujourd'hui formuler un certain nombre de recommandations, que nous adresserons à quatre groupes de personnes : les lexicographes, les enseignants de français, les formateurs de futurs enseignants et les didacticiens du lexique.

4.1 RECOMMANDATIONS À L'INTENTION DES LEXICOGRAPHES

Comme nous l'avons mentionné lors de la présentation de nos résultats relatifs à la gravité des erreurs (VIII-3.3.1.1), notre consultation quasi compulsive des trois dictionnaires retenus dans notre étude nous a permis de constater certains manques dans ces ouvrages. Voici quelques pistes que nous pourrions proposer aux lexicographes, notamment ceux qui développent des ouvrages dans une perspective pédagogique :

1. **revoir la façon de décrire le régime** : Comme nous l'avons souligné lors de l'analyse des erreurs lexicosyntaxiques de notre corpus (VIII-2.2.4.2), les dictionnaires présentent des lacunes importantes dans la façon dont

ils décrivent le régime des UL. Il faudrait développer une méthode pour rendre compte de façon plus complète, accessible et cohérente des structures syntaxiques contrôlées par les UL, en particulier les verbes. Voir apparaître clairement le nombre de compléments associés à un verbe et les prépositions qui peuvent éventuellement les introduire serait à notre avis d'une grande utilité pour les enseignants, les élèves, voire pour les utilisateurs « naïfs » des dictionnaires, pour qui les notions de transitivité et d'intransitivité ne correspondent peut-être pas à grand-chose. Il ne faut pas sous-estimer non plus la valeur des exemples pour illustrer le régime des UL (voir point suivant). La *Lexicologie explicative et combinatoire* a développé une méthode permettant d'indiquer clairement et de façon exhaustive les structures régies par les UL qu'elle décrit; peut-être les lexicographes gagneraient-ils à s'en inspirer?

2. **fournir davantage d'exemples d'utilisation** : Nous avons pu mesurer l'importance des exemples pour trouver des réponses à nos questions tout au long de notre codage, particulièrement en ce qui concerne le régime, et plus spécifiquement les prépositions régies, et la collocation. Il serait donc pertinent que les rédacteurs de dictionnaires prêtent davantage d'attention à la sélection des exemples, de façon à ce qu'ils reflètent les différentes constructions possibles pour chaque UL et qu'ils présentent les principaux collocatifs qui y sont associés (ceux correspondant aux FL Oper, Func, Magn et Real). Structurer les exemples contenant les collocatifs selon une certaine logique, comme le fait le logiciel *Antidote RX*, aiderait aussi les utilisateurs à retrouver plus rapidement l'information à l'intérieur de l'article. Nous croyons aussi que les ouvrages de consultation grand public devraient délaisser les grands classiques de la littérature pour fournir des exemples représentatifs de l'usage moderne.
3. **offrir une meilleure description des locutions** : Un travail considérable sur les locutions mériterait d'être fait, dans le but d'en proposer un traitement plus égal dans les dictionnaires. En effet, les informations qu'on y retrouve relativement aux locutions varient d'un dictionnaire à l'autre, et même à l'intérieur d'un même ouvrage. Généralement, le sens des locutions n'est pas décrit de façon très explicite, sans compter que les différentes acceptions de ce type d'UL sont rarement distinguées.¹⁹⁶ Quant aux renseignements relatifs à l'utilisation des locutions, qui permettraient de mieux saisir leur fonctionnement souvent capricieux (régime, possibilité de passiver l'expression, degré de figement), ils sont à peu près absents; même les exemples d'utilisation se font plutôt rares.

¹⁹⁶ Par exemple, la locution PIED DE NEZ est décrite dans nos trois ouvrages de référence comme un type de grimace, mais aucun ne fait mention de son acception métaphorique (ex. *Le Carnaval de Québec : un pied de nez à l'hiver.*), selon nous plus fréquente si l'on se fie aux premières dizaines d'exemples que retourne le moteur de recherche *Google*.

4. **élargir la notion de paronymie** : Nous avons vu à la section VIII-4.2.3 que la parenté formelle entre des UL est souvent à l'origine d'erreurs lexicales, mais que les remarques des dictionnaires relatives aux paronymes concernent rarement les UL en jeu dans ces erreurs. Peut-être serait-il intéressant de disposer d'ouvrages présentant des séries de mots partageant des caractéristiques formelles (ex. ASSOUIR, ASSOULIR, ASSOUDIR, ASSOUVIR), à partir desquels les enseignants pourraient faire des exercices avec les élèves.

4.2 RECOMMANDATIONS À L'INTENTION DES ENSEIGNANTS

La présente recherche nous a permis d'élucider en partie le rapport des enseignants du français du secondaire à l'erreur lexicale et, un peu plus largement, à l'enseignement du vocabulaire. Voici les suggestions que nous pourrions leur faire à la lumière des constats réalisés.

4.2.1 EN LIEN AVEC L'ENSEIGNEMENT DU LEXIQUE

Nous avons vu que les enseignants, par manque de temps et d'outils, accordent peu d'importance à l'enseignement du vocabulaire. Nous avons aussi constaté que le lexique est souvent abordé par le biais d'autres tâches et vise généralement l'accroissement du stock lexical des élèves plutôt qu'une meilleure compréhension du système lexical et le développement d'une réelle compétence lexicale. En attendant que les didacticiens du lexique les convainquent de l'importance de modifier leurs pratiques d'enseignement en lien avec le lexique, voici quelques suggestions que nous pourrions faire aux enseignants de français du secondaire.

1. **encourager l'exploration et l'utilisation du dictionnaire** : Le dictionnaire demeure la meilleure ressource pour explorer le lexique et éviter l'erreur lexicale. Afin d'encourager les élèves à l'utiliser, l'enseignant doit lui-même bien connaître la structure de ce type d'ouvrage, les types d'informations qu'il contient et les particularités des différents dictionnaires. Une classe devrait (dans l'idéal!) disposer d'une variété de dictionnaires répondant à différents besoins : dictionnaires encyclopédiques du type *Petit Larousse*, dictionnaires de définitions comme *Le Petit Robert*, dictionnaires orientés vers les difficultés d'écriture comme *Le Multidictionnaire*, dictionnaires électroniques comme *Antidote*, dictionnaires de synonymes, de locutions, dictionnaire visuel, thésaurus... et pourquoi pas un *Lexibook* pour encourager les élèves pour lesquels l'orthographe pose particulièrement un problème!

Le rôle de l'enseignant consiste à amener les élèves à se familiariser avec le dictionnaire et à apprendre à le consulter avec aisance. Pour ce faire, les activités en lien avec le dictionnaire ne devraient pas être centrées essentiellement sur l'orthographe et le sens des mots, mais viser aussi la recherche d'informations relevant d'autres dimensions : invariabilité, genre, régime, collocation, locution, registres de langue. Dans cette optique, l'enseignant devrait aussi travailler les marques d'usages et les abréviations présentes dans les ouvrages qu'utilisent les élèves, de façon que ceux-ci en tirent davantage profit.

2. **travailler le sens** : Les erreurs sémantiques étant les plus fréquentes, il serait pertinent que les enseignants travaillent davantage l'aspect sémantique en classe. Ce travail devrait se faire notamment par la comparaison d'UL synonymes à l'aide du dictionnaire dans le but de faire ressortir les nuances de sens qui les distinguent et d'aiguiser l'acuité sémantique des élèves en développant leur désir du mot juste. La mise en lumière de liens sémantiques permettant des regroupements dans les banques de mots construites en préparation à une activité d'écriture constitue une autre façon de travailler le sens.
3. **travailler les registres de langue** : Les erreurs liées à l'emploi d'UL appartenant au registre familier sont fréquentes, et l'influence de la langue orale permet aussi d'expliquer plusieurs erreurs de différents types. Ces constats justifient à notre avis un travail plus systématique sur les registres de langue avec les élèves. Il faut amener les élèves à douter stratégiquement de certains termes qui peuvent appartenir au registre familier; bien entendu, c'est la maîtrise des marques d'usage relatives à cette dimension qui permettra aux élèves de trouver réponse à leurs questions dans le dictionnaire.
4. **utiliser la notion de régime verbal comme outil d'analyse syntaxique** : Voilà plutôt une piste d'enseignement qu'une réelle recommandation. Mais selon nos expériences d'enseignement de la grammaire, se servir des connaissances intuitives qu'ont les étudiants de la structure actancielle du verbe principal d'une phrase leur permet souvent de repérer plus facilement les grands constituants de la phrase qui s'articulent autour de ce verbe.

4.2.2 EN LIEN AVEC LE TRAITEMENT DE L'ERREUR LEXICALE

Notre travail portait précisément sur l'erreur lexicale en production écrite. Voici quelques recommandations en lien avec cette dimension plus spécifique de l'évaluation du lexique. Les recommandations 5 et 6 concernent les interventions à faire en amont d'une production écrite, les numéros 7 à 11 visent plus précisément l'étape de correction, alors que nos dernières suggestions sont en lien avec les interventions après la correction.

5. **fournir aux élèves des outils lexicaux** : En plus du dictionnaire auquel ont droit les élèves durant la rédaction, il pourrait être pertinent de leur fournir des outils maison se rapportant davantage à leurs difficultés lexicales. De tels outils, outre les traditionnelles listes de mots liés à un thème, pourraient prendre la forme de listes compilées par l'enseignant qui présenteraient les anglicismes, les pléonasmes ou les termes familiers souvent présents dans les productions des élèves.
6. **sensibiliser les élèves aux contextes syntaxiques « dangereux »** : Nous avons observé que des types d'erreurs lexicales, principalement lexicosyntaxiques, pouvaient être favorisés par certains contextes syntaxiques. Sensibiliser les élèves au « danger » de la coordination de verbes et à la pronominalisation pourrait une fois de plus les aider à douter stratégiquement lorsqu'ils écrivent.
7. **relever ou pénaliser tout emploi allant à l'encontre de ce que l'on trouve dans les dictionnaires ou divergent par rapport à l'usage** : Toutes les erreurs ne sont pas aussi graves, nous l'avons démontré. Néanmoins, certaines erreurs méritent d'être pointées systématiquement à l'élève : celles qui vont à l'encontre de ce que présentent les ouvrages et celles qui paraissent très peu naturelles par rapport à l'usage. Nous croyons aussi que les erreurs qui entraînent un problème de compréhension devraient être signalées à l'élève systématiquement.
8. **fournir une proposition de correction dans les cas où l'élève ne peut y arriver seul** : Certaines erreurs lexicales, notamment les barbarismes ou les formes analytiques, ne peuvent être corrigées aisément par l'élève, puisqu'il est impossible pour lui d'accéder à la forme voulue à l'aide du dictionnaire. Ainsi, dans de tels cas, il est important de proposer une reformulation à l'élève.
9. **fournir des corrections minimales** : Dans le cas où l'enseignant décide de proposer une correction, il devrait tenter le plus possible de laisser intact le contexte dans lequel l'erreur apparaît afin de ne pas dénaturer le texte de l'élève.
10. **proposer des pistes de correction lorsque l'élève peut par lui-même « réparer » son erreur** : Plusieurs erreurs lexicales peuvent être corrigées par l'élève lui-même s'il se donne la peine de consulter un dictionnaire. Dans le but d'orienter sa recherche, l'enseignant devrait fournir des pistes de correction dans ses annotations. Par exemple, dans le cas d'un problème sémantique, il pourrait écrire « *Es-tu certain que le sens de ce mot est approprié dans cette phrase?* » ou, dans le cas d'une erreur de régime, « *Bon choix de verbe, mais va vérifier dans le dictionnaire avec quelle préposition il doit être utilisé.* ». De tels commentaires permettraient à l'élève d'améliorer par lui-même son texte (Roberge (2006) les nomme d'ailleurs « commentaires mélioratifs »), et l'effort cognitif fourni pour y arriver risque de mener à une meilleure mémorisation des informations en jeu.

11. **développer des codes de correction plus précis relativement au vocabulaire** : Nous l'avons constaté, les annotations relatives au vocabulaire sont vagues et peu systématiques. L'utilisation d'un code d'erreurs plus précis, qui permettrait au moins de distinguer les problèmes sémantiques des erreurs de registre, conduirait sans doute les élèves à tirer davantage profit des commentaires de l'enseignant.

En effet, dans la mesure où l'enseignant décide de relever une erreur sur la copie d'un élève, il est de son devoir de s'assurer que l'élève dispose d'assez d'informations pour savoir pourquoi l'item en question a été relevé. Il faudrait selon nous établir avec les élèves un système d'annotations explicite qui ne laissera pas l'élève perplexe devant une petite vaguelette sous un mot-forme ou les lettres *voc.* écrites dans la marge. Nous croyons que la correction constitue un dialogue entre l'élève et l'enseignant, et la langue employée pour assurer ce dialogue doit être comprise des deux côtés et présenter un certain niveau de précision (Roberge, 2006).

12. **inclure les erreurs lexicales dans le suivi lorsque les élèves doivent retravailler leur texte** : Le second volet de notre recherche a révélé que, lorsque les élèves doivent expliquer leurs erreurs à la suite de la remise d'une rédaction corrigée, les erreurs lexicales sont souvent exclues. Afin d'habituer les élèves à effectuer par eux-mêmes des recherches sur les mots en utilisant les dictionnaires, nous croyons que les erreurs lexicales devraient aussi être incluses dans ce type de suivi.
13. **considérer le sens lors d'interventions concernant les erreurs formelles** : Lorsque l'élève commet une erreur qui semble découler de la parenté formelle entre deux UL, l'enseignant devrait fournir une rétroaction qui ne met pas l'accent que sur la forme, mais qui met aussi en évidence les différences sémantiques entre les UL en jeu.
14. **distinguer les erreurs morphosyntaxiques découlant d'un manque de connaissances lexicales de celles qui relèvent réellement de règles morphosyntaxiques** : Comme les enseignants classent généralement les erreurs lexicales de types morphosyntaxiques parmi les erreurs de grammaire ou de syntaxe, il serait important que leurs interventions didactiques fassent bien ressortir la différence entre les deux types d'erreurs. En effet, la façon de corriger ces erreurs sera radicalement différente selon qu'elles proviennent de la méconnaissance d'une propriété lexicale – ce qui demande la consultation d'un dictionnaire – ou qu'elles sont des erreurs morphosyntaxiques réelles – qui demandent plutôt de faire appel à une règle d'accord ou de consulter une grammaire.

4.3 RECOMMANDATIONS À L'INTENTION DES FORMATEURS D'ENSEIGNANTS

Nos entretiens nous ont permis de constater que l'explication d'erreurs lexicales par les enseignants repose en grande partie sur l'intuition, parce qu'ils possèdent des connaissances insuffisantes en sémantique et lexicologie. Voici donc quelques recommandations à l'endroit des facultés des sciences de l'éducation qui forment les futurs maîtres.

1. **offrir une formation en didactique du lexique** : Comme le souligne Tremblay (2009), la formation que reçoivent les futurs enseignants de français au secondaire en matière de lexique est souvent insuffisante. Il serait primordial de fournir aux futurs maîtres les outils conceptuels nécessaires à l'explication de phénomènes linguistiques aussi complexes que ceux analysés dans notre thèse. Par ailleurs, si cette formation nécessaire en linguistique et en sémantique/ lexicologie est dispensée par des départements de linguistique, il serait important que les cours soient adaptés à aux étudiants en éducation et que des ponts soient faits avec des situations concrètes d'enseignement/ apprentissage. Les prochaines recommandations concernent les objectifs que devraient se fixer de tels cours en regard du traitement de l'erreur lexicale.
 - a. **sensibiliser les futurs enseignants aux différentes facettes de la connaissance d'une UL** : La maîtrise d'un mot comprend un nombre important de connaissances de nature diverse. Un cours de lexicologie/sémantique destiné à de futurs enseignants devrait procurer aux étudiants une meilleure compréhension du système lexical de la langue, et accorder une place particulière au phénomène de la collocation, qui semble inconnu des enseignants que nous avons rencontrés.
 - b. **initier les futurs enseignants à un emploi non naïf des dictionnaires** dans le but de leur permettre d'enseigner efficacement à leurs élèves à retrouver dans ce type d'ouvrages l'information recherchée.
 - c. **amener les futurs enseignants à prendre conscience des interactions entre le lexique et la morphosyntaxe** afin de leur permettre de poser de meilleurs diagnostics sur les erreurs de leurs élèves.

4.4 RECOMMANDATIONS À L'INTENTION DES DIDACTIENS DU LEXIQUE

Bien entendu, la présente thèse fait aussi ressortir quelques recommandations qui nous concernent plus directement, en tant que didacticien du lexique.

1. **reposer des outils aux enseignants :** Nos sujets nous ont mentionné le manque criant d'outils pertinents pour travailler le lexique en classe de français. Il est du devoir des didacticiens de proposer aux enseignants des outils stimulants et efficaces qui permettraient à une véritable didactique du lexique de faire son entrée dans les classes du secondaire. Ces outils pourraient prendre la forme de cahiers d'activités qui toucheraient l'ensemble des propriétés des UL et pourraient entre autres les aborder à partir d'erreurs d'élèves. p
2. **convaincre les enseignants de la pertinence de travailler le lexique :** Nous le savons, les enseignants travaillent peu le vocabulaire en classe. Nous croyons pourtant qu'un travail plus systématique sur le vocabulaire permettrait aux élèves d'acquérir certaines habiletés métalinguistiques et certaines connaissances linguistiques qui pourraient leur être utiles dans leur quête de la maîtrise de la langue écrite. À nous maintenant de convaincre les enseignants qu'une sensibilisation au phénomène de la collocation, la compréhension du concept de régime syntaxique en lien avec la structure actancielle ou une connaissance plus profonde de la façon dont on construit un dictionnaire pourraient aider leurs élèves à mieux écrire.
3. **proposer des interventions didactiques en lien avec les erreurs lexicales :** Nous avons avancé tout au long de cette thèse des pistes d'interventions pour les enseignants en lien avec les erreurs lexicales des élèves. Notre prochaine tâche serait de consigner ces pistes dans un outil d'intervention accessible que les enseignants pourraient utiliser comme guide lorsqu'ils font face à des problèmes de nature lexicale. Un tel outil pourrait notamment présenter une version allégée de notre typologie des erreurs lexicales et mettre en lien les types d'erreurs avec leurs sources possibles et des propositions d'interventions.

EN GUISE DE CONCLUSION

Nous croyons, par nos réflexions théoriques sur l'erreur, nos propositions d'outils d'analyse, nos certaines d'analyses linguistiques et nos rencontres avec des enseignants de français, avoir contribué à éclairer un champ d'études jusqu'ici peu exploré : celui de l'erreur lexicale en langue première. Par cette

recherche, nous apportons aussi notre pierre à l'édifice de la didactique du lexique.

En comprenant mieux quels sont les types d'erreurs commises par les élèves et de quelle façon les enseignants y réagissent, nous sommes mieux outillé pour pouvoir délaissier les recherches purement descriptives et nous tourner maintenant vers des travaux qui proposeront des pistes concrètes aux enseignants en lien avec le traitement de l'erreur lexicale et, plus largement, avec l'enseignement du vocabulaire.

Bien entendu, nous nous sommes concentré ici uniquement sur l'erreur lexicale en production écrite, qui ne constitue qu'une facette de l'évaluation du vocabulaire. Il serait maintenant intéressant de commencer à réfléchir à la façon d'élargir les pratiques d'évaluation du vocabulaire, pour qu'elles permettent non pas de vérifier seulement la maîtrise des unités lexicales utilisées spontanément par l'élève à l'écrit, mais pour qu'elles deviennent un véritable baromètre de la compétence lexicale des élèves.

BIBLIOGRAPHIE

- ABEILLÉ, A. et GODARD, D. (1999). « La place de l'adjectif épithète en français : le poids des mots ». *Recherches Linguistiques*, vol. 28, 9-31.
- ABRIC, J.-C. (1994a). « Les représentations sociales : aspects théoriques ». Dans J.-C. Abric (Ed.), *Pratiques sociales et représentations* (Vol. 11-36). Paris: Presses Universitaires de France.
- ABRIC, J.-C. (1994b). « Méthodologie de recueil des représentations sociales ». Dans J.-C. Abric (Ed.), *Pratiques sociales et représentations*. Paris: Presses Universitaires de France, Coll. Psychologie sociale, 59-82.
- ABRIC, J.-C. (2003). « La recherche du noyau central et de la zone muette des représentations sociales ». Dans J.-C. Abric (Ed.), *Méthodes d'étude des représentations sociales*. Ramonville Saint-Agne: Érès, 59-80.
- ADAM, J.-M. (2008). *La linguistique textuelle : introduction à l'analyse textuelle des discours* (2^e éd.). Paris: Armand Colin, Coll. Collection Cursus, Linguistique.
- AGUSTÍN LLACH, M. P. (2004). « Lexical inconsistencies and deficiencies in English second language learners: proposal of classification ». *Interlingüística*, vol. 15, no 1, 83-92.
- AGUSTÍN LLACH, M. P. (2005a). « A critical review of the terminology and taxonomies used in the literature on lexical errors ». *Miscelánea : a journal of english and american studies*, vol. 31, 11-24.
- AGUSTÍN LLACH, M. P. (2005b). « Responding to different composition topics: a quantitative analysis of lexical error production ». *Glosas didacticas, Revista electrónica internacional (ISSN 1576-7809)*, no 13, 128-140.
- AGUSTÍN LLACH, M. P. (2005c). « The Relationship of Lexical Error and their Types to the Quality of ESL Compositions: an Empirical Study ». *Porta Linguarum*, vol. 3, 45-57.
- ANCTIL, D. (2005). *Maîtrise du lexique chez les étudiants universitaires : typologie des problèmes lexicaux et analyse des stratégies de résolution de problèmes lexicaux*. « Mémoire de maîtrise non publié », Université de Montréal, Montréal.
- ANCTIL, D. (2008). « Caractérisation de la notion d'erreur lexicale et description des problèmes lexicaux d'étudiants universitaires ». Dans *Autour des langues et du langage : perspectives multidisciplinaires, Colloque international des étudiants-chercheurs en linguistique et didactique des langues, Grenoble, 4-7 juillet 2006*. Grenoble: Presses universitaires de Grenoble, 349-356.
- APOSTOLIDIS, T. (2003). « Représentations sociales et triangulation : enjeux théorico-méthodologiques ». Dans J.-C. Abric (Ed.), *Méthodes d'études des représentations sociales*. Ramonville Saint-Agne: Érès, 13-35.
- ARABSKI, J. (1979). *Errors as indicators of the development of interlanguage*. Katowice, Pologne: Uniwersytet Slaski.
- ARNAUD, P. J. L. (1984). « The Lexical Richness of L2 Written Productions and the Validity of Vocabulary Tests ». Dans T. Culhane, C. Klein-Braley & D. K. Stevenson (Eds.), *Practice and Problems in Language Testing*. Colchester: University of Essex, 14-28.
- ARNAUD, P. J. L. (1992). « Objective lexical and grammatical characteristics of L2 written compositions and the validity of separate components tests ». Dans P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics*. London: Macmillan.

- ASTOLFI, J.-P. (1997). *L'erreur, un outil pour enseigner*. Issy-les-Moulineaux: ESF Éditeur, Coll. Pratiques et enjeux pédagogiques.
- BAILLY, D. (1998). *Les mots de la didactique des langues : Le cas de l'anglais (lexique)*. Paris: Ophrys.
- BARRÉ-DE MINAC, C. (2000). *Le rapport à l'écriture : aspects théoriques et didactiques*. Villeneuve d'Ascq, France: Presses universitaires du Septentrion, Coll. Savoirs mieux.
- BARRÉ-DE MINAC, C. (2007). « Écrire pour apprendre : où le rapport à l'écriture est convoqué ». Dans É. Falardeau, C. Fisher, C. Simard & N. Sorin (Eds.), *La didactique du français. Les voies actuelles de la recherche*. Lévis: Les Presses de l'Université Laval, 165-180.
- BARRÉ-DE MINAC, C. (2008). « Le rapport à l'écriture : une notion à valeur euristique ». Dans S.-G. Chartrand & C. Blaser (Eds.), *La rapport à l'écrit : un outil pour enseigner de l'école à l'université*. Namur: Presses universitaires de Namur, 11-23.
- BASTUJI, J. (1978). « Les théories sur le vocabulaire, éléments pour une synthèse ». *Pratiques*, no 20, 75-89.
- BEAUCHESNE, J. (2002). *Dictionnaire des cooccurrences*. Montréal: Guérin, Coll. Au service de l'écriture.
- BECK, I. L. et MCKEOWN, M. G. (1991). « Conditions of vocabulary acquisition ». Dans R. Barr, M. Camail, P. Mosenthal & P. D. Pearson (Eds.), *The Handbook of Reading Research* (Vol. II). New York: Longman, 789-814.
- BEEFUN, H. (2001). « Attitude face aux erreurs dans l'enseignement / apprentissage du français langue étrangère ». *Polifonia (Lisboa, Edições Colibri)*, no 4, 37-51.
- BÉGUELIN, M.-J. (2000). *De la phrase aux énoncés : grammaire scolaire et descriptions linguistiques*. Bruxelles: De Boeck Duculot, Coll. Savoirs en pratique.
- BERGER, C. (2006). « Démarche anthropologique et entretien compréhensif : approcher la réalité de l'Autre ». Dans A. Cain & G. Zarate (Eds.), *L'entretien : ses apports à la didactique des langues*. Paris: Le Manuscrit, 17-27.
- BERRENDONNER, A. (1988). « Normes et variations ». Dans G. Shoeni, J.-P. Bronckart & P. Perrenoud (Eds.), *La langue française est-elle gouvernable? Normes et activités langagières*. Neuchâtel, Paris: Delachaux & Niestlé, 43-62.
- BERTHOUD, A.-C. (1982). « Sur la relative fiabilité du discours métalinguistique des apprenants ». *Encrages*, vol. 8-9, 139-142.
- BERTHOUD, A.-C. (1987). « Les erreurs des apprenants : Au panier ou sous le microscope? ». *Les langues modernes : « Les erreurs des élèves : qu'en faire? »*, vol. LXXXI, no 5, 11-17.
- BERTHOUD, A.-C. et PY, B. (1993). *Des linguistes et des enseignants*. Berne: Peter Lang, Coll. Explorations, série Recherches en sciences de l'éducation.
- BERTRAND, Y. (1987). « Faute ou erreur? Erreur et faute ». *Les langues modernes : « Les erreurs des élèves : qu'en faire? »*, vol. LXXXI, no 5, 70-80.
- BERTUCCI, M.-M. (1997). « L'analyse des erreurs : trois théories ». *Travaux de didactique du français langue étrangère*, vol. 37, 63-88.
- BESSE, H. et PORQUIER, R. (1984). *Grammaire et didactique des langues*. Paris: Hatier.
- BILLIEZ, J. et MILLET, A. (2004). « Représentations sociales : trajets théoriques et méthodologiques ». Dans D. Moore (Ed.), *Les représentations des langues et de*

- leur apprentissage. *Références, modèles, données et méthodes*. Paris: Didier, 31-49.
- BLASER, C. (2007). *Fonction épistémique de l'écrit : pratiques et conceptions d'enseignants de sciences et d'histoire du secondaire*. « Thèse de doctorat non publiée », Université Laval, Québec.
- BOIVIN, M.-C., PINSONNEAULT, R. et PHILIPPE, M.-É. (2003). *Bien écrire : la grammaire revue au fil des textes littéraires* (2e éd.). Laval, Québec: Beauchemin.
- BRENT, E. (1999). « Vers l'élaboration de normes pédagogiques du français ». *Terminogramme. La norme du français au Québec. Perspectives pédagogiques*, no 91-92, 117-130.
- BRODEUR, M., DEAUDELIN, C., BOURNOT-TRITES, M., SIEGEL, L. S. et DUBÉ, C. (2003). « Croyances et pratiques d'enseignants de la maternelle au sujet des habiletés métaphonologiques et de la connaissance des lettres ». *Revue des sciences de l'éducation*, vol. XXIX, no 1, 171-194.
- BULMAN, F. (2003). *Le prépositionnaire : dictionnaire des verbes et des adjectifs pouvant être suivis d'une préposition*. Québec: L'Instant même.
- BUREAU, C. (1985). *Le français écrit au secondaire : une enquête et ses implications pédagogiques*. Québec: Conseil de la langue française, Coll. Documentation du Conseil de la langue française.
- BURT, M. K. et KIPARSKY, C. (1972). *The gooficon : a repair manual for English*. Rowley, Massachusetts: Newbury House.
- CAIN, A. (1987). « Les erreurs des élèves, qu'en faire? ou le triomphe de la raison sur la passion ». *Les langues modernes : « Les erreurs des élèves : qu'en faire? »*, vol. LXXXI, no 5, 7-9.
- CALAQUE, É. et DAVID, J. (2004). *Didactique du lexique : contextes, démarches, supports* (1^{re} éd.). Bruxelles: De Boeck Université, Coll. Savoirs en pratique.
- CALAQUE, É. et GROSSMANN, F. (2000). « Enseignement/apprentissage du lexique ». *Lidil*, no 21.
- CATACH, N., GRUAZ, C.-R. et DUPREZ, D. (1986). *L'orthographe française : traité théorique et pratique avec des travaux d'application et leurs corrigés*. Paris: Nathan, Coll. Université Nathan, information, formation, Linguistique française.
- CAUCHON, J. (2003). *Enseignement de la structure prédicative du verbe au primaire : étude expérimentale à partir du verbe DÉCIDER*. « Mémoire de maîtrise non publié », Université de Montréal, Montréal.
- CAVALLI, M., DUCHÊNE, A., ELMIGER, D., GAJO, L., MARQUILLÓ LARRUY, M., MATTHEY, M., PY, B. et SERRA, C. (2004). « Le bilinguisme : représentations sociales, discours et contextes ». Dans D. Moore (Ed.), *Les représentations des langues et de leur apprentissage. Références, modèles, données et méthodes*. Paris: Didier, Coll. Essais CRÉDIF, 65-99.
- CHARLOT, B. (1997). *Du rapport au savoir. Éléments pour une théorie*. Paris: Anthropos.
- CHARLOT, B. (2003). « La problématique du rapport au savoir ». Dans S. Maury & M. Caillot (Eds.), *Rapport au savoir et didactiques*. Paris: Fabert, Coll. Éducation et sciences.
- CHARTRAND, S.-G. et BLASER, C. (2008). « Du rapport à l'écriture au concept didactique de capacités langagières : apports et limites de la notion de rapport à l'écrit ». Dans

- S.-G. Chartrand & C. Blaser (Eds.), *La rapport à l'écrit : un outil pour enseigner de l'école à l'université*. Namur: Presses universitaires de Namur, 107-127.
- CHARTRAND, S.-G., BLASER, C. et GAGNON, M. (2006). « Fonction épistémique de l'écrit et genres disciplinaires. Enquête dans les classes d'histoire et de sciences du secondaire québécois ». *Revue suisse des sciences de l'éducation*, vol. 28, no 2, 275-293.
- CHAUDENSON, R., MOUGEON, R. et BÉNIAC, É. (1993). *Vers une approche panlectale de la variation du français*. Paris: Diffusion Didier-Érudition.
- CHOMSKY, N. (1965). *Aspects of the theory of syntax*. Cambridge: M.I.T. Press.
- CHOMSKY, N. (1971). *Aspects de la théorie syntaxique*. Paris: Le Seuil (trad. de Chomsky, 1965).
- CLÉMENT, P. (1994). « Représentations, conceptions, connaissances ». Dans A. Giordan, Y. Girault & P. Clément (Eds.), *Conceptions et connaissances*. Berne: Peter Lang, Coll. Exploration. Cours et contributions pour les sciences de l'éducation.
- COQUIN-VIENNOT, D. et GAONAC'H, D. (1995). « Psychologie et didactique : les notions fondamentales ». Dans D. Gaonac'h & C. Golder (Eds.), *Profession enseignant. Manuel de psychologie pour l'enseignement*. Paris: Hachette éducation.
- CORBEIL, J.-C. (1980). *Les choix linguistiques*. Communication présentée au colloque « La qualité de la langue... après la loi 101 » (Québec, du 30 septembre au 3 octobre 1979), Québec : Conseil de la langue française.
- CORDER, P. (1967/1980). « Que signifient les erreurs des apprenants? ». *Langages*, vol. 57, 9-15.
- CORDER, S. P. (1974). « Error analysis ». Dans J. Allen & S. P. Corder (Eds.), *The Edinburgh Course in Applied Linguistics* (Vol. 3). London: Oxford University Press.
- CORDER, S. P. (1981). *Error Analysis and Interlanguage*. Oxford: Oxford University Press.
- COSERIU, E. (1973). « Sistema, norma y habla ». Dans E. Coseriu (Ed.), *Teoría del lenguaje y lingüística general : cinco estudio* (3rd ed.). Madrid: Gredos, 113-177.
- DABÈNE, M. (1987). *L'adulte et l'écriture. Contribution à une didactique de l'écrit en langue maternelle*. Bruxelles: De Boeck.
- DABÈNE, M. (2008). « Écritures, écrits : un monde contrasté de pratiques et de représentations ». Dans S.-G. Chartrand & C. Blaser (Eds.), *La rapport à l'écrit : un outil pour enseigner de l'école à l'université*. Namur: Presses universitaires de Namur, 7-10.
- DALLAIRE, É. (1997). *Évolution du discours institutionnel tenu sur la langue, la norme et la grammaire depuis le Rapport Parent*. Université du Québec à Montréal, Montréal.
- DALLER, H. et PHELAN, D. (2007). « What is in a teacher's mind? Teachers ratings of EFL essays and different aspects of lexical richness ». Dans H. Daller, J. Milton & J. Treffers-Daller (Eds.), *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press, Coll. Cambridge Applied Linguistics, 234-244.
- DAVID, J. et GROSSMANN, F. (2003). « Présentation du dossier "enseigner-apprendre le lexique" ». *La Lettre de l'AIRDF (Association internationale pour la recherche en didactique du français)*, vol. 2, no 33, 3-5.
- DE ROSA, A. S. (2003). « Le "réseau d'associations" : une technique pour détecter la structure, les contenus, les indices de polarité, de neutralité et de stéréotypie du champ sémantique liés aux représentations sociales ». Dans J.-C. Abric (Ed.),

- Méthodes d'étude des représentations sociales*. Ramonville Saint-Agne: Érès, 81-117.
- DE VILLERS, M.-É. (2003). *Multidictionnaire de la langue française* (4^e éd.). Montréal: Québec Amérique.
- DE VILLERS, M.-É. (2005). *Le vif désir de durer : illustration de la norme réelle du français québécois*. Montréal: Québec Amérique.
- DEBYSER, F. (1971). « La linguistique contrastive et les interférences ». *Langue française*, no 8, 31-62.
- DEBYSER, F., HOUIS, M. et ROJAS, C. (1967). *Grille de classement typologique des fautes*. Paris: Bureau pour l'enseignement de la langue et de la civilisation françaises à l'étranger (BELC).
- DELCAMBRE, I. et REUTER, Y. (2002). « Images du scripteur et rapports à l'écriture ». *Pratiques*, no 113/114, 7-28.
- DELOMIER, D. (1980). « La place de l'adjectif en français, bilan des points de vue et théories du XX^e siècle ». *Cahiers de lexicologie*, vol. 37, 5-24.
- DEMAIZIÈRE, F. (1987). « Erreurs et enseignement assisté par ordinateur ». *Les langues modernes : «Les erreurs des élèves : qu'en faire?»*, vol. LXXXI, no 5, 48-53.
- DENANCE, M. (1999). « L'évaluation des fautes de vocabulaire à l'évaluation de l'épreuve uniforme ». *Correspondance (périodique en ligne)*, vol. 4, no 4.
- DOISE, W. (1989). « Attitudes et représentations sociales ». Dans D. Jodelet (Ed.), *Les représentations sociales*. Paris: Presses universitaires de France, Coll. Sociologie d'aujourd'hui, 240-258.
- DREYFUS, M. (2004). *Pratiques et représentations de l'enseignement du lexique à l'école primaire*. Communication présentée au 9^e colloque de l'AIRDF, Québec, 26 au 28 août 2004.
- DUBOIS, J. (1994). *Dictionnaire de linguistique et des sciences du langage*. Paris: Larousse, Coll. Trésors du français.
- DUCROT, O. et SCHAEFFER, J.-M. (1995). *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris: Éditions du Seuil.
- DULAY, H., BURT, M. et KRASHEN, S. (1982). *Language Two*. New York: Oxford University Press.
- DURKHEIM, É. (1898). « Représentations individuelles et représentations collectives ». *Revue de métaphysique et de morale*, vol. VI, Paris : Hachette, 273-302.
- DURKIN, D. (1979). « What classroom observations reveal about reading comprehension instruction ». *Reading Research Quarterly*, vol. 14, 481-533.
- ELLIS, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press, Coll. Oxford Applied Linguistics.
- ENGBER, C. A. (1995). « The Relationship of Lexical Proficiency to the Quality of ESL Compositions ». *Journal of Second Language Writing*, vol. 4, no 2, 139-155.
- ENGH, B. (1971). « En toleransundersköning : tyska elever tolkar svenska elevers språkfel ». *Pedagogisk-psykologiska problem (Lärarhögskolan : Malmö)*, no 137.
- ENKVIST, N. E. (1973). « Should we count errors or measure success? ». Dans J. Svartvik (Ed.), *Errata*. Lund: CWK Gleerup, 16-23.
- FISHER, C. (2004). « La place des représentations des apprenants en didactique de la grammaire ». Dans C. Vargas (Ed.), *Langue et études de la langue. Approches*

- linguistiques et didactiques*. Aix-en-Provence: Publications de l'Université d'Aix-en-Provence, 383-393.
- FORTIER, G. (1994). « Propos sur le vocabulaire ». *Québec français*, no 92, 24-27.
- FREI, H. (1929). *La grammaire des fautes*. Paris: Gauthner.
- GADET, F. (1989). *Le français ordinaire*. Paris: Armand Colin.
- GAJO, L. et SERRA, C. (2000). « Les représentations de l'apprentissage des langues et du bilinguisme dans l'institution éducative ». *Études de linguistique appliquée*, vol. 120, 487-508.
- GAONAC'H, D. (1982). « La notion d'interlangue et la psychologie cognitive du langage ». Dans B. Py (Ed.), « *Acquisition d'une langue étrangère III* », *Actes du colloque organisé les 16-18 septembre 1982 à l'Université de Neuchâtel*. Saint-Denis: Les presses de l'Université Paris VIII - Vincennes, 63-79.
- GARCIA-DEBANC, C. et MAS, M. (1987). « Évaluation des productions écrites des élèves ». *Enjeux*, no 11, 108-122.
- GARCIA-DEBANC, C. et PLANE, S. (2004). *Comment enseigner l'oral à l'école primaire?* Paris: Hatier, Coll. Hatier pédagogie.
- GASS, S. M. et SELINKER, L. (2001). *Second language acquisition: An introductory course* (2^e éd.). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- GILLY, M. (1989). « Les représentations sociales dans le champ éducatif ». Dans D. Jodelet (Ed.), *Les représentations sociales*. Paris: Presses Universitaires de France, Coll. Sociologie d'aujourd'hui, 383-406.
- GIORDAN, A. et DE VECCHI, G. (1987). *Les origines du savoir. Des conceptions des apprenants aux concepts scientifiques*. Neuchâtel, Paris: Delachaux et Niestlé.
- GRANGER, S. (2003). « Error-tagged Learner Corpora and CALL: A Promising Synergy ». *CALICO*, vol. 20, no 3, 465-480.
- GRANGER, S. et MONFORT, G. (1994). « La description de la compétence lexicale en langue étrangère : perspectives méthodologiques ». *AILE*, vol. 3, 55-75.
- GRAWITZ, M. (2001). *Méthodes des sciences sociales* (11^e éd.). Paris: Dalloz, Coll. Précis. Droit public, science politique.
- GREVISSE, M. (1977). *Quelle préposition?* Paris-Gembloux: Duculot, Coll. Les Guides du bon usage.
- GREVISSE, M. et GOOSSE, A. (2008). *Le bon usage* (14^e éd.). Bruxelles: De Boeck ; Duculot.
- GROSSMANN, F., PAVEAU, M.-A. et PETIT, G. (2005). *Didactique du lexique : langue, cognitions, discours*. Grenoble: ELLUG, Université Stendhal.
- GROUPE DIEPE. (1995). *Savoir écrire au secondaire : étude comparative auprès de quatre populations francophones d'Europe et d'Amérique / Groupe DIEPE*. Bruxelles: De Boeck Université, Coll. Pédagogies en développement. Série 1, Problématiques et recherches.
- GUEUNIER, N. (1989). « Corriger des copies: objectifs, langue et interaction ». *Le français aujourd'hui*, no 86, 91-99.
- GUILBERT, L. (1972). « Peut-on définir un concept de norme lexicale? ». *Langue française*, no 16, 29-48.
- HALTÉ, J.-F. (1984). « L'annotation des copies, variété ou base du dialogue pédagogique ». *Pratiques*, no 44, 61-69.

- HALTÉ, J.-F. (1992). *La didactique du français* (1re éd.). Paris: Presses Universitaires de France, Coll. Que sais-je? (2656).
- HAMEL, M.-J. et MILIĆEVIĆ, J. (2007). « Analyse d'erreurs lexicales d'apprenants du FLS : démarche empirique pour l'élaboration d'un dictionnaire d'apprentissage. ». *Revue de l'Association canadienne de linguistique appliquée (ACLA)*, vol. 10, no 1.
- HAUSMANN, F. J. (1979). « Un dictionnaire des collocations est-il possible ? ». *Travaux de littérature et de linguistique de l'Université de Strasbourg*, vol. XVII, no 1, 187-195.
- HENDRICKSON, J. M. (1978). « Error Correction in Foreign Language Teaching: Recent Theory, Research and Practice ». *The Modern Language Journal*, vol. LXII, no 8, 387-398.
- HENRIKSEN, B. (1999). « Three dimensions of vocabulary development ». *Studies in Second Language Acquisition*, vol. 21, 303-317.
- HJELMSLEV, L. (1971). « Langue et parole ». Dans L. Hjelmslev (Ed.), *Essais linguistiques*. Paris: Editions de Minuit. (article paru d'abord en 1943).
- HONVEAU, R. (1995). « Variation linguistique et norme : variables et erreurs graphiques ». *Liaison HESO, De la variation : écarts, erreurs, normes*, no 25-26, 53-64.
- HOTOPF, W. H. N. (1983). « Lexical slips of the pen and tongue: What they tell us about language production ». Dans B. Butterworth (Ed.), *Language production* (Vol. 2). London: Academic Press, 147-200.
- HUET, S. (1997). *Méthodes et contenus pour l'enrichissement du vocabulaire à l'école élémentaire (cycle III)*. Grenoble: C.R.D.P. de l'Académie de Grenoble, Coll. École.
- HYLTENSTAM, K. (1988). « Lexical Characteristics of Near-Native Second-Language Learners of Swedish ». *Journal of Multilingual and Multicultural Development*, vol. 9, no 1-2, 67-84.
- JAMES, C. (1998). *Errors in language learning and use : exploring error analysis*. London: Longman, Coll. Applied Linguistics and Language Studies.
- JAOUI, H. (1979). *Manuel de créativité pratique*. Paris: Éditions de l'Épi.
- JENKINS, J. R. et DIXON, R. (1983). « Vocabulary learning ». *Contemporary Educational Psychology*, vol. 8, 237-260.
- JODELET, D. (1989). « Représentations sociales : un domaine en expansion ». Dans D. Jodelet (Ed.), *Les représentations sociales*. Paris: Presses Universitaires de France, Coll. Sociologie d'aujourd'hui, 47-78.
- JOHANSSON, S. (1975). *Papers in contrastive linguistics and language testing*. Lund Studies in English 50. Lund: CWK Gleerup.
- JOHANSSON, S. (1978). *Studies of error gravity : native reactions to errors produced by Swedish learners of English*. Göteborg: Acta Universitatis Gothoburgensis, Coll. Gothenburg studies in English (44).
- JUHÁSZ, J. (1970). *Probleme der Interferenz*. München: Max Hueber Verlag.
- KAUFMANN, J.-C. (1996). *L'entretien compréhensif*. Paris: Nathan université.
- KOLDE, G. (1981). *Sprachkontakte in gemischtsprachigen Städten. Vergleichende Untersuchung über Voraussetzungen une Formen sprachlicher Interaktion verschiedensprachiger Jugendlicher in den schweizer Städten Biel/Bienne und Fribourg/Freiburg i. U.* Wiesbaden: Franz Steiner.

- L'ÉCUYER, R. (1990). *Méthodologie de l'analyse développementale de contenu : méthode GPS et concept de soi*. Sillery: Presses de l'Université du Québec.
- LABOV, W. (1972). *Sociolinguistic patterns*. Philadelphie: University of Pennsylvania Press.
- LAFONT-TERRANOVA, J. et COLIN, D. (2004). *L'écriture en français et dans deux autres disciplines : représentations d'enseignants*. Communication présentée au 9^e colloque de l'AIRDF, Québec, 26 au 28 août 2004.
- LAFONTAINE, D. (1986). *Le parti pris des mots. Normes et attitudes linguistiques*. Bruxelles: Pierre Mardaga, Coll. Psychologie et sciences humaines (156).
- LAMY, A. (1976). « Pédagogie de la faute ou de l'acceptabilité ». *Études de linguistique appliquée*, vol. 22, 118-127.
- LAMY, A. (1984). « Mes rendez-vous avec la faute : Lettre... ouverte ». *Le français dans le monde*, vol. 185, 77-83.
- LAUFER, B. (1990). « Why Are Some Words More Difficult than Others? Some Intralexical Factors that Affect the Learning of Words ». *IRAL*, vol. 28, no 4, 293-307.
- LAUFER, B. (1992). « Native language effect on confusion of similar lexical forms ». Dans C. Mair & M. Markus (Eds.), *New Departures in Contrastive Linguistics: Innsbrucker Beiträge zur Kulturwissenschaft, Vol. 2*. Innsbruck: University of Innsbruck, 199-209.
- LAUFER, B. et NATION, P. (1995). « Vocabulary size and use: Lexical richness in L2 written production ». *Applied Linguistics*, vol. 16, no 3, 307-322.
- LAUFER, B. et PARIBAKHT, T. S. (1998). « The Relationship Between Passive and Active Vocabularies: Effects of Language Learning Context ». *Language Learning*, vol. 48, no 3, 365-391.
- LAVOIE, N., LÉVESQUE, J.-Y. et CHÉNARD, M. (2004). *Conceptions d'enseignantes du premier cycle du primaire relativement à l'enseignement et à l'apprentissage de l'écriture*. Communication présentée au 9^e colloque de l'AIRDF, Québec, 26 au 28 août 2004.
- LE MANCHEC, C. et BARRÉ-DE MINAC, C. (2004). *Les pratiques dites « pluridisciplinaires » : représentations et conceptions d'enseignants d'école primaire en formation initiale*. Communication présentée au 9^e colloque de l'AIRDF, Québec, 26 au 28 août 2004.
- LEEMAN-BOUIX, D. (1994). *Les fautes de français existent-elles?* Paris: Éditions du Seuil.
- LEFRANÇOIS, P., LAURIER, M., LAZURE, R. et CLAING, R. (2005). *Évaluation de l'efficacité des mesures visant l'amélioration du français écrit du primaire à l'université (tiré à part pour le secondaire)*. Montréal: Université de Montréal, Collège Ahuntsic.
- LEFRANÇOIS, P. et MONTESINOS-GELET, I. (2004). *L'évolution des représentations des futurs enseignants du primaire sur la langue et son enseignement*. Communication présentée au 9^e colloque de l'AIRDF, Québec, 26 au 28 août 2004.
- LEGENDRE, R. (2005). *Dictionnaire actuel de l'éducation* (3^e éd.). Montréal: Guérin.
- LEGENHAUSEN, L. (1975). *Fehleranalyse und Fehlerbewertung. Untersuchungen an englischen Reifeprüfungsnachzählungen*. Berlin: Cornelsen-Velhagen & Klasing.
- LENNON, P. (1991a). « Error : Some Problems of Definition, Identification and Distinction ». *Applied Linguistics*, vol. 12, no 2, 180-196.
- LENNON, P. (1991b). « Error and the very advanced learner ». *International Review of Applied Linguistics*, vol. 29, no 1, 31-44.

- LÉPINE, F. (1995). « Bilan des tests de français à l'admission aux universités québécoises (1987-1994) ». *Revue des sciences de l'éducation*, vol. 21, no 1, 17-36.
- LESSARD-HÉBERT, M., GOYETTE, G. et BOUTIN, G. (1996). *Recherche qualitative : fondements et pratiques* (2^e éd.). Montréal: Éditions Nouvelles, Coll. Éducation.
- LEVENSTON, E. A. et BLUM, S. (1976). « Aspects of Lexical Simplification in the Speech and Writing of Advanced Adult Learners ». Dans P. Corder & E. Roulet (Eds.), *The Notions of Simplification, Interlanguages and Pidgins and Their Relation to Second Language Pedagogy : Actes du 5^e colloque de linguistique appliquée de Neuchâtel, 20-22 mai 1976*. Neuchâtel: Faculté des lettres.
- LOTT, D. (1983). « Analysing and counteracting interference errors ». *English Language Teaching Journal*, vol. 37, 256-261.
- LÜDI, G. et PY, B. (2002). *Être bilingue* (2^e éd.). Berne: Peter Lang, Coll. Exploration. Recherches en sciences de l'éducation.
- MAISONNEUVE, H. (1997). « Une entrevue avec Jean-Denis Moffet, responsable de l'élaboration et de la validation de l'épreuve uniforme de français ». *Correspondance*, vol. 2, no 2.
- MAJEAU, L. (1992). *Le français écrit au secondaire : une réévaluation par la segmentation des textes, l'étude de la cohésion et l'analyse des erreurs lexico-graphiques*. « Mémoire de maîtrise non publié », Université de Montréal, Montréal.
- MARQUILLÓ LARRUY, M. (2003). *L'interprétation de l'erreur*. Paris: CLE International, Coll. Didactique des langues étrangères.
- MARTEL, P. et CAJOLET-LAGANIÈRE, H. (1996). *Le français québécois : usages, standard et aménagement*. Québec: Institut québécois de recherche sur la culture, Coll. Diagnostic (22).
- MARTIN, M. (1984). « Advanced Vocabulary Teaching : The Problem of Synonyms ». *The Modern Language Journal*, vol. 68, no ii, 130-137.
- MASTIN, M. (2001). *L'erreur au service de l'apprentissage*. Centre IUFM de Douai, Douai, France.
- MCCRETTON, E. et RIDER, N. (1993). « Error Gravity and Error Hierarchies ». *Iral*, vol. 31, no 3, 177-188.
- MEARA, P. (1984). « The Study of Lexis in Interlanguage ». Dans A. Davies, C. Criper & A. P. R. Howatt (Eds.), *Interlanguage*. Edinburgh: Edinburgh University Press, 225-239.
- MEARA, P. et BUXTON, B. (1987). « An alternative to multiple choice vocabulary tests ». *Language Testing*, vol. 4, 142-154.
- MEL'ČUK, I. (1992). « Paraphrase et lexique : la théorie Sens-Texte et le « Dictionnaire explicatif et combinatoire » ». Dans I. Mel'čuk, N. Arbatchewsky-Jumarie, L. Iordanskaja, S. Mantha & A. Clas (Eds.), *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques III*. Montréal: Presses de l'Université de Montréal.
- MEL'ČUK, I. (1997). *Vers une linguistique sens-texte : leçon inaugurale faite le vendredi 10 janvier 1997*. Paris, Collège de France: Chaire internationale.
- MEL'ČUK, I., CLAS, A. et POLGUÈRE, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve: Duculot, Coll. Universités francophones.

- MEL'ČUK, I. et POLGUÈRE, A. (2007). *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. Bruxelles: De Boeck & Larcier.
- MEL'ČUK, I., ARBATCHEWSKY-JUMARIE, N., DAGENAI, L., ELNITSKY, L., IORDANSKAJA, L., LEFEBVRE, M.-N. et MANTHA, S. (1988). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques II*. Montréal: Les Presses de l'Université de Montréal.
- MEL'ČUK, I., ARBATCHEWSKY-JUMARIE, N., ELNITSKY, L., IORDANSKAJA, L. et LESSARD, A. (1984). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I*. Montréal: Les Presses de l'Université de Montréal.
- MEL'ČUK, I., ARBATCHEWSKY-JUMARIE, N., IORDANSKAJA, L. et MANTHA, S. (1992). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques III*. Montréal: Les Presses de l'Université de Montréal.
- MEL'ČUK, I., ARBATCHEWSKY-JUMARIE, N., IORDANSKAJA, L., MANTHA, S. et POLGUÈRE, A. (1999). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques IV*. Montréal: Les Presses de l'Université de Montréal.
- MELS. (2003). *Grille d'évaluation de la compétence à écrire. Texte argumentatif, épreuve unique d'écriture, français langue d'enseignement, 5e année du secondaire. Document de présentation*. Québec: Ministère de l'Éducation.
- MIGNE, J. (1969). « Les obstacles épistémologiques et la formation des concepts ». *Éducation permanente*, no 2, 39-66.
- MIGNE, J. (1994). « Pédagogie et représentations ». *Éducation permanente*, vol. 119, no 2, 11-31.
- MILIĆEVIĆ, J. (2006). « A short Guide to the Meaning-Text Linguistic Theory ». *Journal of Koralex*, vol. 8, 187-233.
- MILIĆEVIĆ, J. (2008). « Paraphrase as a Tool for Achieving Lexical Competence in L2 ». Dans S. Vandaele, A. Housen, F. Kulken, M. Pierrard & I. Vedder (Eds.), *Proceedings of the Symposium "Complexity, Accuracy and Fluency in Second Language Learning and Teaching", Bruxelles, March 29-30, 2007*. Bruxelles: Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique, 153-167.
- MILIĆEVIĆ, J. et HAMEL, M.-J. (2007). « Un dictionnaire de reformulation pour apprenants du français langue seconde ». *Revue de l'Université de Moncton, Numéro hors série « Actes des colloques annuels de l'Association de linguistes des provinces atlantiques (APLA) 2005 (Moncton) »*, 145-167.
- MILLET, A., LUCCI, V. et BILLIEZ, J. (1990). *Orthographe mon amour!* Grenoble: Presses Universitaires de Grenoble.
- MINISTÈRE DE L'ÉDUCATION. (2006). *Programme de formation de l'école québécoise. Enseignement secondaire, premier cycle*. Québec: Gouvernement du Québec.
- MOORE, D. (2004a). « Les représentations des langues et de leur apprentissage : Itinéraires théoriques et trajets méthodologiques ». Dans D. Moore (Ed.), *Les représentations des langues et de leur apprentissage. Références, modèles, données et méthodes*. Paris: Didier, Coll. Essais CRÉDIF, 9-22.
- MOORE, D. (Ed.). (2004b). *Les représentations des langues et de leur apprentissage. Références, modèles, données et méthodes*. Paris: Didier.

- MOSCOVICI, S. (1961). *La psychanalyse : son image et son public*. Paris: Presses Universitaires de France.
- MOSCOVICI, S. (1976). *La psychanalyse : son image et son public* (2^e éd.). Paris: Presses Universitaires de France.
- MROUE, M. (2005). *Erreurs lexicales d'apprenants de FLS au Liban en production écrite : essai de typologie*. « Mémoire de maîtrise non publié », Université Stendhal - Grenoble 3, Grenoble.
- NADEAU, M. et FISHER, C. (2006). *La grammaire nouvelle : la comprendre et l'enseigner*. Montréal: Gaëtan Morin, Coll. Chenelière éducation.
- NATION, I. S. P. (1983). « Testing and teaching vocabulary ». *Guidelines*, vol. 5, 12-25.
- NATION, I. S. P. (1990). *Teaching and learning vocabulary*. Boston: Heinle and Heinle.
- NATION, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press, Coll. Cambridge applied linguistics series.
- OLSSON, M. (1972). « Intelligibility ». *The GUME Project*, no 12, Department of Educational Research, Gothenburg School of Education.
- OLSSON, M. (1973). « The Effects of Different Types of Errors in the Communication Situation ». Dans J. Svartvik (Ed.), *Errata*. Lund: CWK Gleerup, 153-160.
- OSTIGUY, L. et TOUSIGNANT, C. (1993). *Le français québécois : normes et usages*. Montréal: Guérin universitaire.
- PATRI, J. et VANSNICK, N. (1992). « Comment aborder les fautes lexicales en classe ? ». *Enjeux*, no 26, 46-55.
- PEREIRA, L. A. (2004). « Des discours sur les pratiques aux pratiques d'enseignement ». Dans C. Barré-De Miniac, C. Brissaud & M. Rispaïl (Eds.), *La littéracie : conceptions théoriques et pratiques de l'enseignement de la lecture-écriture*. Paris: L'Harmattan, 319-332.
- PICOCHÉ, J. (1992). *Précis de lexicologie française : l'étude et l'enseignement du vocabulaire*. Paris: Nathan, Coll. Fac. Linguistique.
- PICOCHÉ, J. et ROLLAND, J.-C. (2001). *Dictionnaire du français usuel : 15 000 mots utiles en 442 articles*. Bruxelles: De Boeck/Duculot.
- POLGUÈRE, A. (1998). « La théorie Sens-Texte ». *Dialangue*, vol. 8-9, Université du Québec à Chicoutimi, 9-30.
- POLGUÈRE, A. (2003). *Lexicologie et sémantique lexicale : notions fondamentales*. Montréal: Presses de l'Université de Montréal, Coll. Paramètres.
- POLGUÈRE, A. (2004a). « La paraphrase comme outil pédagogique de modélisation des liens lexicaux ». Dans É. Calaque & J. David (Eds.), *Didactique du lexique : contextes, démarches, supports*. Bruxelles: De Boeck, Coll. Savoirs en pratique, 115-125.
- POLGUÈRE, A. (2004b). « Savoir consulter un dictionnaire, c'est bien ; savoir comment le construire, c'est mieux ». *Québec français*, no 134, 68-70.
- POLGUÈRE, A. (2007). « "Soleil insoutenable" et "chaleur de plomb": le statut linguistique des greffes collocationnelles ». Dans M.-C. L'Homme & S. Vandaele (Eds.), *Lexicographie et terminologie : compatibilité des modèles et des méthodes*. Ottawa: Presses de l'Université d'Ottawa, 1-24.
- POLGUÈRE, A. (2008). *Lexicologie et sémantique lexicale. Notions fondamentales*. (2^e éd.). Montréal: Les Presses de l'Université de Montréal, Coll. Paramètres.

- POLITZER, R. L. (1978). « Errors of English Speakers of German as Perceived and Evaluated by German Natives ». *Modern Language Journal*, vol. 62, no 5-6, 253-261.
- PORQUIER, R. (1977). « L'analyse des erreurs : problèmes et perspectives ». *Études de linguistique appliquée*, vol. 25, 23-43.
- PORQUIER, R. et PY, B. (2004). *Apprentissage d'une langue étrangère : contextes et discours*. Paris: Didier, Coll. CRÉDIF essais.
- PRÉFONTAINE, C. (1998). *Écrire et enseigner à écrire*. Montréal: Les Éditions Logiques.
- PY, B. (1980). « Quelques réflexions sur la notion d'interlangue ». *Travaux neuchâtelois de linguistique (TRANEL)*, Université de Neuchâtel.
- PY, B. (1984). « L'analyse contrastive : histoire et situation actuelle ». *Le français dans le monde*, vol. 185, 32-37.
- QUIRK, R. et SVARTVIK, J. (1966). *Investigating linguistic acceptability*. The Hague: Mouton.
- QUIVY, M. et TARDIEU, C. (1995). *Réussir l'épreuve sur dossier au CAPES*. Paris: Ellipse, Coll. CAPES/Agrégation d'anglais.
- READ, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press, Coll. Cambridge language assessment series.
- REASON, J. (1990). *Human Error*. Cambridge: Cambridge University Press.
- REBOUL-TOURÉ, S. (2002). « L'enseignement du lexique au collège ». *Le Français Aujourd'hui*, vol. 141, 123-128.
- REID, J. (1990). « Responding to different topic types: A quantitative analysis from a contrastive rhetoric perspective ». Dans B. Kroll (Ed.), *Second language writing: Research insights for the classroom*. Cambridge: Cambridge University Press, 191-210.
- REUTER, Y. (2005). « Définition, statut et valeurs des dysfonctionnements en didactique ». *Repères*, no 31, « L'évaluation en didactique du français, résurgence d'une problématique », 211-231.
- REUTER, Y., COHEN-AZRIA, C., DAUNAY, B., DELCAMBRE, I. et LAHANIER-REUTER, D. (2007). *Dictionnaire des concepts fondamentaux des didactiques*. Bruxelles: de Boeck.
- REY, A. (1972). « Usages, jugements et prescriptions linguistiques ». *Langue Française*, vol. 16 (décembre), 4-28.
- REY, A. (1983). « Norme et dictionnaires ». Dans E. Bédard & J. Maurais (Eds.), *La norme linguistique*. Québec: Conseil de la langue française, 541-569.
- REY-DEBOVE, J. et REY, A. (2002). *Le nouveau petit Robert : dictionnaire alphabétique et analogique de la langue française*. Paris: Dictionnaires Le Robert.
- RICHARDS, J. (1971). « A non-contrastive approach to error analysis ». *English Language Teaching Journal*, vol. 25, 204-219.
- RICHARDS, J. (1976). « The role of vocabulary teaching ». *TESOL Quarterly*, vol. 10, 77-89.
- RIFKIN, B. et ROBERTS, F. D. (1995). « Error Gravity: A Critical Review of Research Design ». *Language Learning*, vol. 45, no 3, 511-537.
- RINGBOM, H. (1987). *The role of the first language in foreign language learning*. Clevedon, Avon, England: Multilingual Matters Ltd., Coll. Multilingual matters.
- ROBERGE, J. (2006). *Corriger les textes de vos élèves : précisions et stratégies*. Montréal: Chenelière Education, Coll. Chenelière/Didactique. Langue et communication.

- ROCQUET, J.-P. (1998). *Ils manquent de vocabulaire ! : lexique et vocabulaire à l'école maternelle et élémentaire*. Marne: CDDP [Centre départemental de documentation pédagogique] Marne.
- SAUSSURE, F. D. (1960). *Cours de linguistique générale*. Paris: Payot.
- SCHNEUWLY, B. (2008). « Le rapport à l'écrit : une notion – deux dimensions en interaction dynamique ». Dans S.-G. Chartrand & C. Blaser (Eds.), *Le rapport à l'écrit : un outil pour enseigner de l'école à l'université*. Namur: Presses universitaires de Namur, 129-138.
- SELINKER, L. (1972). « Interlangage ». *International Review of Applied Linguistics*, vol. 10, 209-231.
- SHAUGHNESSY, M. P. (1977). *Errors and expectations: A guide for the teacher of basic writing*. New York: Oxford University Press.
- SHEEN, R. (1988). « The Importance of Negative Transfer in the Speech of Near-Bilinguals ». Dans D. Nehls (Ed.), *Interlanguage Studies*. Heidelberg: Julius Groos Verlag, 43-57.
- SIMARD, C. (1990). « Les besoins lexicographiques du milieu de l'enseignement au Québec ». Dans *Dix études portant sur l'aménagement de la langue au Québec*. Québec: Conseil de la langue française, 25-51.
- SIMARD, C. (1994). « Pour un enseignement plus systématique du lexique ». *Québec français*, no 92, 28-33.
- SINGLETON, D. (1990). « The cross-linguistic factor in second language learning: a report on some small-scale studies recently conducted at the CLCS ». *CLCS Occasional Paper (Dublin : Trinity College)*, vol. 24.
- SIOUFFI, G. et STEUCKARDT, A. (2001). *La norme lexicale*. Montpellier: Université Paul-Valéry Montpellier III.
- SPENCER, N. J. (1973). « Differences between linguists and non-linguists in intuitions of grammaticality-acceptability ». *Journal of Psycholinguistic Research*, vol. 2, 83-98.
- SPYCHIGER, M. (2006). « La culture de l'erreur – indice d'une conception constructiviste de l'apprentissage ». *Revue suisse des sciences de l'éducation*, vol. 28, no 1, 13-20.
- STAHL, S. A. et NAGY, W. E. (2006). *Teaching word meanings*. Mahwah, N.J.: L. Erlbaum Associates, Coll. The Literacy teaching series.
- STIRMAN-LANGLOIS, M. (1995). « Discours écrit des apprenants et pédagogie de la faute ». *Le français dans le monde : recherches et applications (numéro spécial), La didactique au quotidien : Actes du colloque de Toulon organisé par l'ANEFLE - Campus international - Le français dans le monde*, vol. Juillet 1995, 49-56.
- STUBBS, M. (2001). *Words and phrases: corpus studies of lexical semantics*. Oxford ; Malden, MA: Blackwell Publishers.
- TESNIÈRE, L. (1959). *Éléments de syntaxe structurale*. Paris: Klincksieck.
- TREMBLAY, O. (2003). *Une approche structurée de l'enseignement/apprentissage du lexique en français langue maternelle basée sur la lexicologie explicative et combinatoire*. « Mémoire de maîtrise non publié », Université de Montréal, Montréal.
- TREMBLAY, O. (2009). *Création d'une ontologie des connaissances métalexicales pour l'élaboration d'un module de cours en didactique du lexique destiné aux futurs maîtres au primaire en français langue maternelle*. « Thèse de doctorat non publiée », Université de Montréal, Montréal.

- TRÉVILLE, M.-C. (2000). *Vocabulaire et apprentissage d'une langue seconde : recherches et théories*. Outremont, Québec: Éditions Logiques, Coll. Collection Théories et pratiques dans l'enseignement.
- VAN DER MAREN, J.-M. (1996). *Méthodes de recherche pour l'éducation* (2^e éd.). Montréal: Presses de l'Université de Montréal/De Boeck, Coll. Méthodes en sciences humaines.
- VAN DER MAREN, J.-M. (1999). *La recherche appliquée en pédagogie. Des modèles pour l'enseignement*. Bruxelles: De Boeck, Coll. Méthodes en sciences humaines.
- VANCOMELBEKE, P. (2004). *Enseigner le vocabulaire*. Paris: Nathan, Coll. Les repères pédagogiques. Série Formation.
- VERMERSCH, P. (1994). *L'entretien d'explicitation*. Paris: ESF, Coll. Pédagogies.
- VERMERSCH, P. (2006). *L'entretien d'explicitation* (5^e éd.). Issy-les-Moulineaux: ESF, Coll. Pédagogies outils.
- VERREULT, C. (1999). « L'enseignement du français en contexte québécois : de quelle langue est-il question? ». *Terminogramme. La norme du français au Québec. Perspectives pédagogiques*, no 91-92, 21-40.
- VOGEL, K. (1995). *L'interlangue : la langue de l'apprenant (traduit de l'allemand par J.-M. Brohée et J.-P. Confais)*. Toulouse: Presses universitaires du Mirail, Coll. Interlangues linguistiques et didactiques.
- WARREN, B. (1982). « Common Types of Lexical Errors among Swedish Learners of English ». *Moderna Språk*, vol. LXXVI, no 3, 209-228.
- WIRTHNER, M. (2008). « L'appréhension du rapport à l'écrit par le dispositif didactique ». Dans S.-G. Chartrand & C. Blaser (Eds.), *La rapport à l'écrit : un outil pour enseigner de l'école à l'université*. Namur: Presses universitaires de Namur, 87-105.
- ZIMMERMAN, C. B. (1993). *Teachers Perceptions of and Strategies with Lexical Anomalies*. Communication présentée au Annual Meeting of the Teachers of English to Speakers of Other Languages, Atlanta, April 13-17.
- ZIMMERMANN, R. (1986). « Classification and distribution of lexical errors in the written work of German learners of English ». *Papers and Studies in Contrastive Linguistics*, vol. 21, 31-40.
- ZIMMERMANN, R. (1987). « Form-Oriented and Content-Oriented Lexical Errors in L2 Learners ». *Iral*, vol. 25, no 1, 55-67.
- ŽOLKOVSĀIJ, A. et MEL'ČUK, I. (1967). « O sisteme semantičeskogo sinteza [À propos d'une méthode possible pour la synthèse sémantique] ». *NTI*, no 5, 23-28.
- ZUGHOUL, M. R. (1991). « Lexical Choice: Towards Writing Problematic Word Lists ». *Iral*, vol. 29, no 1, 45-60.

ANNEXE 1

DIFFÉRENTES TYPOLOGIES D'ERREURS LEXICALES

Tableau 25. Typologie des maladresses et des fautes lexicales (Patris et Vansnick, 1992)

MALADRESSES LEXICALES	
1. Intrusion d'un lecte dans un autre	
1.1	Régionalisme ex. : « ... un journal où il <i>fait écrit</i> que l'assassin... » (tour fréquent dans la région montoise);
1.2	Changement de niveau de langue qui rompt avec la dominante du texte ex. : « Il rapporta finalement tout le <i>fric</i> aux gendarmes » [dans une copie homogène, caractérisée par un degré de langue "neutre"].
2. Répétition	
2.1	Reprise excessive du même mot dans une phrase ou dans une suite de phrases;
2.2	« Répétition en écho » ex. : « Je <i>m'occupais</i> à mon <i>occupation</i> favorite quand... »; « Il fallait des <i>preuves</i> pour <i>prouver</i> que... »
3. Imprécision	
3.1	Périphrase au lieu du terme propre ex. : « Il dit des <i>tas d'idées</i> et des <i>raisonnements</i> pour se défendre... »;
3.2	Mot « passe-partout » au lieu du terme propre ex. : « Il découvre dans la mallette des grands papiers où <i>se trouvent</i> des plans.. »

FAUTES LEXICALES	
1. L'erreur concerne une unité du lexique	
1.1	Analogie phonique ex. : Il pénétra dans une usine <i>désinfectée</i> . [« désaffectée »]
1.2	Analogie sémantique ex. : Le malade <i>honore</i> directement son médecin. [« honoraire »]
1.3	Analogie de formation ex. : Les deux voitures se sont <i>collisionnées</i> .
1.4	Confusion de catégories grammaticales ex. : Il voyait un appareil <i>triangle</i> dans le ciel.
2. L'erreur concerne la combinatoire lexicale	
2.1	Non-identification des limites de combinatoire imposées par les traits sémantiques des mots associés ex. : Il voulait <i>capturer</i> ces papiers. Les photos marquées d'une barre étaient <i>décédées</i> récemment (...)
2.2	Non-identification des expressions figées et des limites de leur combinatoire ex. : Il démarra sur <i>des</i> chapeaux de roues. La situation était sans <i>sortie</i> .
3. L'erreur concerne la rection	
3.1	Analogie avec la construction d'un terme synonyme ou appartenant au même champ notionnel ex. : On instruit <i>aux enfants</i> trop de choses inutiles.
3.2	Analogie avec la construction du même terme employé dans un autre environnement morpho-syntaxique ex. : Il se décida <i>d'y jeter</i> un coup d'œil. (« se décider à » / « décider de »)

Tableau 26. Typologie des problèmes lexicaux (Anctil, 2005)

	Dimension touchée	Transgression		Exemples	
A. Altération d'une des dimensions de l'unité lexicale	1. Forme	Erreur d'orthographe lexicale (à l'écrit seulement).		<i>échalotte</i>	
		Barbarisme : altération d'un signifiant qui se manifeste autant à l'oral qu'à l'écrit.		<i>aréoport, lévier de cuisine, de la sphalte, un quipropos.</i>	
	2. Sens	Non-respect d'une composante de sens	Non-respect du typage des actants.	<i>Il ne faut pas déplacer un blessé pour ne pas l'endommager.</i>	
			Non-respect d'une différence spécifique	<i>Elle n'a pas entendu la sonnette, car elle somnolait profondément.</i>	
	3. Propriétés de combinatoire grammaticale	Non-respect du régime	Non-respect de la structure actancielle	<i>Il demande Paul de l'accompagner. Il lui a félicité pour son travail. Le bras droit a démissionné.</i>	
			Utilisation d'une mauvaise préposition pour introduire les compléments. ¹⁹⁷	<i>Rendre visite chez des amis.</i>	
			Non-respect du mode verbal imposé par l'unité lexicale qui régit le verbe dans la subordonnée.	<i>Malgré qu'il est le seul à partir, nous devons l'appuyer.</i>	
			Non-respect d'autres propriétés grammaticales inhérentes à la lexie.		<i>Une avion. La funéraille de mon oncle. Il parle lent.</i>
			Utilisation d'un mauvais verbe auxiliaire pour la formation des temps composés.		<i>J'ai parti tôt ce matin.</i>
			Non-respect du caractère figé d'une locution.		<i>Ses efforts ne valent pas la chandelle.</i>
	4. Propriétés de combinatoire lexicale	Utilisation d'un mauvais collocatif.		<i>Nous payons pour les choix que nous prenons. Heureux comme un paon.</i>	

¹⁹⁷ Nous incluons aussi dans cette classe d'erreurs les problèmes ayant trait aux conjonctions régies par des unités lexicales. Par exemple, la phrase *Son second séjour à Londres fut une expérience meilleure *de ce qu'il avait vécu la première fois.*, l'adjectif MEILLEUR régissant la conjonction QUE.

	Cas	Exemples
B. Utilisation d'une lexie fictive	1. Par la création d'une nouvelle acception d'un vocable selon un patron de polysémie productif.	<i>Il l'attendait sur la rondelle de la station Berri-UQAM (= 'banc de forme arrondie')</i>
	2. Impropropriété : association d'un sens avec le signifiant d'une lexie déjà existante, mais dont le signifié n'entretient aucun lien sémantique avec la nouvelle lexie créée.	<i>Ils ont pris un repas frugal.</i> = 'abondant, copieux'
	3. Par l'application d'une règle de dérivation morphologique généralement productive.	<i>Adapteur, partage, moyen de transportation</i>

	Cas	Exemples
C. Usage douteux en fonction du contexte	1. Changement de registre de langue	<i>Le président a confirmé qu'il en avait plein le cul de ces rumeurs.</i>
	2. Mélange de codes (variantes dialectales)	Dans des textes en français standard : <i>Lorsqu'il faisait chaud, il sortait sa fan.</i> <i>Il l'avait rencontrée à la fin des années septante.</i>
	3. Répétition : excessive ~ en écho	<i>Le commandant commande ses troupes.</i>
	4. Imprécision : périphrase ~ mot générique	<i>Il se dirigea vers la plate-forme où les voyageurs descendaient des trains qui arrivaient.</i> = le quai d'arrivée
	5. Pléonasme	<i>Ils se réunissaient ensemble tous les jeudis.</i>

Tableau 27. Typologie des erreurs lexicales (Milićević et Hamel, 2007)

	Affectant la nomination	Affectant la combinaison	
Erreurs ayant trait au signifié	1. Mauvais quasi-synonyme ex. *C'est aussi une bonne occasion pour regarder (⇒ observer) la nature.	5. Incompatibilité sémantique des sens ex. *La ville n' encourage pas les bicyclettes (⇒ n' encourage pas l' usage des bicyclettes).	
	2. Mauvais terme générique ex. *La bicyclette est un véhicule (⇒ moyen de transport).		
	3. Lexie de sens proche ex. *La bicyclette permet de se rendre à beaucoup d' espaces (⇒ endroits).	6. Incompatibilité pragmatique (du sens et de la situation) ex. *Vous pouvez chercher (⇒ trouver) la plupart des ingrédients au supermarché.	
	4. Sens fictif ex. *... la pluie et les bicyclettes sont des sûretés (⇒ faits) de la vie belge.		
Erreurs ayant trait au signifiant	7. Forme analytique ex. *une bicyclette qui ne bouge pas (⇒ stationnaire)		
	8. Forme erronée ex. * consuiseur (⇒ conducteur) la nature		
	9. Forme fictive ex. *... appartenant à toutes les marches de vie (⇒ couches sociales)		
Erreurs ayant trait au syntactique			Lexicale 10. Mauvaise collocation ex. *Si vous recevez (⇒ attrapez) un coup de soleil, ...
			Syntaxique 11. Mauvais régime ex. *Personne ne s'attendait de ça (⇒ à ça).
			Morphologique 12. Genre 13. Classe flexionnelle
			Stylistique 14. Registre

Tableau 28. Linguistic typology of lexical errors (Zimmermann, 1986)

Error types		Examples
Sense relation errors	*Supernym (→ Hyponym) ¹⁹⁸	I <i>went</i> (→ <i>jumped</i>) upstairs.
	*Hyponym (→ Supernym)	Some of his <i>claims</i> (→ <i>demands</i>) have become less important.
	*Cohyponym ₁ (→ Cohyponym ₂)	A final decision to <i>exterminate</i> (→ <i>eradicate</i>) all dialects
	*Heteronym ₁ (→ Heteronym ₂)	Stairs were being <i>washed</i> (→ <i>scrubbed</i>).
	*Whole (→ Part)	The injury had to be <i>operated</i> (→ <i>sutured</i>).
	*Process (→ Result)	Theodor Siebs had made phonetic <i>recordings</i> (→ <i>records</i>).
Field errors		in a memorial (→ <i>memorandum</i>) by the Foreign Office.
Feature error		I <i>slided</i> (→ <i>slipped</i>) and fell.
Word-formation error	Formally deviant but semantically interpretable	It has been <i>normed</i> (→ <i>become the rule</i>).
	Formally and semantically deficient	The paper is <i>yellowy</i> (→ <i>yellowish? yellowed?</i>).
Collocation error		<i>red-edged</i> (→ <i>red-rimmed</i>) eyes
Non-interpretable form (non-existent lexeme in target language that hasn't been formed by dérivation)		and filled the mouth like a <i>clops</i> (→ <i>lump</i>)
Idiomatic error		These negotiations are <i>cutting a great figure</i> (→ <i>play in important role</i>).
Redundancy error		At a <i>rate</i> (→ ∅) speed of five miles a second
Omission error		I had bitten through ____ (→ <i>edges</i>) my tongue.
Paraphrase errors	Meaning-preserving Formally correct	<i>which had the colour of smoke</i> (→ <i>smoke-coloured</i>)
	Meaning-preserving Formally deviant	<i>in a polite form</i> (→ <i>politely</i>)
	Meaning-changing Formally correct	<i>never greedy</i> (→ <i>always modest</i>)
	Meaning-changing Formally deviant	<i>stood horizontal</i> (→ <i>stuck crosswise</i>)
Stylistic error		The dors of the <i>taverns</i> (→ <i>pubs</i>) are open.
Connotative error		I <i>jerked</i> (→ <i>jumped</i>) upstairs.

¹⁹⁸ La flèche (→) signifie 'au lieu de'.

Tableau 29. Classification of lexical errors (Engber, 1995: 146)

I. Lexical choice

A. Individual lexical items

1. Incorrect – semantically unrelated
 - a. It has some *meanings* to study in the US.
 - b. We can help ourselves with *doing* as international people.
2. Incorrect – semantically close
 - a. They have to *come back* to Indonesia. (*go back*)
 - b. They can study some *strange* subjects. (*unusual*)

B. Combinations

1. Two lexical items
 - a. Young people can *say* their *ideas*.
 - b. I will *bring to go* the development of my country.
2. Phrases
 - a. We have *a lot common*. (*a lot in common*)
 - b. They can discuss their ideas *from the bottom of mind*.
3. Multiple errors involving core lexical items
 - a. Who want to have more exactly their major.
 - b. It is being popular year and year.
 - c. I will get English comprehension perfectly.

II. Lexical Form

1. Derivational errors
 - a. There are a lot of *confliction* between these two countries. (*conflicts*)
 - b. It keeps the class more *activity*. (*active*)
 2. Verb forms
 - a. It isn't a good way for *looking* a job. (*looking for*)
 - b. I want to *make business* with foreign companies. (*do business*)
 3. Phonetically similar, semantically unrelated
 - a. It is difficult to think in the *wild* horizon. (*wide*)
 - b. I *thing* that English will help my country. (*think*)
 4. Word distorted – major spelling errors
 - a. We have learned how to produce a *manifactual* machine.
(*manufactured*)
 - b. We can see the road shows *stimulously*. (*simultaneously*)
-

ANNEXE 2

FORMULAIRES DE CONSENTEMENT

FORMULAIRE DE CONSENTEMENT – ÉLÈVES (1^{ER} OBJECTIF)

Titre de la recherche : Le français écrit au secondaire : description des erreurs d'élèves de 3^e secondaire et du traitement des erreurs par les enseignants

Chercheur : Dominic Anctil

Directrice de recherche : Pascale Lefrançois, Université de Montréal

Codirectrice : Françoise Boch, Université Stendhal-Grenoble 3

A) RENSEIGNEMENTS AUX PARTICIPANTS

1. Objectifs de la recherche

Cette recherche vise à décrire les problèmes de français écrit des élèves du secondaire afin de permettre un enseignement mieux adapté et de développer des outils de correction pour les enseignants.

2. Participation à la recherche

La participation à cette recherche consiste simplement à permettre au chercheur d'avoir accès à un texte que vous avez écrit en classe pour pouvoir le corriger et observer quelles sont les erreurs que vous avez commises. Comme vous aurez déjà produit le texte, la participation à la recherche ne demandera aucun effort supplémentaire de votre part. Si vous acceptez de collaborer à la recherche, vous autoriserez aussi le chercheur à regarder la copie de votre texte corrigée par votre enseignant.

3. Confidentialité

Aussitôt que votre production écrite sera transmise au chercheur, votre nom sera remplacé par un code. De cette façon, il sera impossible de vous identifier pendant la correction ni dans le rapport de recherche. Vos textes seront conservés durant une période de 7 ans dans un bureau fermé à clé, mais seront détruits par la suite.

4. Avantages et inconvénients

En acceptant de participer à cette recherche, vous aiderez le chercheur à proposer des façons d'améliorer l'enseignement du français et la correction de productions écrites.

5. Droit de retrait

Votre participation est entièrement volontaire. Vous êtes libre de vous retirer en tout temps, sur simple avis verbal, sans préjudice et sans devoir justifier votre décision. Si vous décidez de vous retirer de la recherche, vous pouvez communiquer avec le chercheur, au numéro de téléphone indiqué à la dernière page de ce document. Si vous vous retirez de la recherche, les données qui auront été recueillies avant votre retrait seront détruites.

6. Indemnité

Aucune compensation financière ne sera versée pour votre participation à la présente recherche.

B) CONSENTEMENT

On m'a expliqué le projet de recherche et j'accepte d'y participer.

Signature : _____ Date : _____
 Nom : _____ Prénom : _____

Je consens à ce que les données recueillies dans le cadre de cette étude soient utilisées pour des projets de recherche subséquents de même nature, conditionnellement à leur approbation éthique et dans le respect des mêmes principes de confidentialité et de protection des informations. Oui Non

Signature : _____ Date : _____
 Nom : _____ Prénom : _____

Je déclare avoir expliqué le but, la nature et les avantages de l'étude et avoir répondu au meilleur de ma connaissance aux questions posées.

Signature du chercheur
 (ou de son
 représentant) : _____ Date : _____
 Nom : _____ Prénom : _____

Pour toute question relative à la recherche, vous pouvez communiquer avec l'étudiant chercheur par courriel.

Toute plainte relative à votre participation à cette recherche peut être adressée à l'ombudsman de l'Université de Montréal.

Un exemplaire du formulaire d'information et de consentement signé doit être remis au participant

QUESTIONNAIRE

Âge : _____ Sexe : FEMME _____
 HOMME _____

Quelle est la première langue que tu as apprise? _____

Quelle(s) langue(s) parles-tu à la maison? _____

Es-tu toujours allé à l'école en français (depuis la maternelle)?

OUI _____

NON _____ Détails : _____

FORMULAIRE DE CONSENTEMENT – ENSEIGNANTS (1ER OBJECTIF)

Titre de la recherche : Le français écrit au secondaire : description des erreurs d'élèves de 3^e secondaire et du traitement des erreurs par les enseignants

Chercheur : Dominic Anctil

Directrice de recherche : Pascale Lefrançois, Université de Montréal

Codirectrice : Françoise Boch, Université Stendhal-Grenoble 3

A) RENSEIGNEMENTS AUX PARTICIPANTS

2. Objectifs de la recherche

Cette recherche vise à décrire les problèmes de français écrit des élèves du secondaire ainsi que les méthodes de correction mises en oeuvre par les enseignants lors de l'annotation des productions écrites.

2. Participation à la recherche

La participation à ce premier volet de la recherche consiste à permettre au chercheur d'avoir accès à des textes de types narratif ou explicatif que vos élèves auront écrits en classe au cours des mois de mai ou juin 2008. Ces textes constitueront une partie du corpus d'analyse de la recherche. Comme la recherche s'intéresse aussi au traitement des erreurs par les enseignants, notamment à travers la correction, des photocopies des rédactions devront être fournies avant et après la correction; le chercheur s'occupera bien sûr de faire lui-même ces photocopies à ses frais. Comme les textes recueillis seront en fait des rédactions que vous aviez de toute façon prévues, la participation à cette recherche n'implique aucun changement à votre planification pédagogique.

3. Confidentialité

Aussitôt que le chercheur aura récupéré les productions, les noms de tous les élèves seront remplacés par des codes de façon à préserver leur confidentialité. De cette façon, il sera impossible d'identifier les élèves pendant la correction ni dans le rapport de recherche. Les noms des enseignants participant à la recherche seront aussi remplacés par des numéros et ne seront jamais utilisés dans la thèse, pas plus que ne seront mentionnés les établissements scolaires dans lesquels ils travaillent. Toutes les informations recueillies durant ces entrevues seront traitées de manière confidentielle et ne seront utilisées que par le chercheur et son directeur ; en aucun cas elles ne seront divulguées à la direction des écoles. Les textes des élèves seront conservés durant une période de 7 ans dans un bureau fermé à clé, conformément aux règlements de l'Université de Montréal, mais seront détruits par la suite.

4. Avantages et inconvénients

La participation à cette recherche n'implique aucun inconvénient pour vous. Par contre, en acceptant d'y participer, vous aiderez le chercheur à éclairer une dimension jusqu'à maintenant peu abordée en recherche : la question de la correction des productions écrites. Cette recherche devrait fournir certaines pistes pour le développement d'outils de correction de productions écrites destinés aux enseignants.

7. Droit de retrait

Votre participation est entièrement volontaire. Vous êtes libre de vous retirer en tout temps, sur simple avis verbal, sans préjudice et sans devoir justifier votre décision. Si vous décidez de vous retirer de la recherche, vous pouvez communiquer avec le chercheur, au numéro de téléphone indiqué à la dernière page de ce document. Si vous vous retirez de la recherche, les données qui auront été recueillies avant votre retrait seront détruites.

8. Indemnité

Aucune compensation financière ne sera versée pour votre participation à la présente recherche.

B) CONSENTEMENT

Je déclare avoir pris connaissance des informations ci-dessus, avoir obtenu les réponses à mes questions sur ma participation à la recherche et comprendre le but, la nature et les avantages de cette recherche. Après réflexion et un délai raisonnable, je consens librement à prendre part à cette recherche.

Signature : _____ Date : _____
 Nom : _____ Prénom : _____

Je consens à ce que les données recueillies dans le cadre de cette étude soient utilisées pour des projets de recherche subséquents de même nature, conditionnellement à leur approbation éthique et dans le respect des mêmes principes de confidentialité et de protection des informations.	Oui	Non
	<input type="checkbox"/>	<input type="checkbox"/>

Signature : _____ Date : _____
 Nom : _____ Prénom : _____

Je déclare avoir expliqué le but, la nature et les avantages de l'étude et avoir répondu au meilleur de ma connaissance aux questions posées.

Signature du chercheur
 (ou de son
 représentant) : _____ Date : _____
 Nom : _____ Prénom : _____

Pour toute question relative à la recherche, vous pouvez communiquer avec l'étudiant chercheur par courriel.

Toute plainte relative à votre participation à cette recherche peut être adressée à l'ombudsman de l'Université de Montréal.

Un exemplaire du formulaire d'information et de consentement signé doit être remis au participant.

QUESTIONNAIRE ENSEIGNANTS (1ER OBJECTIF)

Quelles étaient les consignes de rédaction données aux élèves? _____

S'agit-il d'une évaluation sommative? _____

La rédaction a-t-elle été faite en classe? _____

Combien de temps était alloué à la rédaction? _____

Est-ce que le recours à des ouvrages de référence était permis en cours de rédaction? Si oui, mentionnez lesquels. _____

Est-ce que cette rédaction a fait l'objet l'objet d'activités de préparation (recherches, lectures, visionnement d'un documentaire, etc.)? Si oui, expliquez brièvement la nature de ces activités. _____

À quelle fréquence durant l'année vos élèves ont-ils écrit des textes comme celui-là? _____

De façon générale, comment évaluez-vous le niveau de maîtrise du français écrit des élèves de ce groupe? _____

Comment décririez-vous le niveau socio-économique de votre école? _____

FORMULAIRE DE CONSENTEMENT – ENSEIGNANTS (2E OBJECTIF)

Titre de la recherche : Le français écrit au secondaire : description des erreurs d'élèves de 3^e secondaire et du traitement des erreurs par les enseignants

Chercheur : Dominic Anctil

Directrice de recherche : Pascale Lefrançois, Université de Montréal

Codirectrice : Françoise Boch, Université Stendhal-Grenoble 3

A) RENSEIGNEMENTS AUX PARTICIPANTS

3. Objectifs de la recherche

Cette recherche vise à décrire les problèmes de français écrit des élèves du secondaire ainsi que les méthodes de correction mises en oeuvre par les enseignants lors de l'annotation des productions écrites.

2. Participation à la recherche

La participation à ce deuxième volet de la recherche consiste à rencontrer le chercheur pour une entrevue de 90 minutes à un moment et dans un lieu que vous choisirez. Cette entrevue portera sur votre traitement des erreurs lors de la correction de productions écrites. Préalablement à l'entrevue, vous devrez transmettre au chercheur cinq rédactions écrites par vos élèves (que vous jugez les plus représentatives de l'« élève moyen ») et sur lesquelles vous aurez effectué la correction; ces textes permettront au chercheur d'orienter ses questions durant l'entretien. L'entrevue sera enregistrée en format audio, puis transcrite.

3. Confidentialité

Aussitôt que le chercheur aura récupéré les productions, les noms de tous les élèves seront remplacés par des codes de façon à préserver leur confidentialité. De cette façon, il sera impossible d'identifier les élèves pendant l'entrevue ni dans le rapport de recherche. Les noms des enseignants participant à la recherche seront aussi remplacés par des numéros et ne seront jamais utilisés dans la thèse, pas plus que ne seront mentionnés les établissements scolaires dans lesquels ils travaillent. Toutes les informations recueillies durant ces entrevues seront traitées de manière confidentielle et ne seront utilisées que par le chercheur et son directeur ; en aucun cas elles ne seront divulguées à la direction des écoles. Les enregistrements des entrevues et les textes des élèves seront conservés durant une période de 7 ans dans un bureau fermé à clé, conformément aux règlements de l'Université de Montréal, mais seront détruits par la suite.

4. Avantages et inconvénients

La participation à cette recherche n'implique qu'un investissement de temps de votre part (90 minutes) et ne demande aucune modification à votre planification en classe. Nous soulignons que la recherche ne vise aucunement à porter un jugement sur la correction faite par les enseignants, mais bien à décrire leurs pratiques de correction.

En acceptant de participer à cette étude, vous aiderez le chercheur à éclairer une dimension jusqu'à maintenant peu abordée en recherche : la question de la correction des productions écrites. Cette recherche devrait fournir certaines pistes pour le développement d'outils de correction de productions écrites destinés aux enseignants.

9. Droit de retrait

Votre participation est entièrement volontaire. Vous êtes libre de vous retirer en tout temps, sur simple avis verbal, sans préjudice et sans devoir justifier votre décision. Si vous décidez de vous

retirer de la recherche, vous pouvez communiquer avec le chercheur, au numéro de téléphone indiqué à la dernière page de ce document. Si vous vous retirez de la recherche, les données qui auront été recueillies avant votre retrait seront détruites.

10. Indemnité

Aucune compensation financière ne sera versée pour votre participation à la présente recherche.

B) CONSENTEMENT

Je déclare avoir pris connaissance des informations ci-dessus, avoir obtenu les réponses à mes questions sur ma participation à la recherche et comprendre le but, la nature et les avantages de cette recherche. Après réflexion et un délai raisonnable, je consens librement à prendre part à cette recherche. Je sais que je peux me retirer en tout temps sans préjudice et sans devoir justifier ma décision.

Signature : _____ Date : _____
 Nom : _____ Prénom : _____

Je consens à ce que les données recueillies dans le cadre de cette étude soient utilisées pour des projets de recherche subséquents de même nature, conditionnellement à leur approbation éthique et dans le respect des mêmes principes de confidentialité et de protection des informations.	Oui	Non
	<input type="checkbox"/>	<input type="checkbox"/>

Signature : _____ Date : _____
 Nom : _____ Prénom : _____

Je déclare avoir expliqué le but, la nature et les avantages de l'étude et avoir répondu au meilleur de ma connaissance aux questions posées.

Signature du chercheur
 (ou de son représentant) : _____ Date : _____
 Nom : _____ Prénom : _____

Pour toute question relative à la recherche, vous pouvez communiquer avec l'étudiant chercheur par courriel.

Toute plainte relative à votre participation à cette recherche peut être adressée à l'ombudsman de l'Université de Montréal.

Un exemplaire du formulaire d'information et de consentement signé doit être remis au participant

CERTIFICAT D'ÉTHIQUE



COMITÉ PLURIFACULTAIRE D'ÉTHIQUE DE LA RECHERCHE (CPÉR)

CERTIFICAT D'ÉTHIQUE

Le Comité plurifacultaire d'éthique de la recherche a examiné le projet de recherche intitulé :

« L'erreur lexicale au secondaire: description des problèmes lexicaux d'élèves de 3e secondaire et du traitement des erreurs par les enseignants »

Soumis par : **Dominic Anctil**
 Directrice de recherche: **Pascale Lefrançois**
 Co-directrice : **Françoise Boch** (Université Stendal-Grenoble 3)

Le Comité a conclu que le projet respecte les normes de déontologie énoncées à la « Politique sur la recherche avec les êtres humains » de l'Université de Montréal.

Tout changement anticipé au protocole de recherche doit être communiqué au CPÉR qui devra en évaluer l'impact au chapitre de l'éthique afin de déterminer si une nouvelle demande de certificat d'éthique est nécessaire.

Toute interruption prématurée du projet ou tout incident grave devra être immédiatement signalé au CPÉR.

25/04/08

François Bowen, Président
 Comité plurifacultaire d'éthique de la recherche
 Université de Montréal

Date d'émission

ANNEXE 3

EXTRAIT DE CODAGE (OBJECTIF 1)

	ID	Nb m	Erreur [correction]	Cert.	Pos. Ens.	Cmt ens.	CdeM	UL	TYPE	Accep	Comp	Idio	Gravité	Commentaire description	Source	Commentaire explication
194	SJ4-18	457	il avait déjà [déjà] atteint le compte de neuf queue.	É	NR		Adv	Lex	F-Orth	2	0	2	4	accent -		
195	SJ4-19	338	les farfadets Alex, Camerone, Jonh, et Rose, font [vont] faire un toure en compagne	É	EP	S	V	Lex	F-Barb	2	1	2	5	confusion de phonèmes		
196	SJ4-20	389	ses souliers étaient accompagnés [ornés/décorés] d'argent.	É	NR		Adj	Lex	S-Fict	2	1	2	5		TrLÉlève	J'imagine que l'élève, qui par ailleurs écrit plutôt bien, sentait que ce n'était pas le meilleur choix de mot.
196	SJ4-20	389	Mais ses parents étaient persuadés que son imagination faisait encore des siennes [faisait (encore) des bêtises]	M	NR		V	Loc	CP-Regis	2	0	0	2	Cette expression apparait seulement dans le Multi, qui la considère familière. Par contre, je trouve que dans le contexte, c'est plutôt correct et la correction proposée n'est guère plus appropriée...	Intégr	
196	SJ4-20	389	Il lui cria qu'il était temps de grandir et qu'elle cesse toutes ces follies [folies]	É	EP	U	N	Lex	F-Orth	2	0	2	4	consonne double +		
197	Sj4-21	308	Dudeley aimait tellement la sirène qu'il voyait tout les soirs au bord de l'eau [!/que...].	É	NA		Adv	Lex	CS-AbsC	2	0	1	3	TELLEMENT est un adverbe qui demande une subordonnée corrélatrice. Sinon, il s'agit d'un adverbe exclamatif qui doit apparaître dans une phrase exclamative.	Intégr	
197	Sj4-21	308	Dudeley couru aller voir [couru voir] son ami.	É	EP	S	V	Lex	S-Pléo	2	0	2	4	COURIR = 'aller, se déplacer rapidement..."		

197	Sj4-21	308	Rendue en dessous du [en dessous du] filet cyclope tribucha	É	EP	U	Prép	Loc	F-Orth	2	0	2	4	lettre muette -		
197	Sj4-21	308	Rendue en dessous du filet cyclope tribucha [trébucha]	É	EP	U	V	Lex	F-Barb	2	0	2	4	confusion de phonèmes		
197	Sj4-21	308	Un jour il vit deux pêcheurs [pêcheurs] la capturer dans un énorme filet.	É	EP	U	N	Lex	F-Orth	2	0	2	4	accent (mauvais choix)		
197	Sj4-21	308	ils vurent [virent] les deux créatures	É	EP	G	V	Lex	F-Flex	2	0	2	4	3e groupe		
197	Sj4-21	308	Les deux espèces [créatures] essayèrent de monter dans le bateau	É	NR		N	Lex	S-Fict	2	1	2	5	ESPÈCE = 'Classe (de personnes, de choses) définie par un ensemble particulier de caractères communs'. Dans ce cas-ci, l'élève ne fait pas référence à une classe de créatures, mais bien à deux "instances" de créatures précises, dont l'une est un cyclope et l'autre un centaure. En fait, il s'agit de deux personnages.	ProxSém	Il y a un lien de sens, puisque les deux personnages appartiennent à deux espèces de créatures bien particulières.
197	Sj4-21	308	Ils réussirent [réussirent] à monter à bord.	É	EP	G	V	Lex	F-Flex	2	0	2	4	2e groupe		
197	Sj4-21	308	Cyclope aussi stupide qu'il soit, en regardant les merveilleuses montagne au loin, il fit le Titanic [fit comme Leonardo di Caprio dans le film <i>Titanic</i>] au bout du bateau.	É	NR		V	Loc	F-Inex	1	1	2	4	La locution FAIRE LE TITANIC n'existe pas et peut poser un problème grave de compréhension pour qui n'a pas vu le film <i>Titanic</i> , qui pourrait alors interpréter cela comme 'couler au fond de l'eau'.	EISém	L'élève sous-entend <i>comme Leonardo di Caprio dans le film Titanic</i>
197	Sj4-21	308	Pendant ce temps, Dudeley aisillait [essayait] de détacher le filet	É	EP	U	V	Lex	F-Orth	2	1	2	5	mauvais graphème/consonne double -/mauvais graphème. Confusion possible avec ASSAILLIR, même si le contexte syntaxique est incompatible.	Dist	La forme <i>essayèrent</i> apparait quelques lignes plus haut.

197	Sj4-21	308	cyclope pila [marcha, posa le pied; écrasa] sur la main de l'un des pêcheurs	É	NR		V	Lex	CP-Regis	2	0	1	3		Intégr	
197	Sj4-21	308	cyclope les assoma [assomma] avec son baton	É	NR		V	Lex	F-Orth	2	0	2	4	consonne double -		
197	Sj4-21	308	cyclope les assoma avec son baton [bâton]	É	EP	U	N	Lex	F-Orth	2	0	2	4	accent -		
198	Sj4-22	389	le plus crain [craint] des clans de vampires.	É	EP	G	Adj	Lex	F-Orth	2	0	2	4	lettre muette -		
198	Sj4-22	389	Il était triste par [du] (le) refu de sa belle mais il continuait de rêver d'elle.	É	NR		Adj	Lex	CS-ChPr	2	0	2	4	<i>X est triste de Y</i>	AnaConst	<i>Il était attristé par le refu de sa belle.</i>
198	Sj4-22	389	Il était triste par le refu [refus] de sa belle mais il continuait de rêver d'elle.	É	EP	U	N	Lex	F-Orth	2	0	2	4	lettre muette -		
198	Sj4-22	389	Elle se laissa donc mourrir, son corps pénétré par les premiers rayons du Soleil [soleil].	É	NR		N	Lex	F-Orth	2	0	2	4	majuscule		
199	Sj4-23	290	hélas après plusieurs minutes, essoufflée à crier aucune réponse de l'autre côter [côté]	É	EP	U	N	Lex	F-Orth	2	0	2	4	mauvais graphème		
199	Sj4-23	290	La nuit tombée, elle s'endormit avec l'espoir de pouvoir sortir d' ici [(de) là]	É	NR		Adv	Lex	CP-Conn	2	0	2	4	ICI est un déictique et ne peut donc pas être utilisé dans la narration pour indiquer un endroit du point de vue d'un personnage.	SApprox	Peut-être aussi l'erreur découle-t-elle d'une difficulté de l'élève à se décentrer dans son écriture.
199	Sj4-23	290	elle entendit qu'elqu'un [quelqu'un] dire : « il y a quelqu'un ici ? »	É	EP	U	N	Lex	F-Orth	2	0	2	4	division des mots-formes	Dist	Le mot est correctement orthographié plus loin dans la même phrase.

199	SJ4-23	290	Puis le sphinx avec sa puissante force [exceptionnelle, surhumaine, incroyable] fit sortir la petite fille.	É	NR		N	Lex	CL-Inex	2	0	2	4	Magn(<i>force</i>)= <i>exceptionnelle, surhumaine</i> , etc. J'aurais peut-être pu coder aussi S-Pléo	ProxSé m	PUISSANCE et FORCE sont des synonymes, c'est donc étrange de les mettre en cooccurrence.
199	SJ4-23	290	elle eu peur par [de] son corps de lion et sa tête humaine	É	EP	S	V	Loc	CS-ChPr	2	0	2	4	<i>X a peur de Y</i>	AnaCons t	être effrayé par /avoir peur de
199	SJ4-23	290	Le gardien reprit les trésors volés et repartit à [vers/retourna à] la pyramide.	É	NR		V	Lex	CS-ChPr	1	0	2	3	Je ne sais pas trop pourquoi, mais le choix de la préposition À avec REPARTIR me dérange dans ce contexte. <i>Il repartit vers la pyramide</i> ou <i>Il retourna à la pyramide</i> me semblent deux formulations plus correctes. Pourtant, si ça avait qqchse comme <i>il repartit au campement</i> , on dirait que ça ne m'aurait pas dérangé... Peut-être est-ce une question de point de vue, REPARTIR prenant comme point d'appui l'endroit du départ, alors que RETOURNER met plutôt l'accent sur la destination (qui nous intéresse davantage dans ce cas, puisque l'on ne sait pas trop où le sphinx est rendu).	AnaCons t	REPARTIR vers/RETOURNER à
200	SJ4-24	295	Les Ragnaröks, petits être étant pareil comme [aux] (les) farfadets	É	NR		Adj	Lex	CP-Regis	2	0	1	3	Il est faut de dire que <i>pareil comme</i> n'existe pas. Par contre, l'emploi de cette préposition avec PAREIL est considéré comme familier.	Intégr	
200	SJ4-24	295	Les Ragnaröks, petits être étant pareil comme les farfadets, étaient plus mesquins, cruels, sans pitié et de plus, ils étaient laids	É	EP	G	N	Lex	CM-Inv	1	0	2	3	PUISSANCE au sens de 'Pouvoir, autorité sur une collectivité.' est selon moi un singulare tantum.	ContSyn	Le nombre est tout le temps un problème avec le déterminant possessif LEUR.

			mais, malgré tout cela leurs puissances [leur puissance] étaient incomparables.														
200	SJ4-24	295	Les Ragnaröks, petits être étant pareil comme les farfadets, étaient plus mesquins, cruels, sans pitié et de plus, ils étaient laids mais, malgré tout cela leurs puissances étaient incomparables [comparables].	M	NR		Adj	Lex	S-Proc	2	2	2	6	INCOMPARABLE = 'Que l'on ne peut comparer à autre chose; essentiellement différent' Pourtant, dans le texte, l'élève semble mettre l'accent sur le fait que les farfadets et les Ragnaröks ne se distinguent que par quelques détails, mais que les deux peuples sont, par ailleurs, très comparables.			
200	SJ4-24	295	Depuis quelques temps [quelque temps], un grondement sourd unissaient les farfadets	É	NR		N	Loc	F-Orth	2	0	2	4	accord			
200	SJ4-24	295	ils allaient s'unir contre Mystia, peu important [peu importait sa puissance, les circonstances...]	M	NR		Adv	Loc	CM-Inv	1	0	2	3	J'ai l'impression qu'ici PEU IMPORTE est utilisé comme préposition et demande par conséquent un complément. Sinon, je sais que PEU IMPORTE peut aussi s'utiliser comme Adv (<i>Peu importe, ils allaient s'unir contre Mystia</i>); par contre, je trouve ça un peu étrange dans ce contexte. Par ailleurs, c'est peut-être l'utilisation de PEU IMPORTE à l'imparfait qui m'embête... Ladverbe PEU IMPORTE est invariable (non conjugable)			
200	SJ4-24	295	la reine malifique [maléfique] se moquait beaucoup des	É	EP	U	Adj	Lex	F-Barb	2	0	2	4	confusion de phonèmes			

			farfadets.														
200	SJ4-24	295	Le jour vint succéder [succéder à] la nuit	É	NR		V	Lex	CS-Trans	2	0	2	4		AnaCons t	Le verbe SUIVRE, qui est un quasi-synonyme, est transitif direct.	
200	SJ4-24	295	ce jour la [-là] serait un grand pas pour les farfadets.	É	EP	G	Adv	Lex	F-Orth	2	0	2	4	accent -/trait d'union -			
200	SJ4-24	295	Elle commençait à s'énerver. Puis, elle commença à se dissoudre [s'évaporer, s'évanouir]. Mystia disparut, impuissante.	É	NR	P	V	Lex	S-Proc	2	0	1	3	Le sens de DISSOUDRE implique un liquide 'former un mélange homogène avec un liquide, les éléments constituants d'un corps'. Par contre, le sens vieilli de DISSOUDRE 'Décomposer (un agrégat, un organisme) par la séparation des parties.' pourrait correspondre au contexte... Cette erreur ne me semble pas si dramatique.	SApprox	L'idée de disparition est tout de même très présente dans DISSOUDRE.	
201	SJ4-25	361	elle le fit cuire grâce aux flammes qu'elle propulsait [lançait, crachait] de sa bouche.	É	NR		N	Lex	CL-Inex	2	0	2	4	Il me semble que les verbes qui "fonctionnent" avec FLAMME sont LANCER et CRACHER (à la limite PROJETER, mais je ne suis pas certain). Dois-je aussi considérer un problème de préposition régie: <i>propulser des flammes de sa bouche</i>	Anaform /AnaSé m	Parenté formelle et sémantique entre PROJETER et PROPULSER, mais je ne suis sûr d'aucun des deux...	

ANNEXE 4

GUIDE DE CONTRE-CODAGE

GUIDE DE CONTRE-CODAGE

La présente recherche vise la description des erreurs lexicales commises par des élèves francophones de troisième secondaire en production écrite. Par erreur lexicale, nous entendons toute erreur linguistique concernant une propriété inhérente à une unité lexicale, que cette erreur soit d'ordre formel, sémantique, collocationnel (combinaison de mots) ou lexicogrammatical (structure régie par le mot, genre du nom, invariabilité, etc.).

Le codage des erreurs est terminé, mais, dans le but de lui assurer une certaine validité, une partie du corpus d'erreurs doit faire l'objet d'un contre-codage. À la suite de nombreuses discussions avec Pascale et après recommandations des membres du jury, j'ai choisi de restreindre l'étape de contre-codage à certains aspects bien spécifiques, soit les dimensions marquées par une plus grande subjectivité et, par conséquent, les plus susceptibles de donner lieu à des divergences entre les différents codeurs.

Mon corpus total comprend plus de 2700 erreurs relevées dans 300 rédactions. Aux fins du contre-codage, j'ai d'abord exclu toutes les erreurs d'orthographe lexicale, auxquelles ma thèse ne s'intéresse que de façon périphérique; ceci limite le corpus à 1300 problèmes. Le contre-codage portera sur 10 % de ce sous-corpus, soit 130 erreurs. Les erreurs retenues pour le contre-codage ont été choisies de façon aléatoire (erreurs 1, 11, 21, 31, 41, etc.) et figurent dans le tableau Excel que vous devrez utiliser pour coder les erreurs (ce fichier est joint au présent courriel).

ASPECTS VISÉS PAR LE CONTRE-CODAGE

Chaque erreur lexicale relevée dans le corpus de textes a été analysée selon trois axes : 1) une description linguistique à l'aide d'une typologie d'erreur lexicale, 2) une évaluation de la gravité de l'erreur sur la base de trois critères et 3) une explication de la source de l'erreur (si possible). Le contre-codage pour lequel vous êtes sollicités concerne les niveaux 2 (gravité) et 3 (source de l'erreur).

ÉVALUATION DE LA GRAVITÉ DES ERREURS

Plusieurs facteurs entrent en jeu lorsque nous posons un jugement de gravité sur une erreur linguistique, qu'elle soit grammaticale, syntaxique ou lexicale. Dans notre recherche, nous avons retenu les trois facteurs qui nous semblaient les plus pertinents pour l'erreur lexicale. Ces facteurs sont les suivants :

1. **L'acceptabilité linguistique**, c'est-à-dire la position de la norme, telle qu'actualisée dans les dictionnaires, par rapport à l'erreur. Une erreur clairement indiquée comme telle dans les ouvrages (ex. emploi abusif de FUTUR_{nom} au sens de 'avenir') ou allant à l'encontre de ce que l'on retrouve dans les dictionnaires (ex. OPAQUE au sens de 'dur') est considérée comme plus grave, en regard de la norme, qu'une erreur qu'on ne peut condamner sur la base de ce qui

apparaît dans les ouvrages (ex. la combinaison de l'adjectif *VIEUX* et du nom *ADOLESCENT* pour former l'expression *les vieux adolescents*, qui signifie 'les jeunes adultes').

2. La **compréhensibilité**, c'est-à-dire l'impact que l'erreur a sur la compréhension du lecteur de la phrase dans laquelle apparaît cette dernière. Plus une erreur compromet la compréhension du message, plus elle est considérée comme grave. Bien sûr, il est souvent possible, par inférence, de reconstruire le sens visé par le scripteur, surtout dans le contexte d'un texte; il est par conséquent plutôt rare qu'une erreur entrave complètement la compréhension du message.
3. L'**idiomaticité**, c'est-à-dire le caractère plus ou moins marqué de l'item considéré comme une erreur par rapport à l'usage « normal ». Plusieurs erreurs sont en effet tout à fait intégrées à l'usage (ex. *JOB* pour signifier 'emploi'), alors que certaines nous paraissent carrément bizarres (ex. **Il eut la grâce idée de demander au nain d'y aller à sa place.*). Une erreur fréquente sera perçue comme moins grave qu'une formulation bizarre.

Le travail de recherche dans les ouvrages de référence ayant déjà été fait dans le but de statuer sur l'acceptabilité linguistique des erreurs du corpus, vous n'aurez à juger les 130 erreurs du corpus de contre-codage qu'en fonction des deux derniers critères, soit la compréhensibilité et l'idiomaticité. Ces deux critères sont très subjectifs et reposent complètement sur vos impressions de lecteurs (compréhensibilité) et vos intuitions de locuteurs du français (idiomaticité); vous n'avez donc aucune recherche à faire pour pouvoir attribuer un degré de gravité aux erreurs.

Vous devrez, pour chacun des deux critères, attribuer une cote de 0 à 2 à chaque erreur, selon le tableau suivant.

Jugement du correcteur	Impact sur la compréhension	n'a aucun impact sur la compréhension	0
		peut être interprété , mais demande un certain effort d'inférence	1
		compromet la compréhension , c'est-à-dire qu'il est impossible de rétablir avec certitude le message visé par l'élève qui a produit l'erreur (aucune correction proposée ou des corrections concurrentes qui divergent sémantiquement)	2
	Perception de l'intégration à l'usage	la formulation semble tout à fait naturelle en français (intégration au code, ex. anglicisme comme <i>opportunité</i>)	0
		la formulation semble maladroite et marquée par rapport à l'usage normal / la formulation semble naturelle, mais pas dans un contexte de rédaction scolaire	1
		la formulation semble carrément boiteuse (on n'entendrait jamais ça)	2

REMARQUES :**Impact sur la compréhension :**

- Comme nous l'avons mentionné plus haut, il est rare qu'une simple erreur compromette complètement la compréhension d'un texte entier (ou d'une phrase complète), surtout qu'il se peut qu'elle n'affecte qu'une information périphérique, ce qui laisse donc intact le « noyau » de sens de l'énoncé dans lequel apparaît l'erreur. Lorsque nous jugeons de l'impact de l'erreur sur la compréhension du message, il est donc important de se demander si l'erreur compromet la compréhension d'une partie de ce dernier. Par exemple, on comprend le sens général de la phrase suivante : *#Alors qu'elle était proche du rivage, un objet **opaque** vint la frapper derrière la tête et Rosalia perdit connaissance.* Pourtant, l'adjectif opaque ne semble pas du tout pertinent dans ce contexte, et il est difficile de déterminer ce que l'élève voulait exprimer exactement; voulait-il signifier que l'objet était dur, qu'il était lourd, massif? Donc, même si le sens général de la phrase reste compréhensible, cette erreur se verrait attribuer une cote de 2 pour la compréhensibilité.
- Un exemple d'erreur qui demande un certain effort d'inférence pour rétablir le sens visé serait la phrase suivante : *#Un jour, l'horloger **aperçut** que les gens vieillissaient très rapidement.* On comprend assez aisément que l'horloger se *rend compte, s'aperçoit* du fait que les gens vieillissent très rapidement, qu'il le *constate* et non qu'il l'*aperçoit* au sens de 'percevoir par la vue, de façon plus ou moins nette'. Si on peut rétablir le sens visé, c'est en s'appuyant sur le contexte de la phrase et en faisant une inférence; par conséquent, cette erreur se verrait attribuer une cote de 1 pour la compréhensibilité.
- Certaines erreurs, au contraire, n'entraînent aucun problème de compréhension (ex. *un nain ***apparaissa**, ***une longue** été, il la colla ***après** un poteau*); de telles erreurs recevraient la cote 0 pour le critère *impact sur la compréhension*.

Perception de l'intégration à l'usage :

- Lorsque vous statuez sur le caractère plus ou moins idiomatique des erreurs observées, il ne faut pas oublier de tenir compte du contexte de production des textes desquels ont été tirées les erreurs. Ces rédactions constituent des évaluations sommatives (souvent de fin d'année), qui requièrent l'utilisation d'un niveau de français standard. Ainsi, demandez-vous toujours « Est-ce que j'aurais pu écrire ça dans une rédaction de fin d'année? ». Ainsi, des emplois pourtant tout à fait naturels pourraient se voir attribuer une cote 1 (ex. PAREIL_{adv} au sens de 'quand même' → *Je l'aime pareil.*).
- Le niveau 2, quant à lui, est réservé aux formulations vraiment bizarres (ex. *un roi **sous le nom de** Grand Robert*) ou qui n'apparaîtraient jamais à l'écrit (ex. **tomber **sur** la connaissance / L'arbre **#s'affala** sur le ruisseau*).

CONSIGNES POUR LE CONTRE-CODAGE

Voici un exemple de ce à quoi ressemble le fichier de codage avec lequel vous aurez à travailler. Les colonnes ombragées sont celles qui concernent directement votre tâche de contre-codage.

	Erreur [correction]	TYPE	Commentaire description	Impact sur la compréhension	Perception de l'intégration à l'usage	Commentaire sur le codage	Explication de la source de l'erreur
1	la flèche s'envolera jusqu'au coeur de la bête et l'animal s'éteindra [mourra]	S-Syn	Selon moi, S'ÉTEINDRE ne fonctionne pas dans ce contexte. On dirait que le sens de ce verbe implique une longue agonie ou du moins une mort douce.	0	1	Le message est tout à fait limpide, mais ça me semble bizarre d'utiliser S'ÉTEINDRE.	Proximité sémantique entre MOURIR et S'ÉTEINDRE. Ce sont des synonymes.
2	Gregory eut l'idée de faire bâtir un endroit où les sans-abri pourraient héberger [loger; être hébergés].	S-Proc		1	2		Mauvaise connaissance de la structure du verbe. <i>X héberge Y, mais Y est hébergé par X</i>
3	Le lendemain soir après le boulot [travail]	CP-Regis	BOULOT signifie vraiment 'travail, emploi', mais appartient au registre familier.	0	0		Le mot <i>boulot</i> est complètement intégré à la langue et on peut même l'entendre à la télévision. En fait, ce n'est même pas clair si ce mot est familier, puisqu' <i>Antidote</i> ne donne aucune indication à cet effet.

La colonne *erreur [correction]* contient un extrait de texte plus ou moins important dans lequel apparaissent l'erreur (en gras) et la correction que j'ai proposée (entre crochets). J'ai tenté de toujours fournir un contexte suffisant pour que l'erreur soit évidente (c'est ce qu'on appelle le *domaine de l'erreur*). La correction que je propose tente toujours de garder le texte de l'élève le plus intact possible (correction minimale). Par conséquent, la correction peut contenir des fautes d'accord; je n'ai réglé que le problème lexical proprement dit (et seulement le problème lexical correspondant au code de classe d'erreurs dans le cas où il y a chevauchement de problèmes, voir paragraphe suivant).

La seconde colonne, *TYPE*, contient la classe de l'erreur, qui correspond à l'une des catégories de ma typologie des erreurs lexicales, reproduite ci-dessous (pour plus de détails sur chacune des classes d'erreurs, voir le document *Survol des classes de la typologie des erreurs lexicales* envoyé dans ce courriel). **Il est très**

important de regarder le code d'erreur, puisqu'il est assez fréquent que des erreurs se superposent et il faut être bien conscient de l'erreur traitée pour pouvoir poser de bons jugements de compréhensibilité et d'idiomaticité. Prenons par exemple la phrase suivante : *Il passait tous ses **passes-temps** à dessiner*. L'emploi de PASSE-TEMPS, ici, présente un double problème. Tout d'abord, ce mot-forme contient une erreur d'orthographe, puisque la forme de PASSE-TEMPS est la même au singulier et au pluriel (pas de *s* à *pas*). Ensuite, l'emploi de PASSE-TEMPS est problématique d'un point de vue sémantique, puisque le sens de cette unité lexicale est 'occupation agréable, divertissement', alors que tout porte à croire que l'élève voulait plutôt dire *temps libres*. Il est donc primordial de bien savoir de quelle erreur on parle lorsque l'on doit poser un jugement : l'erreur d'orthographe n'implique aucun problème de compréhension (=0), tandis que l'erreur sémantique demande un certain effort de compréhension (=1).

La colonne *commentaire description* contient parfois des commentaires qui justifient mon codage relativement au type d'erreur (colonne 2). J'ai choisi de laisser cette colonne dans votre fichier, puisqu'elle peut peut-être vous donner des informations utiles ; par contre, vous n'avez rien à y inscrire.

C'est dans colonne *impact sur la compréhension* que vous attribuerez une cote de 0 à 2 pour la compréhensibilité, selon le barème présenté plus haut. La colonne *perception de l'idiomaticité* reçoit aussi les cotes de 0 à 2 selon le même barème. Une colonne *commentaire sur le codage* a été ajoutée au cas où vous voudriez justifier en une phrase ou en quelques mots les cotes attribuées à l'erreur (seulement lorsque cela vous semble nécessaire ou pertinent). Vous pouvez aussi vous servir de cette colonne pour inscrire tout autre commentaire à propos de l'erreur, par exemple une autre suggestion de correction que vous jugez plus appropriée. **Dans le cas où vous n'auriez pas considéré le problème comme une erreur, mettre la ligne en jaune.**

La colonne *source de l'erreur* vous permet de poser une hypothèse sur la source de l'erreur analysée, dans la mesure où une hypothèse vous vient à l'esprit. Par exemple, dans la phrase **elle est une jeune fille **bordée** [débordante] d'imagination*, on peut supposer que la proximité formelle entre BORDEE et DEBORDANTE est à l'origine de l'erreur. Dans le syntagme *une musique qui le fit s'endormir ***dans** [d'] un sommeil profond*, c'est plutôt une analogie avec les constructions *tomber/somber **dans** un profond sommeil* qui semble permettre d'expliquer l'emploi fautif de la préposition DANS avec le verbe S'ENDORMIR. **Vous n'avez à proposer une explication de la source de l'erreur que lorsqu'une telle explication vous vient à l'esprit** ; toutes les erreurs ne sont pas explicables. Il faut donc se demander, pour chaque erreur, « Pourquoi l'élève a-t-il commis cette erreur ? » Par ailleurs, je ne fournis pas de liste des sources possibles d'erreurs lexicales, puisque cette dimension de ma recherche est très exploratoire et qu'il n'existe pas pour l'instant de liste fermée de scénarios d'explication d'erreurs ; j'ai moi-même ajouté de nouvelles hypothèses tout au long de mon codage. Par contre, je peux tout de même mentionner que mes explications reposent souvent sur la

notion d'analogie (soit formelle, soit sémantique, soit de construction – c'est-à-dire appliquer à un mot une structure qui correspond à un autre) ou sur le fait que l'usage erroné est très fréquent dans la langue familière (ex. **une bouteille en vitre*).

Typologie descriptive des erreurs lexicales

DIMENSION TOUCHÉE	CLASSES D'ERREURS		CODES	
FORME	Mauvaise orthographe lexicale		F-Orth	
	Barbarisme (signifiant déformé)		F-Barb	
	Mauvaise forme fléchie		F-Flex	
	Forme analytique		F-Déf	
	Unité lexicale inexistante		F-Inex	
SENS	Mauvaise utilisation d'un quasi-synonyme		S-Syn	
	Mauvaise utilisation d'une unité lexicale de sens proche		S-Proc	
	Impropropriété	Sens fictif	S-Fict	
		Anglicisme sémantique	S-Ang	
	Non-respect des contraintes sémantiques pesant sur les actants		S-Act	
Pléonasme		S-Pléo		
COMBINATOIRE RESTREINTE	GRAMMATICALE	SYNTAXIQUE (RÉGIME)	Absence d'un complément obligatoire de l'UL	CS-Absc
			Emploi transitif indirect d'un verbe transitif direct	CS-Trans
			Préposition régie absente	CS-AbPr
			Utilisation d'une mauvaise préposition/conjonction	CS-ChPr
			Construction erronée dans un complément	CS-Const
		MORPHO-SYNTAXIQUE	Non-respect de la classe de mots de l'UL	CM-Class
			Non-respect du genre du nom	CM-Genr
			Non-respect du caractère invariable de l'UL	CM-Inv
			Utilisation d'un mauvais verbe auxiliaire	CM-Aux
	Classe résiduelle		CM-Rés	
	LEXICALE	Choix d'un mauvais collocatif (collocation inexistante)		CL-Inex
		Choix d'un collocatif valable, mais mauvaise fonction lexicale		CL-FL
		Non-respect des contraintes d'emploi du collocatif		CL-Cont
		Emploi d'un mauvais dérivé sémantique		CL-Dér
	PRAGMATIQUE	Emploi d'une UL appartenant à un registre de langue inapproprié		CP-Regis
		Emploi d'une UL véhiculant un contenu connotationnel incompatible avec le contexte		CP-Conn
		Emploi d'une UL appartenant à une variété dialectale incompatible avec le contexte		CP-Dial

ANNEXE 5

QUESTIONNAIRE D'ENTREVUE

QUESTIONS POUR LES ENTREVUES AVEC LES ENSEIGNANTS

1^{er} volet

LA MAITRISE DE LA LANGUE ET DU VOCABULAIRE

1. On entend souvent dire que les élèves ne savent plus écrire, qu'ils ne maîtrisent pas la langue. Selon vous, que signifie « maîtriser une langue »?
2. Est-ce important d'avoir un vocabulaire étendu? Pourquoi? Qu'est-ce que ça signifie, « avoir un vocabulaire étendu »?
3. Comment décririez-vous le vocabulaire de vos élèves?
 - a. Lors de l'épreuve unique de la fin du secondaire, les élèves québécois obtiennent année après année de très bons résultats pour le critère vocabulaire. Ces résultats sont-ils selon vous représentatifs de leur niveau de maîtrise du vocabulaire? Si non, qu'est-ce qui explique de tels résultats?

L'ENSEIGNEMENT DU VOCABULAIRE EN CLASSE DE FRANÇAIS AU SECONDAIRE

4. Quel est selon vous le rôle de l'école dans l'acquisition de ce vocabulaire? Est-il important d'enseigner le vocabulaire en classe de français?
5. Enseignez-vous vous-même le vocabulaire en classe?
 - a. Si oui, de quelle façon? Dans quelles circonstances? À quelle fréquence? Dans quels buts?
 - b. Sur quels aspects du vocabulaire travaillez-vous surtout?

LA CORRECTION DE PRODUCTIONS ÉCRITES ET L'ERREUR DE VOCABULAIRE

6. Quel type de grille utilisez-vous lorsque vous corrigez?
7. Tenez-vous compte du vocabulaire dans l'évaluation des productions écrites? Si oui, de quelle façon (on relève, on pénalise)?
 - a. La plupart des grilles de correction proposent une évaluation globale du vocabulaire qui repose sur les critères de richesse, de variété et de précision du vocabulaire. Selon vous, que signifie « avoir un vocabulaire riche, varié et précis »? Quels sont les indices qui vous permettent de poser un jugement sur le vocabulaire utilisé par vos élèves? En d'autres mots, comment se manifestent la richesse, la variété et la précision du vocabulaire?
 - b. Trouvez-vous facile d'évaluer le vocabulaire de vos élèves?
8. De quelle façon notez-vous les erreurs de vocabulaire sur les copies de vos élèves (correction/suggestion, commentaire, renvoi au dictionnaire, soulignement, etc.)?
 - a. Croyez-vous que vos annotations/corrections aident les élèves à améliorer leur vocabulaire?
9. Est-ce que les erreurs de vocabulaire sont fréquentes dans les rédactions de vos élèves?
 - a. Quels types d'erreurs de vocabulaire sont fréquents chez vos élèves? Avez-vous des exemples qui vous viennent à l'esprit?
 - b. En fait, qu'est-ce qu'une erreur de vocabulaire pour vous?

- c. Les rectifications orthographiques sont de plus en plus diffusées. Comment en tenez-vous compte dans votre correction?
10. Faites-vous un suivi lorsque vous avez remis un texte corrigé aux élèves? Doivent-ils retravailler leur texte?
- a. Si c'est le cas, est-ce que les élèves doivent corriger leurs erreurs de vocabulaire?

LES OUVRAGES DE RÉFÉRENCE

11. Utilisez-vous des dictionnaires pour corriger les erreurs de vocabulaire de vos élèves?
- a. Quel est votre dictionnaire préféré? Pourquoi?
- b. Quel usage faites-vous de ce dictionnaire? À quoi sert-il?
12. Quels sont les dictionnaires disponibles dans votre classe?
- a. Vos élèves utilisent-ils le dictionnaire? De quelle façon? En tirent-ils profit?
- b. On parle beaucoup d'anglicismes au Québec. Êtes-vous sensible à l'emploi d'anglicismes?
- c. Que faites-vous des québécismes?

2^e volet

Voici sept erreurs tirées de productions écrites d'élèves de troisième secondaire. Nous allons les regarder une à une et je vais vous poser quelques questions pour chacune.

- Les auriez-vous relevées?
- Est-ce que vous auriez tenu compte de ça dans l'évaluation?
- Où auriez-vous classé cette erreur?
- Comment vous expliqueriez à l'élève que c'est erreur?
- Quelle correction proposez-vous?
- Selon vous, pourquoi l'élève a-t-il commis cette erreur?

LES ERREURS

1. *L'horloger se dirigea **à** la caverne à grande vitesse.*
2. *Aux pieds du volcan Asama **trônait** un village.*
3. *Les publicitaires **portent** une très mauvaise influence sur les jeunes dans la société.*
4. *Dyoll devait absolument passer. Il a trouvé un arbre très grand et l'a couper comme du beurre avec la hach magic. L'arbre **s'est affalé** sur le ruisseau.*
5. *James n'était pas trop convaincu, mais il essaya **pareil**.*
6. *Un premier job d'été peut s'avérer un très bon tremplin pour tes autres **travails** dans le futur.*
7. *Dans **une immeuble** très **luxueuse**, les deux hommes discutent tranquillement.*

ANNEXE 6

LISTE DES CODES UTILISÉS DANS *QDA MINER*

LISTE DES ÉTIQUETTES POUR CODAGE DES ENTREVUES

- Enseignant :
 - caractéristiques de l'enseignants : Caract_ens
 - conception de la *maitrise de la langue* : Concep_maitrise_langue
 - conception de l'EL : Concep_EL
 - conception liée au vocabulaire : Concep_Voc
- Correction
 - grille de correction utilisée : Grille_Corr
 - suivi après correction : Suivi_Corr
 - suivi après correction (EL) : Suivi_Corr_EL
 - annotation : Corr_Annotation
 - prise en compte du vocabulaire dans la correction : Voc_Corr_Note
 - priorités de correction : Priorite_Corr
 - attribution d'une note pour la rédaction : Redaction_Note
 - attribution d'une note pour le vocabulaire : Voc_Note
 - contribution des annotations à l'amélioration du vocabulaire : Annot_Amelio_Voc
 - difficulté de la tâche d'évaluation du vocabulaire : Facilite_Corr_Voc
- Erreurs_Eleves
 - description du vocabulaire des élèves : Voc_eleves
 - erreurs les plus fréquentes chez les élèves : Err_frequentes
 - EL les plus fréquentes chez les élèves : EL_frequentes
- Pratiques
 - pratiques d'enseignement du vocabulaire : Pratiques_Ens_Voc
 - pratiques d'écriture : Pratique_Ecriture
- Divers
 - utilisation du dictionnaire (enseignant) : Dico_Ens
 - utilisation du dictionnaire (élève) : Dico_Eleve
 - anglicismes : Anglicismes
 - rectifications orthographiques : Rectif_Orth
 - québécoismes : Quebecismes
- Explication_Err
 - erreur 1 : Err1
 - erreur 2 : Err2
 - erreur 3 : Err3
 - erreur 4 : Err4
 - erreur 5 : Err5
 - erreur 6 : Err6
 - erreur 7 : Err7

