



HAL
open science

Le vocabulaire médical du point de vue des trois fonctions primaires.

Karim Chebouti

► **To cite this version:**

Karim Chebouti. Le vocabulaire médical du point de vue des trois fonctions primaires.. Linguistique. Université Paris-Nord - Paris XIII, 2014. Français. NNT : 2014PA131006 . tel-01233280

HAL Id: tel-01233280

<https://theses.hal.science/tel-01233280v1>

Submitted on 24 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS 13

« U.F.R. LETTRES, SCIENCES DE L'HOMME ET DES SOCIÉTÉS »



THÈSE DE DOCTORAT

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PARIS 13

EN SCIENCES DU LANGAGE

LE VOCABULAIRE MÉDICAL DU POINT DE VUE

DES TROIS FONCTIONS PRIMAIRES

Présentée et soutenue publiquement par

Karim CHEBOUTI

Le 08 avril 2014

DIRECTEUR DE THÈSE

M. Pierre-André BUVET

JURY

M. Salah MEJRI, Université Paris 13

M. Xavier BLANCO, Université Autonome de Barcelone

M. André DUGAS, Université du Québec à Montréal

M. Pierre-André BUVET, Université Paris 13

UNIVERSITÉ PARIS 13

« U.F.R. LETTRES, SCIENCES DE L'HOMME ET DES SOCIÉTÉS »



THÈSE DE DOCTORAT

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PARIS 13

EN SCIENCES DU LANGAGE

LE VOCABULAIRE MÉDICAL DU POINT DE VUE

DES TROIS FONCTIONS PRIMAIRES

Présentée et soutenue publiquement par

Karim CHEBOUTI

Le 08 avril 2014

DIRECTEUR DE THÈSE

M. Pierre-André BUVET

JURY

M. Salah MEJRI, Université Paris 13

M. Xavier BLANCO, Université Autonome de Barcelone

M. André DUGAS, Université du Québec à Montréal

M. Pierre-André BUVET, Université Paris 13

A mes parents

*« La vie, c'est comme une bicyclette,
il faut avancer pour ne pas perdre l'équilibre. »*
Albert Einstein

Remerciements

Mes premiers remerciements sont adressés à mon directeur de recherche Monsieur Pierre-André Buvet. Je le remercie de fond de mon cœur d'avoir accepté de diriger mes travaux de recherche. Ses conseils sur la thèse et sur la linguistique en général ont été très bénéfiques pour moi. Sa disponibilité m'a permis d'avancer et de résoudre toutes les difficultés liées à la thèse.

J'adresse mes sincères remerciements et ma gratitude au professeur Salah Mejri pour m'avoir accueilli au laboratoire LDI pour préparer ma thèse. Il a mis à ma disposition des moyens inégalés pour travailler dans de bonnes conditions. Je le remercie également pour ses nombreux conseils et sa sympathie.

Je remercie Monsieur André Dugas et Monsieur Joël Eline pour leurs relectures de la thèse et leurs nombreux conseils.

Je remercie également les doctorants du laboratoire avec qui j'ai partagé beaucoup de moments, parfois difficiles, mais chaleureux. Les discussions enrichissantes que nous avons partagées m'ont incité à réfléchir à d'autres problématiques linguistiques et d'autres perspectives. Je remercie Adila Ameer, Lichao Zhu, Zina Sibachir, Marco Fasciolo, Kangho Lee, Motoyuki Miyagawa, Marie-Anatolie Jaozandry, Ndeye Fatou Thiaw, Pham Thi Nhung..., ainsi que les étudiants du Master TILDE.

Je remercie les membres du laboratoire pour leurs conseils et leur sympathie. J'adresse mes remerciements à Ahmed Elchheb pour son aide en informatique. Je remercie également Aude Grezka, Fabrice Issac, Xavier-Laurent Salvador, Arabelle Baudette, et Maëlle Le Bihan pour leur sympathie.

J'exprime ma gratitude et mes remerciements à toute ma famille qui m'a beaucoup soutenu dans cette démarche. Je remercie particulièrement mes parents ainsi que Boussad et Nabila pour m'avoir encouragé.

TABLE DES MATIERES

LISTE DES SYMBOLES ET ABREVIATIONS.....	XIV
INTRODUCTION.....	ERREUR ! SIGNET NON DEFINI.
CHAPITRE 1.....	18
PROBLÉMATIQUE.....	18
Introduction	18
1. La langue de la médecine du point de vue des trois fonctions primaires.....	18
1.1. Unité d'analyse.....	18
1.1.1. Le mot : unité d'analyse traditionnelle.....	19
1.1.2. Réfutation du mot comme unité d'analyse.....	19
1.1.3. La phrase : autre unité d'analyse.....	20
1.2. Les niveaux discursifs.....	21
1.2.1. Niveau logico-sémantique.....	21
1.2.2. Niveau énonciatif.....	22
1.2.3. Niveau interprétatif.....	23
1.3. Notion d'emploi.....	24
1.3.1. Généralité sur la notion d'emploi.....	24
1.3.2. La notion d'emploi prédicatif.....	24
1.3.3. La notion d'arguments.....	25
2. De la langue générale à la langue spécialisée.....	26
2.1. Traitement commun.....	27
2.1.1. Traitement de la polysémie.....	27
2.1.2. Traitement du figement.....	27
2.1.3. Traitement de la polymorphie.....	28
2.2. Traitement spécifique à la médecine.....	28
2.2.1. Complexité référentielle.....	29
2.2.2. Complexité énonciative.....	29
2.2.3. Complexité textuelle.....	30
2.3. Articulation entre langue générale et langue de la médecine.....	30
2.3.1. Analyse morphologique.....	31
2.3.2. Analyse syntaxique.....	31
2.3.3. Analyse sémantique.....	32
3. Dictionnaires électroniques.....	33
3.1. Le dictionnaire des prédicats.....	33
3.1.1. Choix des mots.....	33
3.1.2. Traitement.....	34
3.1.3. Hypothèses.....	34

3.2. Le dictionnaire des arguments.....	35
3.2.1. Choix des mots.....	35
3.2.2. Traitement.....	35
3.2.3. Hypothèses.....	36
3.3. Le dictionnaire des humains.....	36
3.3.1. Choix des mots.....	37
3.3.2. Traitement.....	37
3.3.3. Hypothèses.....	38
Conclusion.....	39
CHAPITRE 2.....	40
ÉTAT DE L'ART.....	40
Introduction.....	40
1. Etudes morphologiques.....	40
1.1. Mots composés.....	40
1.1.1. Les composés savants.....	41
1.1.2. Les composés non savants.....	43
1.1.3. Ellipse.....	43
1.2. Autres mots construits.....	44
1.2.1. Siglaison.....	45
1.2.2. Acronymie.....	47
1.2.3. Dérivation.....	47
1.3. Productivité lexicale.....	49
1.3.1. Néologie.....	49
1.3.2. Emprunt.....	50
1.3.3. Éponymie.....	52
2. Etudes sémantiques.....	53
2.1. Analyse lexicale.....	53
2.1.1. De l'hégémonie gréco-latine à l'émergence des langues médicales nationales.....	53
2.1.2. La scientificité.....	54
2.1.3. L'universalisation du vocabulaire médical.....	55
2.2. Analyse syntactico-sémantique.....	56
2.2.1. Polysémie du vocabulaire médical.....	56
2.2.2. Actualisation et noms médicaux.....	58
2.2.3. Appropriation et noms médicaux.....	59
2.3. Analyse de discours.....	60
2.3.1. Problèmes stylistiques.....	60
2.3.2. Langue générale et langue spécialisée.....	62
2.3.3. Dialogues médecin-patient.....	63
3. Études applicatives.....	64
3.1. Traduction.....	65
3.1.1. Didactique de la traduction médicale.....	65

3.1.2. Outils pour la traduction médicale	66
3.1.3. Difficultés de la traduction médicale.....	67
3.2. Recherche documentaire	68
3.2.1. Ressources documentaires.....	68
3.2.2. Moteurs de recherche	70
3.2.3. Bases de données médicales : l'exemple de MeSH.....	71
3.3. TAL	73
3.3.1. Le vocabulaire de la médecine et le traitement automatique.....	74
3.3.2. Quels outils pour le traitement des textes médicaux ?.....	75
3.3.3. Les principales difficultés.....	77
Conclusion.....	79
CHAPITRE 3.....	80
MÉTHODOLOGIE	80
Introduction	80
1. Cadre théorique : modèle de données	80
1.1. Outils d'analyse.....	80
1.1.1. La phrase élémentaire.....	81
1.1.2. Le postulat	81
1.1.3. Principes d'analyse.....	82
1.2. La théorie des trois fonctions primaires	82
1.2.1. Fonction prédicative	82
1.2.2. Fonction argumentale	83
1.2.3. Fonction actualisatrice.....	84
1.3. La notion d'emploi	84
1.3.1. Niveaux linguistiques.....	84
1.3.2. Descripteurs de définition.....	86
1.3.3. Descripteurs de condition	88
2. Corpus : collection de données.....	91
2.1. Constitution du corpus.....	91
2.1.1. Dictionnaires	91
2.1.2. Web	92
2.1.3. Littérature	92
2.2. Types de corpus.....	93
2.2.1. Les termes spécialisés	93
2.2.2. Les termes non spécialisés	93
2.2.3. Les termes moyennement spécialisés.....	94
2.3. Sources	94
2.3.1. Bases de données.....	94
2.3.2. Forums.....	95
2.3.3. Revues	95
3. Dictionnaire : structuration et description des données.....	96

3.1. Structuration.....	96
3.1.1. Dictionnaires électroniques.....	96
3.1.2. Macrostructure.....	97
a) des noms prédicatifs.....	97
b) des noms argumentaux.....	98
c) des verbes prédicatifs.....	98
d) des adjectifs prédicatifs.....	98
3.1.3. Microstructure.....	99
3.2. PRED_DIC.....	99
3.2.1. Particularités du dictionnaire.....	99
3.2.2. Noms et Adjectifs.....	102
3.2.3. Verbes.....	103
3.3. ARGU_DIC, HUM_DIC et ENCY_DIC.....	104
3.3.1. ARGU_DIC.....	104
3.3.2. HUM_DIC.....	106
3.3.3. ENCY_DIC.....	107
Conclusion.....	108
CHAPITRE 4.....	109
LES EMPLOIS PRÉDICATIFS DE LA LANGUE DE LA MÉDECINE.....	109
Introduction.....	109
1. Propriétés morphologiques.....	109
1.1. Dérivation.....	110
1.1.1. Dérivation suffixale.....	110
1.1.2. Dérivation préfixale.....	112
1.1.3. Dérivation impropre.....	114
1.2. Composition.....	115
1.2.1. La composition savante.....	115
1.2.2. La composition non savante.....	117
1.2.3. La composition mixte.....	118
1.3. Autres mots construits.....	118
1.3.1. Siglaison.....	119
1.3.2. Acronymie.....	120
1.3.3. Troncation et mots valises.....	120
2. Propriétés sémantiques.....	121
2.1. Racines prédicatives.....	121
2.1.1. Prédicats à forme unique.....	122
2.1.2. Prédicats à forme nominale et adjectivale.....	122
2.1.3. Prédicats à forme nominale et verbale.....	123
2.1.4. Prédicats à forme verbale, nominale et adjectivale.....	124
2.2. Types et classes sémantiques.....	126
2.2.1. Type état/action/événement.....	126
2.2.2. Classes sémantiques.....	128

2.3.3. Aspect inhérent.....	130
2.3. Les emplois prédicatifs.....	131
2.3.1. Les emplois nominaux.....	131
a. La classe maladie.....	132
b. Les sous-classes maladies	133
b. 1. Maladies diverses	133
b. 2. Maladies contagieuses.....	134
b. 3. Maladies chroniques.....	134
b. 4. Inflammation	136
b. 5. Symptôme.....	136
c. La classe attaques	137
d. La classe accidents	137
e. La classe soins	138
2.3.2. Les emplois verbaux.....	139
a. Verbes spécialisés.....	139
b. Verbes de la langue générale.....	140
c. Verbes de la langue générale très utilisés en médecine	141
2.3.3. Les emplois adjectivaux	142
3. Propriétés combinatoires	143
3.1. Construction	143
3.1.1. Construction des verbes prédicatifs.....	143
3.1.2. Construction des adjectifs prédicatifs.....	144
3.1.3. Construction des noms prédicatifs.....	144
3.2. Combinatoire interne	145
3.2.1. Séquences opaques	145
3.2.2. Séquences transparentes	146
3.2.3. Détermination standard	148
3.3. Combinatoire externe	149
3.3.1. Les verbes.....	149
3.3.2. Les adjectifs.....	150
3.3.3. Détermination non standard	151
Conclusion.....	152
CHAPITRE 5.....	153
LES NOMS ARGUMENTAUX DU DOMAINE DE LA MÉDECINE	153
Introduction	153
1. Propriétés morphologiques.....	153
1.1. Noms simples	154
1.1.1. Noms spécialisés	155
1.1.2. Noms non spécialisés	156
1.1.3. Noms transdisciplinaires	157
1.2. Noms dérivés.....	158

1.2.1. Dérivation préfixale.....	158
1.2.2. Dérivation suffixale.....	159
1.2.3. Troncation.....	160
1.3. Noms composés.....	161
1.3.1. Les composés savants.....	161
1.3.2. Composés non savants.....	162
1.3.3. Composés mixtes.....	163
2. Propriétés sémantiques.....	165
2.1. Noms de nature biologique.....	166
2.1.1. Parties du corps.....	166
2.1.2. Micro-organismes.....	168
2.1.3. Liquides organiques.....	170
2.2. Matériel médical.....	171
2.2.1. Instruments.....	171
2.2.2. Appareils d'observation.....	172
2.2.3. Appareils de mesure.....	174
2.3. Locatifs.....	174
2.3.1. Lieux.....	175
2.3.2. Parties de lieux.....	176
2.3.3. Emplois métonymiques.....	177
3. Propriétés combinatoires.....	178
3.1. Verbes prédicatifs.....	178
3.1.1. Verbes spécialisés.....	178
3.1.2. Verbes non spécialisés.....	180
3.1.3. Actualisation des noms élémentaires lorsqu'ils se combinent avec des verbes prédicatifs.....	182
3.2. Adjectifs prédicatifs.....	183
3.2.1. Adjectifs spécialisés.....	184
3.2.2. Adjectifs non spécialisés.....	184
3.3. Noms prédicatifs.....	185
3.3.1. Noms spécialisés vs noms non spécialisés.....	186
3.3.2. Noms simples vs noms composés.....	187
a) noms simples.....	188
b) noms composés.....	188
Conclusion.....	189
CHAPITRE 6.....	190
LES EMPLOIS DES NOMS D'HUMAINS DU DOMAINE MÉDICAL.....	190
Introduction.....	190
1. Propriétés morphologiques.....	191
1.1. Les noms simples.....	191
1.1.1. Les noms spécialisés.....	191

1.1.2. Les noms non spécialisés	192
1.1.3. Les noms transdisciplinaires.....	193
1.2. Les noms dérivés.....	194
1.2.1. Dérivation suffixale.....	194
1.2.2. Dérivation préfixale.....	195
1.2.3. Dérivation impropre	195
1.3. Autres noms construits	196
1.3.1. Composition	196
1.3.2. Troncation	197
1.3.3. Ellipse.....	199
2. Propriétés sémantiques	200
2.1. Humains prédicatifs.....	200
2.1.1. La prédicativité des noms d'humains	200
2.1.2. Fonction prédicative	201
2.1.3. Fonction argumentale	202
2.2. Les humains argumentaux.....	203
2.2.1. Fonctionnement global	203
2.2.2. Fonctionnement particulier.....	204
2.2.3. Les dérivés des noms propres.....	205
2.3. Humains collectifs.....	205
2.3.1. Noms non ambigus.....	206
2.3.2. Noms ambigus.....	206
2.3.3. Autres métonymies.....	207
3. Classifications sémantiques et combinatoires.....	208
3.1. Profession et fonction.....	208
3.1.1. Noms appropriés.....	208
3.1.2. Verbes appropriés.....	209
3.1.3. Adjectifs appropriés	210
3.2. Malades	211
3.2.1. Noms appropriés.....	211
3.2.2. Verbes appropriés.....	212
3.2.3. Adjectifs appropriés	213
3.3. Autres noms d'humains.....	214
3.3.1. Noms appropriés.....	214
3.3.2. Verbes appropriés.....	215
3.3.3. Adjectifs appropriés	216
Conclusion.....	216
CHAPITRE 7.....	217
DICTIONNAIRE POLYFONCTIONNEL.....	217
Introduction	217
1. Structuration des données lexicales	217

1. 1. Bases de données.....	217
1.1.1. Définition des bases de données.....	218
1.1.2. Les différentes bases de données.....	218
1.1.3. Notre choix : PHP SQL.....	219
1.2. Tableaux CSV et Excel	220
1.2.1. Définition des fichiers Excel et CSV.....	220
1.2.2. Conversion des fichiers Excel vers CSV.....	221
1.2.3. Exploitation des fichiers CSV	222
1.3. Fichiers XML	222
1.3.1. Définition des fichiers XML	222
1.3.2. Exploitation	223
1.3.3. Les inconvénients du XML	224
2. Description extensive.....	225
2.1. Diversité métalinguistique et extralinguistique	225
2.1.1. Informations métalinguistiques et diachronie	225
2.1.2. Informations métalinguistiques et synchronie	226
2.1.3. Informations extralinguistiques.....	226
2.2. Sources	227
2.2.1. Sources lexicographiques	228
2.2.2. Sources provenant des corpus	228
2.2.3. Créations personnelles.....	229
2.3. Mode de structuration.....	229
2.3.1. Formalisation métalexographique	230
2.3.2. Dictionnaires coordonnés et non coordonnés.....	231
2.3.3. Dictionnaires généraux et spécialisés.....	232
3. Exploitation lexicographique	232
3.1. Utilisations diverses	233
3.1.1. Dictionnaire monolingue.....	233
3.1.2. Dictionnaire historique	234
3.1.3. Dictionnaire plurilingue.....	234
3.2. Présentation globale	235
3.2.1. Morfetik.....	236
3.2.2. TLFi.....	236
3.2.3. Dictionnaires de synonymes.....	237
3.3. Présentation ciblée.....	238
3.3.1. Terminologique	238
3.3.2. Historique	239
3.3.3. Par niveaux de langue.....	240
Conclusion.....	240
CHAPITRE 8.....	242
APPLICATION : INTERFACE D'INTERROGATION DU LANGAGE DE LA MÉDECINE.....	242

Introduction	242
1. La base de données.....	242
1. 1. Le langage SQL.....	242
1.1.1. Définition.....	243
1.1.2. Utilisation	243
1.1.3. Justification du choix.....	244
1.2. Structure de la base de données.....	244
1.2.1. Définition de la base.....	244
1.2.2. Structure de la base.....	245
1.2.3. Fonctionnement de la base	246
1.3. Structure des tables.....	246
1.3.1. Table relative aux prédicats.....	246
1.3.2. Table relative aux arguments.....	248
1.3.3. Table relative aux humains.....	248
1.3.4. Table relative au dictionnaire encyclopédique	248
2. Interface d'interrogation	249
2.1. Le langage PHP.....	249
2.1.1. Définition.....	249
2.1.2. Fonctionnement.....	250
2.1.3. Justification du choix.....	252
2.2. Le langage HTML / CSS.....	252
2.2.1. Définition du HTML	252
2.2.2. Définition du CSS	253
2.2.3. Utilisation	253
2.3. Les formulaires.....	254
2.3.1. Définition.....	254
2.3.2. Structuration des formulaires	255
2.3.3. Fonctionnement des formulaires	256
3. Mode d'emploi.....	257
3.1. Requête sur les prédicats	258
3.1.1. Présentation du formulaire	258
3.1.2. Présentation de la requête.....	260
3.1.3. Script PHP	260
3.2. Requête sur les arguments.....	261
3.2.1. Présentation du formulaire	261
3.2.2. Présentation de la requête.....	262
3.2.3. Script PHP	262
3.3. Requête sur les humains	263
3.3.1. Présentation du formulaire	264
3.3.2. Présentation de la requête.....	264
3.3.3. Script PHP	265
3.4. Requête sur le dictionnaire encyclopédique.....	265
3.4.1. Présentation du formulaire	266
3.4.2. Présentation de la requête.....	266
3.4.3. Script PHP	267

Conclusion.....	268
CHAPITRE 9.....	269
PRÉSENTATION DES RÉSULTATS.....	269
Introduction	269
1. Accueil	270
2. Contact	271
3. Dictionnaires linguistiques	272
3.1. Dictionnaire des prédicats	273
3.1.1. Requête.....	273
3.1.2. Résultat.....	275
3.2. Dictionnaire des arguments	277
3.2.1. Requête.....	277
3.2.2. Résultat.....	278
3.3. Dictionnaire des humains	280
3.3.1. Requête.....	280
3.3.2. Résultat.....	281
4. Dictionnaire encyclopédique	282
4.1. Requête.....	282
4.2. Résultat.....	283
Conclusion.....	284
CONCLUSION GÉNÉRALE	285
BIBLIOGRAPHIE	289
INDEX DES TERMES	302

LISTE DES SYMBOLES ET ABREVIATIONS

Adj	adjectif
Adv	adverbe
Det	déterminant
Hum	humain
eta	état
eve	événement
act	action
N	nom
V	verbe
Vsup	verbe support
Prep	préposition
X0	argument 1
X1	argument 2
X2	argument 3
PRED_DIC	dictionnaire des prédicats
ARGU_DIC	dictionnaire des arguments
HUM_DIC	dictionnaire des noms d'humains
ENCY_DIC	dictionnaire encyclopédique
=	désigne l'équivalence
≠	désigne la non-équivalence
*	désigne l'incompatibilité
Pred	prédicat
Arg	argument
Act	actualisateur

INTRODUCTION

L'information médicale apparaît comme un objet pluridimensionnel du fait de la diversité des discours qu'elle engendre : publications scientifiques, documents pédagogiques, textes de vulgarisation, campagnes promotionnelles ou sanitaires à visée préventive ou thérapeutique, communiqués de politique de santé, écrits juridiques, etc. Par ailleurs, sa prise en charge par les nouveaux médias liés à l'évolution du Web (flux rss, blog, forum, Twitter, etc.) donne lieu à de nouveaux types de discours qui la rendent également plus complexe.

La compréhension des textes de la médecine passe par celle de son vocabulaire, cela implique l'utilisation des ressources et des outils d'analyse linguistique qui ne se limitent pas à des listes terminologiques, c'est-à-dire des listes de termes associés univoquement à des concepts. Les descriptions à l'origine des ressources linguistiques sont souvent minimales ; il s'agit le plus souvent d'informations morphosyntaxiques relatives aux parties du discours et d'informations sémantiques importées d'ontologies. Elles ne permettent pas d'étudier d'une manière satisfaisante des phénomènes linguistiques fondamentaux dans la perspective du traitement automatique des langues. Pour remédier à ces défaillances, nous proposons d'exploiter des ressources linguistiques qui intègrent plus de descripteurs, notamment les descripteurs de définition et les descripteurs de condition (P. A. Buvet 2009).

La thèse traite de l'information médicale et étudie plusieurs phénomènes linguistiques afin d'expliquer leur fonctionnement et leurs spécificités pour faciliter la compréhension du discours médical.

Cette étude s'inscrit dans le cadre des travaux de recherche menés au laboratoire LDI. Elle nous a conduit à élaborer quatre dictionnaires électroniques de la langue de la médecine, en l'occurrence le dictionnaire des prédicats, le dictionnaire des arguments, le dictionnaire des humains et le dictionnaire encyclopédique.

L'analyse du vocabulaire médical s'inscrit dans la lignée des travaux qui ne distinguent pas les discours spécialisés des discours de la langue générale quant à leurs principaux modes de fonctionnement linguistiques (Gautier 2000, G. Gross et M. Mathieu-Colas 2001, Z. Harris 1988 et P. Lerat 1995). Elle a comme champ d'expérimentation le domaine de la médecine. Nous montrons que la langue de la médecine fonctionne comme la

langue générale du point de vue syntactico-sémantique. Pour ce faire, nous proposons d'analyser systématiquement plusieurs phénomènes linguistiques dans les discours médicaux : la polysémie prédicative (par exemple, *prendre* équivaut à *mesurer* dans *prendre le pouls* et à *ingérer* dans *prendre un traitement*) ; le figement (*corps géniculé médial*, *scissure de Sylvius*, etc.) et la collocation (par exemple, *elle accouche (une mère + *un père)* ; la polymorphie prédicative (*anesthésier*, *anesthésiant*, *anesthésique*) et argumentale (*front*, *frontal*). Les descriptions et les classifications effectuées dans le cadre de ces analyses nous ont permis de constituer des ressources lexicales pour le Traitement Automatique des Langues.

Cette étude a été l'occasion pour nous d'effectuer un classement des termes médicaux selon les trois catégories suivantes : les emplois prédicatifs, les noms argumentaux et les noms d'humains médicaux. Chaque catégorie a donné un dictionnaire électronique spécifique pour le traitement automatique.

La thèse a été également l'occasion d'exploiter les ressources (dictionnaires électroniques) pour élaborer une interface d'interrogation qui permet de donner des informations linguistiques et encyclopédiques en rapport avec chaque terme des dictionnaires. Les informations linguistiques associées aux termes dans les dictionnaires sont d'ordre morphologique, syntaxique et sémantique. Les informations encyclopédiques sont formulées sous la forme de classes sémantiques, de domaines et de sous-domaines.

Notre thèse est constituée de neuf chapitres. Ils seront présentés de la façon suivante : Dans le premier chapitre, nous abordons la problématique de la thèse. Elle s'articule autour du fonctionnement syntactico-sémantique de la langue de la médecine du point de vue de l'articulation entre la langue générale et les langues spécialisées et sur l'utilisation des dictionnaires électroniques pour le traitement automatique.

Dans le deuxième chapitre, nous présentons l'état de l'art sur le vocabulaire médical. Il comportera trois parties, la première est consacrée aux études morphologiques, la deuxième aux études sémantiques et la troisième aux études applicatives du vocabulaire médical. Cette dernière comporte les études sur la traduction, sur la recherche documentaire et sur le traitement automatique de la langue de la médecine.

Le troisième chapitre est consacré à la méthodologie que nous avons suivie pour parvenir à notre objectif. Il s'agit de la méthodologie du laboratoire LDI. Elle porte la théorie

des trois fonctions primaires, notamment la fonction prédicative, la fonction argumentale et la fonction actualisatrice, puis la notion de la phrase élémentaire qui sert du socle pour l'analyse des mots et enfin les corpus et les dictionnaires électroniques.

Le quatrième chapitre est consacré à l'analyse morphologique, sémantique et syntaxique des prédicats de médecine.

Dans le chapitre cinq, nous analysons les noms argumentaux du vocabulaire médical selon leurs propriétés morphologiques, sémantiques et combinatoires.

Le chapitre six traite de la morphologie, de la syntaxe et de la sémantique des noms d'humains médicaux. Nous traitons les noms d'humains selon qu'ils ont une fonction prédicative ou une fonction argumentale, et selon qu'ils appartiennent à la catégorie des humains médecins ou à celle des humains malades. Ainsi, leur propriétés syntaxiques sont différentes, notamment quand il s'agit de la combinatoire avec les verbes appropriés.

Dans le chapitre sept, nous présentons les outils informatiques qui nous permettent d'élaborer une interface d'interrogation¹, ensuite nous présentons le dictionnaire selon la diversité métalinguistique, les ressources et les modes de structuration, enfin nous présentons son exploitation et son mode de fonctionnement.

Le chapitre huit présente l'interface d'interrogation, son mode de structuration, les outils exploités pour sa réalisation et son mode d'emploi, tout en sachant que chaque dictionnaire a son emploi particulier.

Le chapitre neuf est consacré à la présentation des résultats obtenus avec notre base de données de chaque dictionnaire, suivis de commentaires.

¹ Cela nous permettra de choisir un outil pour élaborer notre interface d'interrogation par nos dictionnaires électroniques.

CHAPITRE 1

PROBLÉMATIQUE

Introduction

Notre thèse a pour objet l'étude des prédicats, des arguments et des humains de la langue de la médecine. Nous partons du principe que l'unité minimale d'analyse est la phrase simple. Les mots hors contexte ont des sens vagues. C'est dans le cadre de la phrase que leurs sens se précisent. Nous adoptons une méthodologie et un traitement communs à ceux utilisés pour la description de la langue générale parce que leurs principaux modes de fonctionnement sont identiques (P. Lerat 1995), afin d'étudier certains phénomènes linguistiques comme la polysémie, la polymorphie, le figement, l'inférence, etc. Néanmoins, le discours médical a certaines spécificités, par exemple son appartenance à un domaine spécialisé, ce qui engendre une certaine complexité énonciative et textuelle².

1. La langue de la médecine du point de vue des trois fonctions primaires

Nous définissons la phrase comme l'unité minimale d'analyse linguistique, puis, nous étudions les trois niveaux d'analyse de la phrase (le niveau logico-sémantique, le niveau énonciatif et le niveau interprétatif), et enfin nous évoquons la notion d'emploi des unités linguistiques, notamment l'emploi prédicatif et l'emploi argumental.

1.1. Unité d'analyse

Toute méthode d'analyse et de description des mécanismes langagiers qui explique le fonctionnement d'une langue opte pour le choix d'une unité minimale d'analyse. Dans la tradition terminologique, c'est l'étude du mot qui prime dans l'analyse linguistique. Mais cette approche n'est pas satisfaisante car les mots prennent leurs sens en contexte, c'est-à-dire dans la phrase (G. Gross & M. Mathieu-Colas 2001).

² Le discours médical n'est pas accessible aux non initiés.

1.1.1. Le mot : unité d'analyse traditionnelle

La notion de mot est parmi les phénomènes linguistiques les plus étudiés et demeure sans doute le plus complexe. Une revue trimestrielle (Mot) dont le thème diffère dans chaque numéro lui est d'ailleurs consacrée. Par exemple, le numéro 15 Octobre 1987 est consacré à la dénomination, sous le titre « comment nommer ? ». Le numéro 39 Juin 1994 est consacré aux mots de l'environnement et de l'écologie.

André Martinet, dans *Éléments de linguistique générale* (1970) distingue, dans la double articulation, le phonème du monème qu'il considère comme l'unité minimale ayant du sens. Le monème est à son tour divisé en deux catégories, le monème de la grammaire (morphème) et le monème du lexique (lexème).

Dans les *Cours de linguistique générale* (1967) de Ferdinand De Saussure, on évoque la notion de mot sous un autre angle. On considère que toute unité linguistique (mot) est dotée d'un signifiant, qui « désigne l'image acoustique du mot » et d'un signifié qui « désigne le concept, c'est-à-dire la représentation mentale d'une chose ».

Nous constatons que, traditionnellement, l'analyse linguistique était centrée sur le mot. Il est considéré comme l'unité fondamentale de la langue.

1.1.2. Réfutation du mot comme unité d'analyse

Du point de vue de la théorie des trois fonctions primaires, nous considérons que les mots doivent être étudiés en contexte, c'est-à-dire dans des phrases où ils prennent leurs sens. Considérer en syntaxe les mots comme des entités isolées, dotées de formes acoustiques et de sens est une aberration. Le fonctionnement des unités lexicales se fait dans la phrase où elles entretiennent des relations syntaxiques et sémantiques (F. Cusin-Berche 2003). Par exemple, *livre* porte un sens différent, selon qu'il est précédé par un pronom personnel : *il livre la marchandise*, ou un déterminant : *il a acheté un livre*. Cet exemple montre à quel point les mots peuvent varier en emploi. Des exemples plus complexes montrent qu'il est difficile de distinguer les différents sens des mots :

Il traite la vache de la fermière

Il traite la fermière de vache

Dans ces exemples, le verbe *traiter* est ambigu. Dans le premier cas, il s'agit de soins, dans le second, d'une insulte (P. A. Buvet 2009).

Dans d'autres exemples, l'ambiguïté est encore plus importante :

La belle porte le voile

Dans l'exemple ci-dessus, presque tous les mots de la phrase sont ambigus. Autrement dit, la phrase peut avoir deux interprétations différentes. Dans le premier cas, nous pouvons l'interpréter de la façon suivante : *la belle (fille) porte le voile*. Dans le second, nous pouvons l'interpréter comme : *la belle porte (entrée) voile ce qui se trouve dans l'autre pièce*. Dans le premier cas, *porte* est un verbe et *voile* est un nom, tandis que dans le second, *porte* est un nom et *voile* est un verbe (G. Gross & M. Mathieu-Colas 2001).

En guise de conclusion, citons F. Cusin-Berche (2003) : « un avocat ne peut pas être à la fois à la vinaigrette et à la cour... »

1.1.3. La phrase : autre unité d'analyse

La phrase est l'unité minimale d'analyse. Elle est constituée de plusieurs unités linguistiques. Chaque unité doit être définie en contexte, en rapport avec les autres éléments de la phrase avec lesquels elle entre en relation. Dissocier les éléments linguistiques constituant la phrase et les étudier, chacun de manière isolée, leur donnerait des sens vagues et imprécis, sauf dans le cas des termes scientifiques monosémiques. Certes, les verbes *disséquer*, *hospitaliser*, *intuber*, etc. n'ont pas besoin d'être insérés dans des phrases pour que l'on puisse décrire leurs sens respectifs. Néanmoins, les vocabulaires des langues de spécialité sont constitués, en majorité, de mots de la langue générale.

Puisque la plupart des mots sont polysémiques, ils ne doivent pas être traités en dehors du cadre de la phrase. Prenons par exemple le verbe *traiter* ; son sens change selon le contexte :

Il traite le patient

Il traite le dossier

Il le traite de ...

Dans le premier cas, le verbe *traiter* signifie *soigner*, dans le deuxième, *étudier* et dans le troisième, *insulter*. Ainsi, pour connaître le sens du verbe *traiter*, il faut l'insérer dans une phrase, dans laquelle les classes d'arguments jouent un rôle fondamental dans la désambiguïsation des termes polysémiques.

1.2. Les niveaux discursifs

Il existe trois niveaux d'analyse de la phrase : le niveau logico-sémantique, le niveau énonciatif et le niveau interprétatif.

1.2.1. Niveau logico-sémantique

L'analyse syntactico-sémantique de la phrase est basée sur la structure prédicat/argument. Selon Z. Harris (1976), les prédicats et les arguments doivent entretenir des relations logico-sémantiques dans toute phrase. C'est-à-dire que si l'on s'appuie sur l'axe paradigmatique défini par Saussure (1967), la liste des paradigmes est ouverte, mais elle pose des contraintes d'ordre logico-sémantique. Par exemple, le sujet de prescrire une ordonnance peut se combiner avec tous les noms d'humains de médecine habilités à exécuter cette action.

Cette notion peut s'expliquer par la combinatoire des termes dans le cadre de la phrase. Il s'en suit que chaque terme inséré dans une phrase doit se combiner avec les autres éléments de la même phrase, sans mettre en péril le sens de celle-ci.

L'infirmier panse la blessure

Le chirurgien suture la plaie

**Le plombier panse le tuyau*

**Le serrurier suture la porte*

Nous constatons que la liste des noms qui se combinent avec les verbes *panse* et *suture* est ouverte, mais il y a un blocage lorsqu'il s'agit du non sens. Ainsi, il n'y a pas de compatibilité entre certains noms et les verbes ci-dessus. Autrement dit, seules les personnes ayant des capacités et éventuellement des diplômes requis peuvent appliquer des pansements

sur des plaies ou sur des blessures. Cela s'applique notamment à la langue générale. Le verbe *manger* ne se combine qu'avec les êtres vivants ayant la faculté de manger des aliments, sauf dans le cas de la métaphore où les orateurs cherchent des effets stylistiques.

1.2.2. Niveau énonciatif

L'analyse de la phrase, selon la structure prédicat-argument, consiste à distinguer les éléments de la phrase selon leurs fonctions, la fonction prédicative, la fonction argumentale et la fonction actualisatrice (S. Mejri 2009). La fonction prédicative et la fonction argumentale concerne la construction de la phrase et la relation du prédicat avec les arguments qu'il sélectionne, comme nous l'avons montré au niveau logico-sémantique. La fonction actualisatrice concerne l'actualisation du prédicat et des arguments (X. Banco & P. A. Buvet 2009).

Le niveau énonciatif concerne tous les éléments qui inscrivent la phrase dans une situation d'énonciation. Ils prennent en charge les notions de temps, d'espace et de modalité, etc. Une phrase non actualisée perd beaucoup d'informations importantes relatives à la situation de communication.

Paul faire AVC

Paul a fait un AVC hier

La phrase ci-dessus, lorsqu'elle n'est pas actualisée, apporte une information principale concernant l'accident vasculaire cérébral fait par Paul, mais le passé composé nous apporte une information supplémentaire qui est le passé et *hier* précise le moment de l'*AVC*. Nous pouvons toujours apporter une information supplémentaire par un actualisateur comme dans :

Paul a fait un AVC hier à midi

Avec un actualisateur, nous pouvons rajouter une information supplémentaire par rapport au lieu :

Paul a fait un AVC hier à midi, chez lui

1.2.3. Niveau interprétatif

Il ne s'agit pas ici de voir le sens des phrases hors contexte et surtout, en dehors des situations de communication. Selon Robert Martin (1983), seul le contexte peut rendre certaines phrases synonymes. Par exemple, si on dit « *il fait froid* », on veut dire « *tu pourrais fermer la fenêtre* ».

Le sens se construit dans une situation de communication. Il n'existe pas dans l'absolu et hors contexte (P. Charaudeau 1992). Autrement dit, le sens et le résultat d'une activité langagière se concrétisent dans l'intention de transmettre des messages.

Linguistiquement, les énoncés, hors contexte, expriment des vérités absolues. Nous montrons cela dans les phénomènes d'inférence et de paraphrase :

Il est hospitalisé
= *Il est malade*

Ainsi, les exemples ci-dessus sont en relation inférentielle. Dans le cas contraire, on met en question la véracité de l'énoncé « *il est hospitalisé* ». Donc, on peut dire : *il est hospitalisé parce qu'il est malade*, ou : *il est malade puisqu'il est hospitalisé*. Par contre, les deux phrases qui sont en juxtaposition : *il est hospitalisé, il n'est pas malade* sont en contradiction.

La paraphrase est un énoncé qui a le même sens que l'énoncé initial. La définition de la paraphrase suffit pour montrer que si le premier énoncé est vrai, le second est vrai, et si le premier est faux, le second est faux aussi (R. Martin 1983).

Il a la jaunisse
Il souffre d'un ictère

L'interprétation des énoncés en contexte est différente. Dire *il est malade ce type de quelqu'un qui prend des risques énormes* ne signifie pas qu'il a une maladie quelconque, mais qu'*il met sa vie en danger*.

Il est malade ce type
= *Il met sa vie en danger*
≠ *Il a une maladie*

Il est important de mentionner que l'énoncé est considéré depuis le contexte dans lequel il prend un sens ; par contre, la phrase hors contexte sert de support pour comprendre l'énoncé.

1.3. Notion d'emploi

L'une des tâches les plus importantes de notre travail est l'étude des emplois de chaque prédicat de la langue de la médecine et de ses arguments afin de dresser la liste des phrases possibles qui peuvent être construites autour de chaque emploi prédicatif.

1.3.1. Généralité sur la notion d'emploi

Les principes d'analyse développés au laboratoire Lexique Dictionnaires Informatique, s'appuient sur trois facteurs importants en linguistique : le lexique, la syntaxe et la sémantique (I. Sfar 2007).

Si l'on tient compte du principe qui postule que chaque phrase est constituée d'un prédicat et des arguments, et que les mots s'analysent dans le cadre de la phrase, nous constatons que la plupart des prédicats sont polysémiques. Si la polysémie est considérée comme le fait qu'un mot peut avoir plusieurs sens, la notion d'emploi est plus complexe et plus subtile. Un emploi prédicatif dépend du domaine d'argument du prédicat.

1.3.2. La notion d'emploi prédicatif

Les mots ont des emplois différents d'une phrase à une autre. Les mots *voile* et *cache* ne peuvent pas être interprétés de la même façon lorsqu'ils sont précédés d'un article ou d'un groupe nominal.

Il cache son visage

Il se voile la face

Il enlève le cache du moteur

Il fait de la voile

Dans les deux premiers cas, il s'agit de verbes, dans les deux derniers cas, de substantifs. Cette ambiguïté paraît évidente et simple pour un humain, mais ne l'est pas pour un ordinateur, d'où la nécessité d'élaborer des dictionnaires électroniques pour décrire le vocabulaire médical, dans la perspective du traitement automatique. Cet exemple montre clairement que l'étude et la description des mots doit se faire dans leur contexte phrastique. Dans les langues spécialisées, il y a des verbes propres à un domaine, comme *hospitaliser* ou *cautériser* pour ce qui est de la langue de la médecine. Il y a également des verbes polysémiques dont l'un des emplois, au moins, appartient à une langue spécialisée. Si l'on prend l'exemple du verbe *voir* dans *voir un médecin* et *voir un ami*, le premier emploi, qui équivaut à *consulter*, est propre au domaine de la médecine.

Les exemples *voile* et *cache* montrent que les mots peuvent être interprétés comme des substantifs ou comme des verbes. La notion d'emploi, parfois, concerne seulement les verbes. Le verbe *prendre* signifie *ingurgiter* dans *prendre du paracétamol*, *boire* dans *prendre de l'eau*, *boire de l'alcool* dans *prendre un verre*, *manger* dans *prendre un sandwich*, *écrire* dans *prendre des notes*, *voyager* dans *prendre un train*.

Nous constatons qu'à chaque fois qu'on change les arguments, le prédicat change d'emploi. Le vocabulaire médical n'échappe pas à cette caractéristique des prédicats, puisqu'il est formé, en majorité, des mots de la langue générale.

1.3.3. La notion d'arguments

Nous ne pouvons pas parler de l'emploi prédicatif sans tenir compte des arguments avec lesquels il se combine. L'un des rôles de la description des arguments est de réduire la polysémie. La nature sémantique des arguments aide à la reconnaissance de chaque emploi prédicatif. Si nous reprenons l'exemple du verbe *prendre*, nous comprenons qu'il faut à chaque fois chercher la nature sémantique de l'argument en position de complément afin de déduire l'emploi prédicatif. Le verbe *prendre* ne peut signifier *ingurgiter* si l'on ne définit pas *paracétamol* comme appartenant à la classe des *médicaments*. Il en est de même pour la classe

des *aliments* pour *sandwich*, *boisson* pour *eau*, *alcool* pour *un verre*, *informations* pour *notes* et *moyen de transport* pour *le train*.

Les classes d'arguments servent à réduire l'ambiguïté des prédicats, mais nous constatons que les mêmes prédicats ne se combinent pas systématiquement avec tous les arguments de la même classe. Ainsi, le verbe *prendre* ne se combine pas avec tous les noms de médicaments.

Il a pris du paracétamol

Il a ingurgité du paracétamol

Il a pris une pommade

**Il a ingurgité une pommade*

Le verbe *prendre* signifie « ingurgiter » lorsqu'il a *paracétamol* en position de complément, mais il ne signifie pas « ingurgiter » lorsqu'il s'agit d'une *pommade*, tout en sachant que *paracétamol* et *pommade* appartiennent à la classe (médicament). Dans ce cas, il serait important de préciser la sous-classe à laquelle l'argument appartient.

2. De la langue générale à la langue spécialisée

Nous considérons que les langues de spécialités n'ont aucune caractéristique spécifique par rapport à la langue générale (P. Lerat 1997³). Ainsi, toutes les théories linguistiques s'accordent sur le fait que les langues spécialisées ne sont pas des simples nomenclatures (*ibid*). Dans cette perspective, nous proposons d'étudier la langue de la médecine avec les mêmes outils utilisés pour la langue générale.

Dans cette partie, nous montrons le traitement commun à la langue de la médecine et à la langue générale, ensuite, nous nous intéressons au traitement spécifique du vocabulaire médical et enfin nous traitons l'articulation entre la langue générale et la langue de la médecine.

³ Selon Lerat « Aucune théorie linguistique, quelle qu'elle soit, n'a jamais isolé le fonctionnement des langues spécialisées de celui des langues naturelles en général ».

2.1. Traitement commun

Si l'on considère les langues spécialisées comme étant des nomenclatures, il serait impensable d'étudier les phénomènes linguistiques tels que la polysémie, le figement et la polymorphie, dans le vocabulaire médical. Cela est aujourd'hui faisable parce que les terminologues ne considèrent plus les langues spécialisées comme des listes de mots mais des langues ayant un fonctionnement semblable à celui de la langue générale.

2.1.1. Traitement de la polysémie

La polysémie n'est pas un phénomène linguistique réservé uniquement à la langue générale. Le phénomène de la polysémie est très fréquent dans le vocabulaire médical. Ce dernier est composé de beaucoup de termes de la langue générale comme *prendre*, *suivre*, *consulter*, etc. Par exemple, *prendre* équivaut à *mesurer* dans *prendre le pouls* et à *ingérer* dans *prendre un traitement*, *suivre* équivaut à *soigner* dans *le médecin suit un patient* et à *marcher derrière* dans *l'enfant suit sa mère* et *consulter* équivaut à *faire un examen médical* dans *le médecin examine un patient* et à *regarder* dans *il consulte sa montre*.

Nous constatons que le phénomène de la polysémie est très fréquent dans le vocabulaire médical. Ne pas le prendre en considération engendrerait beaucoup de problèmes d'ambiguïté dans l'analyse du discours médical.

2.1.2. Traitement du figement

Le vocabulaire médical a connu un essor considérable grâce à la diversification et au foisonnement de son vocabulaire qui résulte du développement qu'a connu la médecine ces dernières années. Le besoin de dénommer de nouveaux objets, appareils, instruments et techniques de soins, etc., a eu un énorme impact sur la linguistique. La composition fait partie des procédés de formation les plus exploités dans les langues spécialisées (G. Gross 1996). La langue de la médecine, dont on voit l'évolution constante du vocabulaire, a beaucoup recours à la composition de termes pour former de nouvelles unités significatives.

Nous étudions trois sortes de noms composés ; les composés savants, les composés non savants et les composés mixtes.

Nous observons, en premier lieu, que les noms composés ont des degrés de figement différents : le nom composé *dent de sagesse* est plus transparent que *pomme d'Adam*, *culotte de cheval* ou *corps géniculé médical*.

Le figement lexical n'est pas négligeable en langue de spécialité, notamment dans la langue de la médecine. Son étude facilite la compréhension du vocabulaire médical qui est déjà très complexe.

2.1.3. Traitement de la polymorphie

L'étude relative à la structure prédicat-argument nécessite le traitement de la polymorphie. Il s'agit de l'étude des formes appartenant à une même racine, qu'elle soit prédicative comme dans *il respecte ses parents*, *il est respectueux de ses parents* et *il a le respect envers ses parents* ou argumentale comme dans *il a une maladie du cœur*, *il a une maladie cardiaque*. Dans le vocabulaire médical, une même racine, comme *anesth-*, peut donner plusieurs formes différentes, notamment la forme verbale *anesthésier*, la forme nominale *anesthésie* et la forme adjectivale *anesthésiant*.

La polymorphie ne concerne pas uniquement la catégorie des prédicats, mais aussi les arguments. Certains arguments ont une forme nominale et une forme adjectivale comme *front*, *frontal*, *cheveu*, *chevelu*, *intestin*, *intestinal*, etc.

2.2. Traitement spécifique à la médecine

Les langues de spécialité ont certaines spécificités qui les différencient de la langue générale. Ainsi, le vocabulaire médical appartient à un domaine et des sous-domaines propres à la santé. Il est caractérisé par une complexité énonciative. Les spécialistes et les non spécialistes ne partagent pas les mêmes discours à cause de la complexité textuelle du vocabulaire médical. Il est doté d'une terminologie spécialisée, alors que les personnes non initiées utilisent plutôt les mots de la langue générale pour désigner des phénomènes compliqués liés aux noms de maladies.

2.2.1. Complexité référentielle

Certaines notions extralinguistiques, introduites dans les dictionnaires électroniques sont notamment utiles dans la désambiguïsation des unités lexicales, dans les dictionnaires électroniques qui sont destinés au traitement automatique de la langue. Parmi celles-ci, il y a les indications de domaine et de sous-domaine. Le rattachement variable d'un mot à un domaine en fonction du degré de précision que l'on attribue à ce type d'information explique la distinction entre les champs domaine et sous-domaine (P. A Buvet & M. Mathieu-Colas 1999). Par exemple, le mot *virus* peut être considéré par les lexicographes comme relevant du domaine de la médecine, ou plus précisément, de la microbiologie. Dans les dictionnaires électroniques, nous adoptons les deux premiers codes qui donnent les informations sur les mots, de type domaine et sous-domaine.

La notion de classe sémantique sur laquelle s'appuient les dictionnaires électroniques dans la description des unités lexicales est primordiale pour réduire la polysémie, mais les notions de domaine et de sous-domaine demeurent aussi indispensables. Les noms *bistouri*, *stéthoscope* et *toise bébé* ont en commun d'appartenir à la classe d'objet (matériel médical) et diffèrent seulement par leur appartenance respective aux sous-domaines de la chirurgie, de la médecine générale et de la maternité/ pédiatrie.

Les indications de domaines d'emploi figurant dans les dictionnaires électroniques peuvent aussi concourir, dans le cadre d'analyse automatique des textes, à désambiguïser des unités lexicales.

2.2.2. Complexité énonciative

Si la langue générale est plus ou moins accessible à tout le monde, la langue de la médecine, elle, est plutôt réservée aux spécialistes. Ils utilisent leur propre vocabulaire scientifique et technique. Les patients, qui ne possèdent pas nécessairement des connaissances dans le domaine médical, ne sont pas en mesure de comprendre le discours médical parce qu'il est scientifique et contient beaucoup de mots techniques (P. Faure 2010).

Cela pose beaucoup de problèmes de compréhension dans un dialogue médecin/patient. Pour se faire comprendre, les praticiens de la santé utilisent souvent les mots

de la langue générale lorsqu'ils communiquent avec leurs patients, comme on le constate dans les forums sur la santé.

Les praticiens de la santé utilisent une terminologie qui devrait satisfaire à la fois la curiosité des patients voulant avoir des renseignements et des précisions sur leurs maladies, et les exigences scientifiques de la langue de la médecine.

En français, le patient ne maîtrise pas forcément les termes techniques, sauf s'il est initié. Il utilise plutôt les mots « saignement de nez » au lieu d'*épistaxis*, « maladie de colon » pour *colopathie*. Les mots techniques sont réservés aux spécialistes du domaine.

2.2.3. Complexité textuelle

Dans cette étude, nous sommes confrontés à un vocabulaire ayant deux niveaux de langue, un niveau soutenu réservé aux spécialistes et aux chercheurs et un niveau non soutenu utilisé plutôt par le public en général et les patients en particulier. Cette variation nous impose de nous occuper à la fois des textes spécialisés, tirés des magazines, des revues spécialisés et des thèses, et des publications de vulgarisation, tirés des revues de vulgarisation du vocabulaire médical ainsi que des forums sur la santé et le bien-être.

Le vocabulaire médical est un vocabulaire scientifique. Son analyse nécessite de s'occuper des deux registres de langue, tout en sachant que l'un est constitué des termes techniques et scientifiques et l'autre des termes de la langue générale qui servent de synonymes approximatifs des termes techniques. Néanmoins, la terminologie scientifique et la terminologie générale sont toutes les deux destinées à l'expression de phénomènes scientifiques, parfois extrêmement compliquées comme les maladies ou les molécules des médicaments. Par exemple, *la leucémie* s'appelle *cancer de sang* chez les non initiés qui s'expriment sur les forums que nous avons consultés, comme *esanté.fr*.

2.3. Articulation entre langue générale et langue de la médecine

Nous étudions la langue de la médecine du point de vue morphologique, syntaxique et sémantique.

2.3.1. Analyse morphologique

L'analyse morphologique des prédicats consiste à présenter, dans le dictionnaire, toutes les formes que peut avoir un prédicat donné. La première étape consiste à définir leur mode de formation (dérivation, composition, etc.), puis à donner toutes les formes possibles d'une même racine. Ainsi, nous obtenons jusqu'à trois prédicats de la même racine prédictive : forme verbale *cautériser*, forme nominale *cautérisation* et forme adjectivale *cautérisé*.

Le médecin cautérise la plaie

Le médecin procède à la cautérisation de la plaie

La plaie est cautérisée

Certaines racines prédictives ne permettent de former qu'une forme d'un emploi prédictif. Un prédicat est dit autonome lorsqu'il correspond à un emploi unique. L'emploi nominal du prédicat *carcinome* est la seule instance de ce prédicat dans la phrase :

Il a un carcinome

2.3.2. Analyse syntaxique

Le prédicat est l'élément qui établit la relation dans la phrase. Il se combine avec les arguments. Les relations qui existent entre les prédicats et les arguments sont d'ordre hiérarchique. Le nombre d'arguments qu'un prédicat sélectionne varie d'un type de prédicat à un autre.

Les prédicats monadiques entretiennent une relation logique avec un seul argument qu'ils prédominent et qu'ils catégorisent. Ainsi dans *il se blesse*, *il a guéri* et *il est malade*, les prédicats *se blesser*, *guérir* et *être malade*, sont rattachés à un seul argument *il* en tant que sujet, en position de X0.

Les prédicats dyadiques s'accordent avec deux arguments ; leurs relations peuvent être, soit unidirectionnelles comme dans *Paul aime Marie*, soit interactionnelles comme dans *Paul et Marie s'aiment*. Le premier argument est sujet, en position de X0 et le deuxième est complément, en position de X1.

Les prédicats triadiques apparaissent dans des phrases plus longues, c'est-à-dire qu'ils s'accordent avec trois arguments. Par exemple *le médecin ampute Paul de sa jambe*. Le premier fonctionne comme sujet en position X0, le deuxième est complément en position X1 et le troisième est complément en position X2.

2.3.3. Analyse sémantique

Les prédicats sont caractérisés par leurs classes sémantiques. Leur appartenance à une classe ou à une sous-classe est fondée sur la synonymie et les domaines d'arguments communs.

Les emplois prédicatifs polysémiques partagent une même forme mais ils ne se rapportent pas aux mêmes prédicats car ils ont des caractéristiques sémantiques et syntaxiques différentes. C'est le cas du verbe *consulter* :

Paul consulte un médecin

Le médecin consulte un dictionnaire

La nature des compléments rend compte de la polysémie de *consulter* : le verbe est quasi synonyme de *interroger* lorsqu'il est suivi de noms de (profession médicale), quasi synonyme de *lire* lorsqu'il est suivi de noms de (texte). Les deux emplois de *consulter* se rapportent à deux prédicats distincts : un prédicat de (procédure médicale) et un prédicat de (lecture).

Cet exemple montre la complexité de la sémantique, non seulement dans la langue générale, mais aussi dans le vocabulaire spécialisé, notamment la langue de la médecine dans laquelle on utilise énormément de verbes polysémiques. Leurs emplois sont discriminés par les classes et les sous-classes d'arguments.

3. Dictionnaires électroniques

Les dictionnaires seront présentés sous forme de tableaux qui comportent différentes données relatives à la microstructure et à la macrostructure des dictionnaires que nous avons élaborés. Les données et les informations qu'elles comportent sont différentes d'un dictionnaire à l'autre.

Un dictionnaire électronique est un dictionnaire contenant du lexique. Contrairement au dictionnaire traditionnel qui utilise le support papier, le contenu du dictionnaire électronique est stocké sur un CD-ROM ou sur internet. Il présente à l'utilisateur plusieurs avantages comme la grande capacité du volume de stockage, la simplicité d'utilisation et l'accès facile à l'information et la recherche du lexique (A. L. Jousse & O. Tremblay 2006).

3.1. Le dictionnaire des prédicats

Le dictionnaire des prédicats contient des mots ayant des emplois prédicatifs. Il est constitué d'une façon plus ou moins exhaustive. Le choix des mots n'est pas exclusif à une classe ou un domaine donné, par contre, leur organisation dans le dictionnaire est soumise à des exigences qui répondent aux besoins du traitement automatique.

3.1.1. Choix des mots

La macrostructure du dictionnaire (PRED_DIC) est constituée de noms, de verbes et d'adjectifs qui sont employés comme des prédicats ; autrement dit, ils ont au moins un emploi prédicatif. Il s'agit de mots simples (*cancer*), de mots composés, (*hernie discale*) ou de mots dérivés (*gingivite*).

Pour notre étude, nous avons opté pour des mots distincts et variés. Nous avons collecté des mots techniques et hautement spécialisés comme *ictère*, *splénite*, *splénectomie*, etc.), des mots moyennement spécialisés comme *nausée*, *cancer*, et *grippe*, etc. ainsi que des non spécialisés tels que *jaunisse* et *tour de reins*.

Les classifications utilisées dans le dictionnaire des prédicats correspondent en grande partie aux ontologies médicales. Nous avons exploité celles de l'arborescence MeSH⁴.

3.1.2. Traitement

Les classes sémantiques des arguments sont constituées de noms élémentaires, c'est-à-dire des noms qui fonctionnent uniquement comme des arguments. L'élaboration d'une classe de prédicats s'effectue selon le domaine d'arguments, afin de traiter les problèmes d'ambiguïté liés aux prédicats (*ibid*). La constitution d'une classe de prédicats s'effectue aussi sur la base de la synonymie et de l'hyponymie (P. A. Buvet 2009). Lorsqu'une classe sémantique est constituée sur la base de la synonymie, elle est définie à partir d'un prédicat exemplaire. Il s'agit du prédicat le plus représentatif de la classe, les autres prédicats sont considérés comme périphériques (P. A. Buvet & A. Grezka 2007).

Dans le cas de l'hyponymie, les classes sémantiques des prédicats sont présentées sous des formes nominales, en position sujet et en position de complément du nom, par exemple : *le diabète est une maladie*.

Lorsque les prédicats n'ont pas tous une forme nominale, la notion d'anaphore infidèle permet de justifier la classe (G. Kleiber 1991). Par exemple, la classe (soin dentaire) est constituée de prédicats comme *brosser, nettoyer, cureter, faire un détartrage*. Avec l'anaphore infidèle *ces soins dentaires*, on peut construire des phrases simples qui permettent de reprendre la classe des (soins dentaires). Par exemple : *Luc (brosse + nettoie + curete + faire un détartrage de) ses dents*.

3.1.3. Hypothèses

La description des prédicats de médecine consiste à démontrer leur fonctionnement ainsi que leurs éventuels emplois, dans la langue de la médecine. Pour chaque prédicat donné, nous donnons sa description morphologique afin de traiter les problèmes liés aux racines prédictives, sa description sémantique afin d'étudier son emploi prédictif et sa description combinatoire dans le but de réduire son ambiguïté.

⁴ MESH est une liste de termes qui fonctionne sous forme d'arborescence. Elle est utilisée pour l'analyse documentaire dans le domaine biomédical.

3.2. Le dictionnaire des arguments

Le dictionnaire des arguments sera consacré aux noms élémentaires. Ces derniers fonctionnent uniquement comme arguments dans des phrases et jamais comme prédicats. Leur rôle est secondaire par rapport à celui des prédicats, mais ils sont indispensables sur le plan syntaxique et sémantique.

3.2.1. Choix des mots

La macrostructure du dictionnaire (ARGU_DIC) est constituée de noms élémentaires. Ils ont toujours une fonction argumentale. Le corpus regroupe tous les noms du domaine de la médecine classés en sous-domaine comme l'anatomie (*parties du corps, liquide organique*), et d'autres disciplines périphériques ou annexes comme la biologie ou la pharmacologie. Ce vaste choix n'est pas fortuit puisque la médecine est fortement liée à ce type de disciplines. Prenons l'exemple des *virus*, des *bactéries* et des *champignons* ; ils appartiennent au domaine de la microbiologie, mais sont néanmoins fréquemment employés en médecine parce qu'ils sont souvent à l'origine de beaucoup de maladies et d'infections.

Certes, la pharmacologie est un domaine autonome, mais il entretient des liens étroits avec la médecine ; la pharmacologie existe grâce à la médecine qui a besoin de médicaments pour les soins. Cela intervient dans la pratique quotidienne de la profession des deux cotés. Un pharmacologue ne peut pas négliger la médecine et le médecin ne peut pas négliger la pharmacologie. En résumé, on peut dire *tel médicament est bon pour telle maladie, telle maladie nécessite tel médicament*.

3.2.2. Traitement

Conformément aux principes de la notion de classes d'objets, nous commençons par la spécification et la définition sémantique des entrées du dictionnaire des arguments. Par la suite, l'objectif est de trouver, à chaque fois, le prédicat approprié. Ainsi, nous considérons qu'un ou plusieurs prédicats donnés sont nécessaires et surtout suffisants pour caractériser la classe (M. Mathieu-Colas 1998). Autrement dit, chaque prédicat ou groupe de prédicats définit une classe donnée. Cela se fait par discrimination, c'est-à-dire que c'est le prédicat qui sélectionne les arguments d'une classe donnée.

Le muscle se déchire, se froisse

L'os casse

Les veines gonflent

Ces exemples montrent clairement que les verbes *déchirer*, *froisser*, *casser* et *gonfler* ne se combinent pas de manière exclusive avec tous les noms des parties du corps. Par contre, chaque verbe se combine à un mot qui appartient à une classe ou à une sous-classe définie, comme la classe des *partie du corps os*, *liquide organique*, *partie du corps cellule*, etc.

En contrepartie, les arguments ont un rôle majeur dans la définition de l'emploi prédicatif. C'est-à-dire que le prédicat porte un sens selon la classe des arguments qui l'entourent. Autrement dit, ils réduisent l'ambiguïté des prédicats.

Nous appliquons aux arguments un codage qui fait en sorte qu'on distingue systématiquement les classes et les sous-classes, dans le but d'appliquer automatiquement les mêmes prédicats appropriés à toutes les entrées de la même classe.

3.2.3. Hypothèses

Nous définissons, d'abord, les classes d'arguments, grâce aux verbes prédicatifs. Chaque classe d'arguments doit être regroupée et catégorisée par des prédicats appropriés à la classe.

Avec la description des prédicats appropriés, nous obtiendrons un nombre défini de classes qui seront catégorisées par les prédicats appropriés.

3.3. Le dictionnaire des humains

Nous tenons à expliquer, en premier lieu, le choix d'élaborer un dictionnaire électronique destiné exclusivement au traitement des noms d'humains. Sachant que les noms d'humains peuvent fonctionner soit comme des prédicats (*Paul est chirurgien*), soit comme des arguments (*le chirurgien a opéré Paul*), ils peuvent donc préalablement apparaître dans le dictionnaire des prédicats pour les premiers et dans le dictionnaire des arguments pour les

seconds. Toutefois, ce choix à été effectué après avoir constaté que les noms d'humains font partie des catégories de noms qui sont les moins traitées dans ce domaine.

3.3.1. Choix des mots

La macrostructure du dictionnaire des humains est constituée des noms d'humains utilisés dans le vocabulaire médical. Ils sont classés selon leur morphologie, leur classe sémantique ainsi que les domaines et les sous-domaines auxquels ils appartiennent. Le choix des noms d'humains s'est focalisé sur le principe de l'exhaustivité. D'ailleurs, dans le dictionnaire des humains médicaux, nous pouvons constater que les noms sont organisés et structurés selon les différents domaines de la médecine comme la chirurgie, la pédiatrie, la neurologie, la chirurgie dentaire, etc. Les classes sémantiques des noms d'humains sont principalement structurées par classement général comme *humain malade* et *humain médecin* ou par classement précis comme *humain anesthésiste*, *humain chirurgien*, *humain neurologue*, etc.

Certains noms d'humains de la langue générale sont recensés dans le dictionnaire des humains, tels que *père*, *mère*, *adulte*, *mère porteuse*, etc. Cela vient de la fréquence d'utilisation de ces noms dans le vocabulaire médical. Par exemple, on évoque les noms *père* et *mère* et *mère porteuse* dans le traitement des maladies génétiques, ou les mots *adulte*, *adolescence* et *enfant* dans le but d'adapter les traitements médicaux à chaque catégorie de personnes.

3.3.2. Traitement

Le traitement des noms d'humains consiste à les prendre en charge dans les deux fonctions distinctes. Il s'agit de la fonction prédicative et de la fonction argumentale. Certains noms d'humains peuvent fonctionner comme prédicats :

Il est chirurgien

Il est médecin

Ils peuvent fonctionner aussi comme des arguments.

Le chirurgien travaille aujourd'hui

Le médecin exerce aujourd'hui

Dans les deux premiers cas, les noms *chirurgien* et *médecin* fonctionnent comme des prédicats. Par contre, dans les deux derniers exemples, ils fonctionnent comme des arguments. Les verbes *travailler* et *exercer* fonctionnent comme des prédicats.

D'autres noms d'humains de la langue de la médecine fonctionnent toujours comme des arguments comme les noms propres (*Ambroise Paré, Avicenne, Robert Debré, etc.*)

Cette variation dans le fonctionnement des noms d'humains en médecine nécessite de les distinguer, premièrement, s'ils sont plutôt des noms communs ou des noms propres. (Les premiers peuvent avoir, soit un emploi prédicatif, soit un emploi argumental). Ensuite, il faut distinguer les emplois des noms communs qui peuvent être variables, d'une phrase à l'autre.

Contrairement aux noms qui figurent dans le dictionnaire des arguments qui sont toujours des arguments, puisqu'ils sont des noms élémentaires, ceux du dictionnaire des humains peuvent avoir des emplois différents.

3.3.3. Hypothèses

Le statut des humains est particulier du fait qu'ils peuvent fonctionner comme des prédicats ou comme des arguments dans le cadre de la phrase. Dans le dictionnaire des humains, nous montrons les différentes fonctions des noms d'humains de la langue de la médecine et les conditions dans lesquelles ils ont une fonction prédicative, les conditions dans lesquelles ils ont une fonction argumentale.

Nous montrerons ici, avec les outils linguistiques, la nature des classes sémantiques de chaque argument pour définir les emplois prédicatifs. L'utilisation des types sémantiques des arguments faciliterait la discrimination des différents emplois.

Conclusion

Nous constatons que le vocabulaire médical est analysable avec les mêmes outils linguistiques exploités dans l'étude de la langue générale. Cela est possible du fait que les deux discours (de la langue général et de la langue de la médecine) ont une structuration linguistique similaire.

L'objectif de notre étude est de rendre compte du fonctionnement de la langue de la médecine dans le cadre de la théorie des trois fonctions primaires.

CHAPITRE 2

ÉTAT DE L'ART

Introduction

La langue de la médecine a fait l'objet de nombreuses études. Nous présenterons l'approche morphologique de la langue de la médecine, puis l'approche sémantique qui englobe l'analyse lexicale, l'analyse syntactico-sémantique ainsi que l'analyse de discours, et enfin l'approche applicative, en rapport avec la traduction et le traitement automatique des langues.

1. Etudes morphologiques

Dans les études morphologiques, on aborde les questions qui se rapportent à la composition savante et mixte ainsi qu'à l'ellipse. On examinera également d'autres mots construits, notamment, la siglaison, l'acronymie et la dérivation, et enfin on abordera la productivité lexicale. Il s'agit pour cette dernière de la néologie, de l'emprunt et de l'éponymie.

1.1. Mots composés

Le vocabulaire de la médecine est très complexe à cause de son origine gréco-latine. Cette complexité s'observe au niveau morphologique et concerne les mots composés. Ils sont difficilement compréhensibles si l'on ignore la morphologie et l'étymologie du vocabulaire médical.

Pour décrire de façon précise les locutions, il faut tout d'abord s'intéresser à leur construction interne. Le vocabulaire médical comporte toutes sortes de locutions, notamment les noms composés qui ont des constructions différentes et variées, mais qui répondent aux normes grammaticales de la combinatoire.

La composition nominale est très importante dans le vocabulaire médical. Elle est le fait d'assembler deux ou plusieurs mots simples ou composés pour former de nouveaux substantifs (J-C. Sournia1997).

Nous mentionnons les travaux portant sur les composés savants et les composés non savants, puis des études qui traitent de la question de l'ellipse, c'est-à-dire des phénomènes d'abréviation et de troncation des mots composés.

De nombreux linguistes ont étudié la diversité du vocabulaire médical et la liberté de formation des noms composés qui permettent de dénommer les maladies, les parties du corps, les organes, les examens médicaux, les opérations chirurgicales, etc. Nous présentons ici les principaux résultats de leurs études.

1.1.1. Les composés savants

Les mots composés savants sont le plus souvent des locutions nominales. Ils constituent une grande partie du vocabulaire médical. Les mots composés dits savants résultent de la juxtaposition de deux ou plusieurs mots d'origine grecque ou latine.

Il existe deux types de mots composés savants. Ceux du premier type sont soudés comme c'est le cas de la plupart des spécialités de la médecine.

Il y a une abondance de composés de type NN (nom + nom), c'est-à-dire, formés de deux bases nominales. Chacun des exemples ci-dessous est formé de deux termes significatifs et indépendants.

<i>Composés</i>	<i>mot simple</i>	<i>sens</i>	<i>mot simple</i>	<i>sens</i>
<i>traumatologie</i>	<i>traumato</i>	<i>blessure</i>	<i>logie</i>	<i>science</i>
<i>neurologie</i>	<i>neuro</i>	<i>système nerveux</i>	<i>logie</i>	<i>science</i>
<i>dermatologie</i>	<i>dermato</i>	<i>peau</i>	<i>logie</i>	<i>science</i>
<i>urologie</i>	<i>uro</i>	<i>urine</i>	<i>logie</i>	<i>science</i>
<i>pathologie</i>	<i>pathos</i>	<i>maladies</i>	<i>logie</i>	<i>science</i>

Ceux du second type sont formés de deux termes autonomes, significatifs et séparés par des blancs dans le cas de *fusariumpoae*, *fusariumsolani*. Il s'agit d'un mode de composition facile et très fréquent qui consiste à juxtaposer deux termes pour désigner une nouvelle unité significative par exemple : *candidose cutanée*, *candidose buccale*, *céphalée primitive*, *arythmie sinusale*, *atrésie pulmonaire*, etc.

De nombreux mots composés de la terminologie médicale comportent un trait d'union. Il est parfois nécessaire et parfois inutile (J-C. Sournia 1997).

Selon H. Joly et C. Murcia (2005), il est nécessaire de garder le trait d'union dans le cas où le premier élément de la séquence est un mot complet comme *crimée-congo*, *criger-najjar*, *epstein-barr virus*. Le trait d'union est inutile lorsque le premier élément formant le mot composé est une préposition (*à*, *en*, *sous*, *sus*, *anté*, *rétro*, etc.), par exemple *suspubien*, sauf quelques exceptions dans le but de respecter la prononciation comme *sous-claviculaire*, *pré-éclampsie*, *anté-hypophyse*.

Lorsque le premier élément se termine par une voyelle qui sert souvent de moyen de liaison (« o » en général), on supprime le trait d'union parce qu'elle assure la soudure des deux termes, à l'instar de *cardiovasculaire*, *anatomopathologie*, *auriculoventriculaire*, etc.

Le trait d'union reste néanmoins indispensable lorsque deux voyelles distinctes se rencontrent et doivent être prononcées toutes les deux, c'est le cas de *génito-urinaire* (J. Chevallier 1987).

1.1.2. Les composés non savants

Les mots composés non savants du vocabulaire médical sont très souvent des locutions nominales formées avec des mots du français courant.

De son côté, J-C. Sournia (1997) soutient que cette apposition est ancienne en français, mais que son développement est récent. Elle utilise deux mots ayant chacun leur sens, les unit pour désigner un nouveau concept. *La santé publique, le bloc opératoire, un centre hospitalier, un praticien hospitalier et l'assurance maladie*, etc. sont des « unités définissables ».

Il ajoute que la composition nominale est la juxtaposition de deux substantifs sans intermédiaire, « ni même le trait d'union qui s'impose de moins en moins » alors qu'il était fréquent. On cite comme exemple en médecine : *médecin conseil, effet retard, insuline retard, cellule hôte*, etc.

La composition nominale a donné lieu à la formation d'un nombre important de locutions nominales en médecine, notamment dans le cas des noms de maladies par rapport aux organes qu'elles touchent comme *carie dentaire, atrésie intestinale, atrésie pulmonaire, inflammation cutanée*, etc. comme on utilise tout simplement l'adjectif de relation pour évoquer l'origine de la maladie : *grippe aviaire, grippe porcine*, etc. ou toute autre motivation.

On utilise parfois ce mode de composition qui est relativement facile pour éviter les termes intermédiaires explicatifs inutiles. Ainsi *risque alcool* est l'équivalent de « risque auquel sont exposés les consommateurs excessifs de boissons alcoolisées ».

1.1.3. Ellipse

Le vocabulaire médical se caractérise par des constructions elliptiques. Elles sont relativement importantes. G. Skvortsov (1994) a étudié 2000 termes médicaux (essentiellement dans le domaine de l'oto-rhino-laryngologie). Il a montré que de nombreuses constructions elliptiques sont de type N A (nom et adjectif) dans lesquelles le substantif est systématiquement omis, par exemple : *le facial* pour *le nerf facial*, *l'acoustique* pour *le nerf acoustique*, *le temporal* pour *l'os temporal*, *la carotide* pour *l'artère carotide*, *la parotide* pour *la glande parotide*, etc.

L'analyse de ces exemples montre clairement que le phénomène de l'ellipse concerne les groupes nominaux les plus fréquents et les plus connus dans le domaine de l'ORL. Il s'agit souvent de la chute d'une partie du nom composé désignant une partie du corps comme *nerf*, *os*, *muscle*, *glande*, qui sont très connus ou d'une maladie. Dans beaucoup de cas, les adjectifs qui subsistent après l'ellipse ne s'emploient qu'avec les substantifs éliminés. Cela facilite la compréhension de l'ellipse et la récupération de la séquence d'origine. Par exemple, l'adjectif masculin *facial* ne s'emploie qu'avec le substantif *nerf*. L'adjectif *carotide* est toujours accompagné du substantif *artère*, etc.

L'ellipse est un phénomène linguistique très répandu dans le vocabulaire médical. Elle est décrite dans les dictionnaires médicaux. Dans le dictionnaire franco-russe de la médecine d'A. N. Roubakine (1974), il y a beaucoup de mots composés tronqués. Leur description renvoie toujours aux substantifs supprimés. Par exemple, nous trouvons devant le mot *carotide* la remarque : « à propos de l'artère » ; devant le mot *assourdissant*, on trouve : « à propos du bruit », etc.

L'ellipse est considérée comme une sorte de phraséologisme qui porte un seul sens et dont les unités constitutives perdent leur sens initial. Dans ce genre de construction, le substantif porte le sens générique et l'adjectif porte le sens discriminant de la séquence, de telle sorte que l'adjectif devient l'élément essentiel de la construction. Ainsi, l'élimination du substantif devient « évidente » et même recommandée puisque l'utilisation des deux lexèmes (substantif et adjectif) dans la même séquence peut être considérée comme un pléonasme. Après l'ellipse, le terme est un adjectif qui subit une transformation. Ainsi, le *nerf facial* (facial = adjectif) devient *le facial* (facial = substantif).

1.2. Autres mots construits

Dans ce chapitre, nous rapportons les travaux effectués sur la problématique de la siglaison et de l'acronymie dans le vocabulaire médical (Bulletin terminologique et linguistique numéro 62, 1991). Après avoir défini ces deux notions, nous faisons un rappel historique pour montrer que la siglaison touche à toute la sphère scientifique, puis nous décrivons les principes qui régissent l'usage de la siglaison et de l'acronymie et nous précisons quels sont leurs procédés de formation. Nous discutons également de la dérivation dans le domaine de la médecine.

1.2.1. Siglaison

Le vocabulaire médical s'est enrichi au long des siècles de beaucoup de mots composés, de locutions, et d'affixes, si bien qu'il y a des mots trop longs et trop compliqués à lire, à écrire et à prononcer. Par exemple, *Syndrome dysgénésique et respiratoire porcin*.

Pour contourner ce problème, les praticiens de la médecine ont eu recours à la siglaison car c'est un procédé qui est plus économique sur le plan communicatif (J-C. Sournia 1997).

Il y a deux cas de figure selon que la siglaison porte sur une unité monolexicale ou une unité polylexicale (S. Mejri 2008).

La siglaison est une forme de troncation d'une seule forme graphique du mot. Elle est dénommée, au sens large, une abréviation (J-C. Boulanger 1978). Elle peut se faire de plusieurs façons différentes, ce qui entraîne plusieurs appellations. Lorsque l'unité monolexicale est représentée par l'unique lettre initiale, on la dénomme sigle simple (Le Larousse 1976). Par exemple (OMS) *organisation mondiale de la santé*. Lorsque elle est représentée par la lettre initiale suivie d'une ou de plusieurs lettres de ce même mot, on la dénomme un sigle composé (Le Petit Robert 1990). Si elle est représentée par la lettre initiale puis des lettres choisies arbitrairement dans le mot, on la dénomme sigle acronymique (D. Gehenot 1976 : 127). Ces procédés de troncation englobent également l'acronymie.

La réduction d'un syntagme à la lettre initiale de chacun de ses composants est dénommée *sigle* (Le Petit Robert 1990).

La siglaison est apparue à l'époque des Grecs. Elle favorise généralement la majuscule au détriment de la minuscule. L'apparition des sigles sans les occurrences des unités monolexicales ou polylexicales donne lieu à des différences d'interprétation lorsqu'on n'est pas spécialiste du domaine qui justifie la création du sigle. Par exemple APPF signifie (Artérite Pulmonaire Plexi Forme). Certains sigles du domaine médical sont très connus du public, comme SIDA (syndrome d'immunodéficience acquise), alors que d'autres ne le sont pas, comme IVG qui signifie *insuffisance ventriculaire gauche*.

La sélection des lettres se fait parfois d'une façon arbitraire. Par exemple, en gynécologie et en obstétrique, la locution nominale *test de tolérance à la contraction utérine*

est tantôt tronqué sous forme TUC, (G. Muller, 1982 : 77), tantôt tronqué sous forme TTUC (J-P. Verreault 1990 : 133).

La siglaison est considérée comme un phénomène propre aux langues de spécialité. Les savants dénomment de nouvelles notions et parallèlement ils utilisent la siglaison pour abréger les séquences trop longues. *Sclérose en plaque* donne lieu à SEP. Les sigles sont utilisés plus souvent que leurs séquences d'origine par les spécialistes.

On rencontre parfois quelques difficultés dans l'interprétation de certains sigles. L'IVG signifie en cardiologie *insuffisance ventriculaire gauche*. Mais, après l'adoption en France de la loi de (1975) qui légalisait l'avortement provoqué, le même sigle correspond à *interruption volontaire de grossesse*. Cela provoque des confusions au niveau de l'interprétation de ce genre de sigle, notamment en dehors du contexte. Le contexte dans lequel on utilise ce genre de sigle élimine les zones d'ombre parce qu'il aide les interlocuteurs à se situer. Par exemple, L'IVG fait référence à *l'insuffisance ventriculaire gauche* dans les services de cardiologie et à *l'interruption volontaire de grossesse* dans ceux de la maternité. Dans ce sens, J-C. Sournia (1997) ajoute que « le sigle appartient à une langue et n'est compréhensible que dans son cadre précis ».

Il est impossible d'établir une liste exhaustive des sigles de la langue de la médecine même en se limitant à un sous-domaine tel que L'ORL ou l'ophtalmologie parce qu'ils varient d'une région à une autre, et certains connaissent des périodes d'utilisation éphémères.

On considère que ce genre d'abréviation est à éviter, et quelque soit la longueur du terme, le lecteur ou l'interlocuteur attend qu'il soit entièrement écrit ou prononcé parce que le sigle n'est pas toujours facile à comprendre, surtout dans le cas où c'est une convention. Il appartient alors à un groupe de travail ou à un service hospitalier particulier.

La lexicalisation de la siglaison s'appelle l'acronymie. Autrement dit, de la siglaison naît l'acronymie.

1.2.2. Acronymie

L'acronyme est la réduction d'un groupe de mots par l'abréviation, par la ou les premières lettres de ces mots, à l'instar du phénomène de la siglaison montré ci-dessus, mais le résultat est une suite de syllabes qui se prononce comme un mot simple.

L'acronyme n'est pas synonyme de sigle : premièrement, l'acronyme n'est pas un simple sigle mais un sigle lexicalisé. Il devient comme un mot simple du point de vue de la prononciation ; deuxièmement tous les sigles ne produisent pas des acronymes.

L'acronyme se lexicalise lorsque son usage se généralise dans la langue, il porte par exemple la marque du pluriel, et lui-même devient une source de dérivation comme *SIDA* qui est à l'origine de *sidatique*, *sidéen* et *sidéenne*. L'acronyme s'intègre donc dans la langue au point de ne plus être distingué des mots, et ne peut plus être confondu avec les sigles, à condition que ses lettres ne soient pas épelées à la prononciation comme DDS (dent de sagesse).

Dans la langue de la médecine, l'acronyme est souvent employé sans le syntagme qui l'explique. Pour autant, il est généralement bien interprété par l'interlocuteur.

L'acronyme se présente souvent graphiquement sans point d'abréviation. Cela implique qu'il est perçu comme un mot simple par les spécialistes.

Contrairement aux sigles qui interdisent la formation des dérivés, les acronymes « une fois transformés en substantifs, fournissent des bases nouvelles et donnent parfois naissance à des dérivés » (J-C. Boulanger 1978 : 72-73). Ce phénomène est accru lorsque l'acronyme est un terme spécialisé. Notons que les acronymes lexicalisés n'ont pas toujours la possibilité de donner lieu à des formes dérivées. S'il y a peu d'acronymes dérivés dans les langues spécialisées, ce serait à cause de l'évolution rapide des notions. Les nouveaux acronymes entraînent la disparition des anciens de l'usage (George 1977 : 40).

1.2.3. Dérivation

Les termes médicaux sont généralement composés de lettres et de syllabes. Le principe de la composition de ces dernières n'est fixé par aucune règle liée à une discipline précise.

Quelques informations générales peuvent nous rappeler les significations des suffixes et des préfixes rajoutés aux radicaux de la médecine. Les mots simples sont souvent d'origine celtique ou germanique, alors que les mots dérivés et ou composés proviennent du grec et du latin. Seuls les mots dérivés seront traités dans cette partie (J-C. Sournia1997).

Les racines constituent le noyau central des mots. Certaines d'entre elles sont spécifiques et précises. Par exemple, *gastr-* s'utilise en français pour désigner l'estomac. *Enter* s'applique à l'intestin grêle. Le radical *col* convient uniquement au gros intestin. *Hydr-* désigne l'eau et *hidr-* désigne la sueur, etc.

Par contre, d'autres racines du vocabulaire médical prêtent à confusion. L'utilisation de la racine *cardi-* sert à la fois pour désigner tout ce qui s'applique au cœur (cardiopathie), et à la contraction involontaire du *cardia* gênant le passage des aliments de l'œsophage vers l'estomac et provoquant une inflammation (cardiospasmе). J. C. Sournia (1997) souhaite qu'on réserve *cardio-* au cœur et *cardia* à la jonction entre l'œsophage et l'estomac pour éviter les confusions.

De nombreux préfixes ne sont pas propres à la médecine. Autrement dit, ils ne sont pas spécialisés et portent un sens très général. *A, an et al* signifient par exemple une absence. *Ab* signifie une séparation et un éloignement. *All* est utilisé pour désigner tout ce qui est différent. *Di* implique une dualité. *Dis* évoque une anomalie de fonctionnement ou une difficulté. *Ec, ecto, ex* et *exo* expriment une idée d'extérieur, contrairement à *en, endo* et *in* qui évoquent l'intérieur.

Les suffixes semblent être moins importants que les préfixes de point de vue de leurs valeurs sémantiques. *Ase* désigne enzyme. Avant, les parasitologues ont annexé déjà ce suffixe pour de nombreuses parasitoses comme *helminthiase, ascaridiase*. Mais beaucoup d'infections parasitaires sont désignées par *ose* comme *hydatidose* et *filariose*.

Le suffixe *Ose* est aussi utilisé pour désigner certains glucides, dans la mesure où il est utilisé pour dénommer de nombreux produits chimiques. Il s'oppose au suffixe *ite* comme dans le cas de *arthrose* (sans inflammation) et *arthrite* (avec inflammation).

Le suffixe *ite* désigne les maladies inflammatoires. J. Hamburger (1982) mentionne que le suffixe *ite* est souvent rajouté à une base savante pour signifier une inflammation ;

otite, gingivite, dermatite, conjonctivite, etc. mais on utilise parfois *ite* pour des maladies non inflammatoires, par exemple *ostéite*.

1.3. Productivité lexicale

La langue de la médecine évolue constamment du fait de l'apparition de nouvelles maladies, de nouveaux symptômes, de nouveaux examens médicaux ainsi que de nouveaux instruments. Cela nécessite de dénommer ces nouvelles entités afin d'assurer la communication en médecine.

1.3.1. Néologie

Les progrès de la médecine impliquent l'introduction de nouveaux termes dans le vocabulaire médical.

Leur intégration dans le vocabulaire de la médecine est souvent difficile du fait qu'ils se prêtent à de contresens en n'étant pas employés à de bon escient. Leur rejet catégorique freine l'évolution et l'enrichissement du vocabulaire médical.

Toutefois, l'évolution constante du savoir médical et la nécessité de communiquer sur les acquis récents impliquent le recours à des néologismes.

Il s'avère finalement que les termes nouveaux facilitent la pratique de la médecine, bien que les professionnels de la santé et le public ne parviennent pas toujours à les utiliser en contexte (J-P. Boissel & F. Dürr 2004).

Les néologismes sont fondamentaux dans la communication en médecine, de telle sorte que les linguistes se doivent de les faire connaître à la fois aux praticiens et aux patients afin d'améliorer et de faciliter l'intercompréhension dans les dialogues médecin-médecin et médecin-patient.

La langue de la médecine a beaucoup recours à la composition de termes pour former de nouvelles unités significatives. L'un des meilleurs exemples pour illustrer cela est la création d'un nouveau concept, « *bébé médicament* », qui s'est faite après la découverte des

effets méticuleux du sang du cordon ombilical qu'on récupère des nouveaux nés pour soigner des bébés malades.

1.3.2. Emprunt

Si l'on parle de langue spécialisée, plutôt que de vocabulaire spécialisé, c'est parce que chaque profession s'exprime dans une langue donnée, dans laquelle les spécialistes expriment les savoirs et les compétences qui caractérisent leur secteur d'activité.

La langue de la médecine est une langue spécialisée qui permet aux praticiens et aux chercheurs d'échanger des informations d'une façon efficace. Elle est souvent difficilement accessible aux non spécialistes.⁵

Depuis quelques années, les langues médicales se sont beaucoup transformées à cause de l'influence de la langue anglaise, en raison du développement lié à la recherche médicale dans le continent nord-américain.

Selon P. Faure (2010), l'influence de l'anglais sur la langue médicale est courante. Cela n'a pas été empêché par l'action du ministère de la santé publique et de la sécurité sociale Français, dans les années soixante dix, qui a créé dans son département une commission du langage médical. Cette commission a comme tâche d'établir des listes d'anglicismes couramment employés dans la documentation médicale, et de leur trouver des équivalents en français afin que la terminologie française s'impose dans le vocabulaire médical.

Le vocabulaire médical du français est en train de changer en raison de la mise en place de nomenclatures internationales rédigées en anglais ; c'est le cas notamment du vocabulaire de l'anatomie qui utilise de plus en plus de termes des nomenclatures internationales, au détriment de la terminologie habituelle. En français, le terme *atrium* est préféré au terme *oreillette* qui tend à disparaître. De même, le *nerf vague* est remplacé par un autre terme, *nerf pneumogastrique*. Ces changements sont le résultat de l'influence anglo-saxonne.

⁵ En anglais, on parle même de *medspeak* (langue médicale) ou encore *medicalese*.

Les transformations du vocabulaire médical sont soutenues par l'organisation mondiale de la santé (OMS). Cette organisation impose aux langues indo-européennes les dénominations communes internationales (DCI). Cette pratique consiste à nommer les médicaments par leur molécule, ce qui revient à favoriser l'anglais, car les pays anglo-saxons, notamment les USA sont très innovants et font de la recherche de grande qualité.

L'observation du vocabulaire médical français permet de constater l'influence de l'anglais sur le lexique français à beaucoup de niveaux. Les langues européennes comme le français, ont recours aux emprunts, souvent par manque d'équivalents dans la langue cible.

L'influence anglaise va loin au-delà du phénomène de l'emprunt. On crée de nouveaux verbes sous l'influence de l'anglais. Par exemple, on dit *la patiente a été clampée et césarisée* par emprunt de *clamped* et *sectioned*.

Certes, quand une langue ne dispose pas d'un mot, elle l'emprunte à une autre langue. En ce qui concerne le vocabulaire médical français, cette solution est due, selon J-C. Sournia (1997), à la « paresse ou au snobisme des importateurs ». Les emprunts posent des problèmes en français au niveau de l'écrit et de l'oral. Souvent, les gens ne savent pas comment les lire et les accentuer, ni comment les écrire et les prononcer.

Pour l'auteur, la francisation des termes anglais serait une solution pour en finir avec toutes les difficultés liées aux emprunts. *Clearance* peut se prononcer et s'écrire *clairance*, comme le cas de *banding* qui est devenu *bandage* après son introduction il ya une vingtaine d'années à la langue médicale française. Il mentionne aussi les sens attribués aux emprunts dans la langue cible, sens qu'ils n'avaient pas dans la langue source. Par exemple, *sludge* signifie en anglais une boue ordinaire ; or, en français, il désigne une boue particulière qui se dépose dans l'intérieur des artères. Néanmoins, Laborit (1974) avait défendu une thèse linguistique qui aurait favorisé, voire rendu obligatoire les emprunts entre les langues spécialisées. Pour lui, « le langage scientifique appartient à celui qui trouve. Il est non seulement juste mais indispensable que la paternité de la découverte... ». C'est-à-dire, les termes scientifiques appartiennent à la science et non pas à telle ou telle langue.

1.3.3. Éponymie

Selon le Grand Larousse Universel, *éponyme* vient du grec *epônumos*, (*epi* équivaut à *sur* et *onoma* à *nom*), à propos du fondateur légendaire d'une cité qui aurait donné son nom à cette cité. Par exemple, *Athena* est éponyme d'*Athènes*.

En médecine, on nomme des signes, des syndromes, des maladies, des interventions et des instruments par le nom du médecin qui les décrit (J. Hamburger 1982).

Dans la terminologie médicale, le concept *éponyme* est rendu par l'imprécision terminologique des noms *syndrome*, *maladie*, *symptôme* et *signe*. Par exemple, les spécialistes francophones de la santé utilisent indifféremment les termes *symptôme*, qui est de nature subjective, et *signe*, qui est de nature objective, pour désigner les manifestations d'un état pathologique. Or les anglophones différencient symptôme et signe. Le premier est une manifestation subjective comme le mal de tête, et le second est une constatation objective comme les battements irréguliers du cœur. En France, on distingue les signes physiques que l'on peut observer et les signes fonctionnels, de nature subjective. On parle aussi de signes généraux pour parler par exemple des manifestations qui seraient une conséquence de la maladie sur l'ensemble du corps, comme la fièvre (J. Hamburger 1982).

On commet parfois des confusions entre une appellation et son dérivé. En anglais, par exemple, on confond *Addison's disease* (destruction des glandes surrénales) par rapport au médecin Addison, et « *addisonism* » qui signifie tout simplement la présence d'une pigmentation cutanée anormale avec une grande fatigabilité (S. Monin 1996).

Les éponymes ne sont pas constitués des seuls substantifs mais aussi des verbes et des adjectifs. « *Eponymer* » serait le fait d'emprunter le nom propre d'un individu ou celui d'un savant et de l'associer à un référent. Ce procédé est très courant en langues spécialisées, notamment en médecine, où le besoin est permanent de dénommer de nouvelles techniques, de nouveaux instruments, de nouvelles opérations, des théories, des maladies, et de nouveaux syndromes. Les éponymes ne sont pris en compte que s'ils sont reconnus et attestés par les locuteurs spécialistes du domaine en question.

Selon Van Hoof (2001), les éponymes banalisés sont tous les termes substantivés *parkinsonisme*, verbalisés *pasteuriser*, ou adjectivés *parkinsonien*, par opposition à l'éponyme qui figure en tant que nom propre comme *maladie de parkinson*.

2. Etudes sémantiques

Les études sémantiques impliquent l'analyse lexicale, l'analyse syntactico-sémantique et l'analyse de discours.

2.1. Analyse lexicale

De manière générale, la langue évolue constamment ; cette évolution est plus rapide quand il s'agit des vocabulaires spécialisés puisqu'ils parlent de domaines en perpétuelle évolution.

Nous montrerons dans cette partie comment le vocabulaire médical a d'abord été dominé par le grec, puis a progressé depuis l'antiquité grecque jusqu'à maintenant pour se voir finalement dominé progressivement par l'anglais.

2.1.1. De l'hégémonie gréco-latine à l'émergence des langues médicales nationales

Les premiers documents de la médecine remontent à Hippocrate entre le cinquième et le quatrième siècle avant J.-C. Cette période est connue pour la domination scientifique grecque, notamment en médecine. Au début du premier siècle après J.-C. au moment où le grec est toujours la langue de la médecine, un aristocrate romain appelé Aulus Cornelius, surnommé l'Hippocrate latin, rédige *De Medicina*, une encyclopédie médicale en latin où il a conservé des mots grecs à cause du manque d'équivalents latins.

Au moyen âge, une autre langue avait influencé celle de la médecine, il s'agit de la langue arabe, grâce à des érudits comme Avicenne, dont le Canon est resté pendant des siècles la référence des praticiens en Occident.

C'est à la renaissance que la période médicale latine commence réellement, au moment où le grec est de moins en moins compris. Par la suite, les travaux majeurs en médecine sont systématiquement traduits en latin. La terminologie de la médecine ne change pas mais s'enrichit progressivement. Par la suite, on a vu dans chaque pays la substitution progressive du latin par les langues.

L'évolution de la langue de la médecine a évolué en même temps que l'évolution de la médecine elle-même. L'histoire de la médecine montre que la terminologie médicale utilise de moins en moins le grec et le latin et développe une terminologie propre à chaque langue. Néanmoins, la terminologie gréco-latine reste très utilisée en médecine.

En s'adressant à des patients, les médecins optent pour un discours plus facile et plus compréhensible en essayant à chaque fois d'utiliser une terminologie courante. Les praticiens de la médecine opèrent des alternances entre langue générale et langue spécialisée (mots courants et mots techniques) pour se faire comprendre par des patients qui ne possèdent généralement pas de connaissances dans ce domaine. Par exemple, dans certaines langues germaniques, les termes d'origine gréco-latine restent réservés à l'élite. En allemand, le diabète se dit par les non initiés *Zuckerkrankheit*, littéralement « maladie du sucre ».

En français, un patient ne maîtrise pas forcément les termes techniques, sauf s'il est initié. Il dit plutôt *tour de rein* au lieu de *lumbago* ou *mal de cœur* pour *nausée*. Mais on voit de plus en plus de personnes non initiées utiliser les termes techniques de la médecine grâce à l'accès facile à l'information sur internet et la diffusion de séries et de documentaires médicaux à la télévision.

2.1.2. La scientificité

Pascaline Faure (2010) parle de l'appropriation de la profession médicale par la langue de spécialité.

Le savoir médical progresse et se précise au cours des siècles. On a commencé par les théories philosophiques de l'antiquité, puis on est passé à la description anatomique et enfin à l'étude des organes et de leurs fonctions. La précision progressive du savoir médical influence la langue de la médecine qui se dote d'un vocabulaire propre permettant aux spécialistes de communiquer d'une façon précise et concise.

Le discours de la médecine devient de plus en plus scientifique. Désormais, on précise même le groupe auquel la maladie appartient (*grippe A*) ou le type de souche (*grippe H1N1*).

A l'époque, les praticiens de la médecine utilisaient les abréviations, les sigles et les acronymes dans le but de cacher des réalités inquiétantes pour les patients. Par exemple, on utilise **C** pour *cancer* et **néo** pour *néoplasme*.

Aujourd'hui, la langue de la médecine connaît un recours sans précédent aux abréviations, aux sigles et aux acronymes qui ont intégré le jargon médical, non pas pour dissimuler aux patients leurs maladies graves, mais ils font partie des mots de la communication quotidienne dans les milieux hospitaliers.

Les acronymes sont de plus en plus nombreux, et leur large utilisation fait que les non initiés ignorent même les mots qui sont à la base de leur formation. Par exemple, *SIDA* n'est pas toujours reconnu comme l'acronyme de *Syndrome d'Immunodéficience Acquise*. En anglais, on peut même trouver un acronyme au début ou au milieu d'un mot : *picornavirus* (un virus à RAN), *vipoma* (une tumeur qui sécrète une vasoactive intestinale peptide, VIP).

2.1.3. L'universalisation du vocabulaire médical

Depuis les années 1950, il y a un changement au niveau du vocabulaire de la médecine, notamment dans le vocabulaire de l'anatomie pour laquelle la terminologie habituelle est remplacée par des nomenclatures internationales. En français, *oreillette* tend à laisser la place à *atrium*. Le *nerf vague* est remplacé par *nerf pneumogastrique*. Ces changements sont dus à l'influence anglo-saxonne.

Ces nomenclatures sont souvent formées à partir de racines grecques choisies en fonction du rôle ou de la place précise d'un organe ou bien de la caractéristique principale d'une maladie. Par exemple, en français, *maladie d'Addison* devient *insuffisance surrénalienne lente*. Ce processus est encouragé par l'organisation mondiale de la santé (OMS), qui a imposé par exemple, les dénominations communes internationales, dans le domaine pharmaceutique ; il s'agit de nommer les médicaments par leurs molécules et non pas par leurs noms commerciaux.

Pour Pascaline Faure (2010), l'arrivée de moyens de communication et d'information de plus en plus rapide sont conduits les scientifiques à travailler avec une langue commune. Pour ce faire, ils n'ont que le choix d'utiliser une langue médicale fortement influencée par

l'anglais. Cette langue connaît un nombre importants de chercheurs et de publications, notamment, de revues nord-américaines dans le domaine de la médecine.

L'influence de l'anglais sur les langues médicales indo-européennes est remarquable, surtout au niveau du lexique. Les emprunts se font parfois obligatoires à cause du manque d'équivalents dans la langue cible. Il arrive également les langues empruntent un autre terme à l'anglais, ce qui donne des doublons qui sont de plus en plus employés.

Cette influence est due à la domination scientifique des Etats-Unis. Elle peut s'expliquer également par la flexibilité de l'anglais car cette langue facilite la création de nouveaux termes et sa concision favorise la communication rapide et l'économie de la langue.

2.2. Analyse syntactico-sémantique

L'analyse syntactico-sémantique présuppose l'analyse phrastique. Autrement dit, il ne s'agit pas d'étudier les mots de façon isolée, mais dans des phrases où leur sens est déterminé par leur syntaxe. De ce point de vue, la phrase est considérée comme l'unité minimale d'analyse (G. Gross 1996).

2.2.1. Polysémie du vocabulaire médical

Les mots de la langue ne sont pas tous homogènes de point de vue des trois fonctions primaires. Certains peuvent fonctionner comme des prédicats, d'autres comme des arguments.

Les noms dit élémentaires sont toujours des arguments et ne fonctionnent jamais comme des prédicats dans des phrases, par exemple, *table* <meuble> et *main* <partie du corps>. Par contre, les noms comme *fièvre* <*maladie*>, *colère* <*affect*>, peuvent fonctionner comme des prédicats ou comme des arguments. Il s'agit de noms prédictifs (P. A. Buvet, C. Girardin, G. Goss et C. Groud 2005).

<i>Prédicat</i>	<i>N0</i>	<i>NI</i>	<i>V. support</i>	<i>Exemple</i>
<i>fièvre</i>	<i>Pierre</i>		<i>avoir</i>	<i>Pierre a la fièvre</i>
<i>colère</i>	<i>Pierre</i>		<i>être</i>	<i>Pierre est en colère</i>
<i>souffre de</i>	<i>Pierre</i>	<i>fièvre</i>		<i>Pierre souffre de la fièvre</i>
<i>évoque</i>	<i>Pierre</i>	<i>la colère de Marie</i>		<i>Pierre évoque la colère de Marie</i>

Une description du vocabulaire médical en termes de classes d'objets est proposée dans « description de la langue de la médecine » G. Gross & Mathieu-Colas (2001) ; ces deux linguistes démontrent que les vocabulaires spécialisés ne sont pas formés de nomenclatures ou de listes de noms isolés, correspondant aux objets du monde qui résultent du besoin de les dénommer. Les langues spécialisées fonctionnent exactement de la même manière que la langue générale. Les textes, même les plus spécialisés, sont constitués en grande partie de mots de la langue générale. C'est pourquoi les outils d'analyse de la langue générale sont exploitables pour les langues spécialisées

Les arguments sont catégorisés en traits et en classes sémantiques en fonction des prédicats qui les mettent en relation. Nous avons adopté le codage suivant :

G : catégorie grammaticale (*nm, nf, nmf*)

T : trait syntaxique

C : classe d'objets

L'analyse des arguments regroupe tous les substantifs de la médecine qui sont toujours arguments comme les humains non prédictifs : (*Pasteur, Ambroise Paré, Avicenne*), les animaux d'expérimentation : (*souris, lapin*), les plantes médicinales (*goyave*), les micro-organismes : (*virus, bactérie, champignon*), les parties du corps : (*main, tête, pied, etc.*)

Les prédicats sont classés selon qu'ils relèvent des traits *état* (états pathologiques *anémie*, adjectifs d'états *anémique*), *action* (consultations, examens, soins, prescriptions,

opérations, ablations, anesthésies), ou *événement* (maladies infectieuses, inflammations, attaques, blessures et adjectifs causatifs comme analgésique). Ainsi, ils peuvent être des verbes, des noms et des adjectifs.

2.2.2. Actualisation et noms médicaux

Considérer la phrase comme une seule unité d'analyse implique qu'elle est constituée d'un prédicat, le noyau de la phrase, et d'arguments qui lui sont subordonnés. Il s'ensuit que les arguments sont définis par les prédicats dont ils dépendent.

Le prédicat et ses arguments seuls ne constituent pas un énoncé complet et compréhensible. Les actualisateurs sont aussi des éléments constituants de la phrase. Leur adjonction aux prédicats et aux arguments fait appel à des procédés morphologiques et des procédés syntaxiques. La détermination est le procédé syntaxique propre aux substantifs qu'ils soient des prédicats (*Paul a le diabète*) ou bien des arguments (*Paul souffre du diabète*) (P. A. Buvet 2009).

Les prédicats verbaux sont actualisés par la conjugaison tandis que les prédicats nominaux sont actualisés par les verbes supports et les déterminants.

(1) *Paul a le diabète* (*avoir* est verbe support, *le* est déterminant)

(2) *Paul a subi deux opérations chirurgicales* (*subir* est verbe support, *deux* est déterminant)

La nature sémantique des prédicats explique la sélection des déterminants. Les noms de maladies sélectionnent différemment leurs déterminants, c'est-à-dire, qu'ils se combinent avec tel ou tel déterminant selon leur nature sémantique.

(3) *Paul a (le + un + *deux + *quelques + *plusieurs) SIDA*

(4) *Paul a (*l'+un+deux+quelques+plusieurs) herpès*

La différence entre (3) et (4) porte sur les prédicats. On constate que les noms de maladies comme SIDA, qui fonctionne ici comme un prédicat, ne permettent pas la quantification (on ne peut pas avoir deux, quelques ou plusieurs SIDA). Par contre les noms de maladie comme *herpès* permettent la quantification. Les noms de maladies comme *grippe*

et *fièvre* acceptent, non pas le quantitatif mais le fréquentatif puisqu'elles sont des maladies passagères et que l'on peut avoir plusieurs fois dans une période déterminée.

(5) *Paul a eu plusieurs fois la grippe/ la fièvre.*

Le fréquentatif est exprimé ici par l'actualisateur *plusieurs fois*

(6) *Paul diagnostique un début de cancer*

(7) *Paul a un début de cancer*

Dans l'exemple (6), le substantif *cancer* est un argument du prédicat *diagnostiquer*. A ce niveau, il est compatible avec le déterminant nominal à valeur aspectuelle inchoative *un début de*.

On constate que le déterminant *un début de* s'applique aussi au substantif *cancer* en position de prédicat. Pour les exemples (6) et (7), on peut conclure que c'est l'aspect intrinsèque du substantif qui sélectionne le déterminant et non sa position en tant qu'argument ou prédicat.

P. A. Buvet (2008) montre les limites d'utilisation des verbes supports à aspect inchoatif *attraper* et *contracter*. La substitution de ces verbes supports au verbe *avoir* pose quelques conditions. L'utilisation du verbe *attraper* est plutôt réservée à l'usage familier de la langue : *Il a attrapé la crève*. Par contre, *contracter* est plutôt réservé au niveau soutenu : *Il a contracté l'influenza*.

2.2.3. Appropriation et noms médicaux

La notion de prédicat ne concerne pas uniquement la classe des verbes, elle concerne aussi d'autres catégories de discours comme les noms et les adjectifs. Une même racine prédicative peut donner jusqu'à trois prédicats : *anesthésier*, *anesthésie* et *anesthésiant* sont formés à partir de la racine prédicative *anesthési*. (G. Gross 2001).

Selon P. A. Buvet (1998), les noms de maladies fonctionnent comme des prédicats dans le vocabulaire médical dans les deux constructions phrastiques en *avoir* suivantes :

Paul a une grippe

Paul a une acné sur la joue

Il existe des verbes spécifiques (leur nombre est limité) à la langue de la médecine. Ils renvoient à ce domaine par leur contenu sémantique (X. Blanco 2000).

Anesthésier (implique *anesthésie* et *anesthésiste*)

Hospitaliser (implique *hôpital* et *hospitalisation*)

Soigner (implique *médecin* et *patient*)

Suturer (implique *chirurgien* et *patient*)

Souffrir (implique *patient* et *blessure* ou *maladie*)

L'introduction des verbes spécifiques à la médecine dans des phrases modifie leurs schémas d'argument. Par exemple, lorsqu'on introduit le verbe *diagnostiquer* et *guérir de* dans la phrase : *Paul a la grippe*, le nom de maladie grippe devient argument dans :

Le médecin diagnostique la grippe chez Paul

Paul est guéri de la grippe

Les verbes *diagnostiquer* et *guérir* prennent la position de prédicat, et *grippe* devient un argument.

2.3. Analyse de discours

En linguistique, on distingue souvent la langue générale et les langues spécialisées ou langues de spécialités. Or, l'analyse syntactico-sémantique des langues dites spécialisées montre que ces dernières partagent exactement les mêmes propriétés du point de vue de leurs caractéristiques syntaxiques et terminologiques avec la langue générale. Les langues spécialisées sont avant tout des langues naturelles ayant comme fonction la communication des savoirs spécialisés (P. Lerat1995).

2.3.1. Problèmes stylistiques

La notion de style est souvent associée à la littérature et à la poésie où le sens passe par « représentation imagée », en utilisant des procédés comme la métaphore. Si le discours littéraire est fait ainsi, le discours spécialisé, dont la priorité est de communiquer des réalités

de façon claire et objective, doit viser la neutralité, l'objectivité et l'absence de connotations. En d'autres termes, le vocabulaire spécialisé est censé écarter tout emploi imagé ou figuré. Dorothy Nakos (1994) démontre le contraire. Il y a foisonnement d'images dans les documents médicaux.

Les problèmes d'ambiguïté et de manque de précision sont cruciaux dans les discours médicaux. Selon Dorothy Nakos (1994), les termes connotés ne peuvent pas rapporter la réalité d'une façon adéquate. A l'inverse, les termes dénotés n'arrivent jamais à décrire la réalité d'une façon absolue. Autrement dit, le principe de la monosémie n'existe pas, même en langue de spécialité. La polysémie renforce la richesse d'une langue.

Il y a des métaphores voilées et des métaphores apparentes en médecine. Concernant les métaphores voilées, elles sont qualifiées de « métaphores mortes » à cause de la difficulté dans leur compréhension, par des raisons étymologiques. C. Murray Turbayne (1972) dit que la plus grande partie du langage est composée de métaphores « mortes » ou « dormantes ».

En pathologie, le mot *cirrhose* signifie la coloration jaunâtre du foie. La métaphore paraît aussi dans le mot *arachnoïde* ou *arachnodactylie*, dont la racine *arach-* signifie en grec araignée : le premier terme désigne la forme d'une membrane de méninge et le second terme une malformation pathologique des doigts.

Pour certaines métaphores voilées, les connaissances linguistiques suffisent pour les déchiffrer, tandis que d'autres nécessitent des connaissances linguistiques, médicales et historiques.

Si les métaphores voilées de la médecine comme *cirrhose* mobilisent des connaissances pour les déchiffrer, notamment le grec et le latin, les métaphores non voilées, elles, sont faites en français. Elles prennent un sens figuré par les procédés de rapprochement d' « une chaîne référentielle à une autre pour faire naître des images franches et parlantes ».

S. Ullmann (1957) parle de la transposition des noms d'organes humains, considérée comme l'une des formes universelles de la métaphore. Il l'illustre avec des exemples comme *tête* appliquée à *un arbre*, *une épingle*, *un pont*, etc. De même, beaucoup de noms d'organes humains dérivent des noms appartenant à d'autres domaines comme : *animal*, *végétal*, *minéral*, *architecture*, *géographie*. Cela concerne l'anatomie, mais aussi la pathologie.

Quelques termes de ces domaines servent à construire des syntagmes comme : *nerf vestibulaire, colonne vertébrale, écorce cérébrale*, etc.

2.3.2. Langue générale et langue spécialisée

Après avoir vu quelques similitudes entre la langue générale et les langues spécialisées du point de vue syntactico-sémantique, nous essayons d'appréhender le vocabulaire de la médecine du point de vue de l'articulation entre langue générale et langue spécialisée.

Pour Christian Balliu (2001), « un texte médical » n'équivaut pas à un « texte écrit par un médecin », même spécialisé. L'auteur donne un exemple d'un texte écrit par un médecin, qu'on considère comme un texte médical par sa terminologie. Pourtant, le texte lui-même fait penser à un roman policier. C'est l'histoire d'un chercheur à la quête d'un virus, traqué comme un malfaiteur.

Wimmer (1982), cité par Bernd Spillner (1992), distingue la langue spécialisée et la langue générale à partir des critères suivants :

- La langue de spécialité est, selon lui, connue par sa précision, son univocité déterminative, son économie, son invariance situationnelle, son rapport à la matière et son niveau théorique. Par contre, la langue générale est connue par son indétermination, son ambiguïté, ses redondances, sa multiplicité situationnelle et sa multiplicité thématique.
- La langue de spécialité vise la précision sémantique, la systématisation conceptuelle et la neutralité émotive. Les scientifiques mettent en avant l'objectivité qui garantit le caractère scientifique de leur discours. Or, on a remarqué que la terminologie médicale ne garantit en aucune manière la vérité ni la précision du discours. En effet, en médecine, il y a une distinction entre le vocabulaire savant et le vocabulaire populaire ; ainsi *exsangue* équivaut à *anémique* et *tumeur cervicale* équivaut à *cancer du cou*. Cette différence se produit dans d'autres langues indo-européennes comme dans l'allemand médical où il y a une façon de dire *diabète* dans la langue générale : *Zuckerkrankheit*, littéralement, « maladie du sucre ».

Dans un autre article, Christian Balliu (2005), préfère parler volontairement de vocabulaire de spécialité et non pas de langue de spécialité, disant que le français ne se subdivise pas en « sous-langue » : langue médicale, langue juridique, langue d'informatique, etc. mais qu'il existe des discours spécialisés pour chaque métier ou chaque domaine.

Concernant le discours médical, il existe une littérature médicale qui est une manière d'écrire ou de parler propre à la médecine et aux médecins, caractérisée par des nomenclatures spécifiques à ce domaine. Dans le cas d'un dialogue entre un médecin et un patient non initié, le médecin opte plutôt pour une terminologie courante et facile.

2.3.3. Dialogues médecin-patient

La médecine est un domaine spécialisé, son vocabulaire progresse, se précise et se spécialise au cours des siècles, en même temps que son évolution scientifique. La précision sans cesse renouvelée du savoir médical influence son vocabulaire. On observe, par exemple, l'évolution dans la langue médicale à travers l'utilisation fréquente des éponymes et des métaphores dans la dénomination des maladies. Désormais, on appelle les maladies par les noms de leurs agents pathogènes (*Herpes Zoster, papillomavirus, Helicobacter*), Et on précise même le groupe auquel la maladie appartient comme la *grippe porcine* : (*grippe A*) dont le type de souche (*grippe H1N1*), grippe aviaire dont le type de souche (*grippe H5N1*).

En plus de ces procédés, la médecine recourt massivement à l'utilisation des abréviations, des sigles et des acronymes (P. Faure 2010).

Les patients qui ne possèdent pas nécessairement des connaissances dans le domaine médical, ne sont pas en mesure de comprendre le discours technique des médecins. Cela est dû à l'évolution de plus en plus rapide du vocabulaire médical qui tend à se préciser par la concision, notamment le recours aux mots techniques et aux abréviations.

Pour qu'il y ait compréhension dans un dialogue médecin/patient, on remarque que les praticiens de la santé alternent entre la langue générale et la langue de spécialité pour comprendre et se faire comprendre par le patient.

Les praticiens de la santé utilisent une terminologie qui devrait satisfaire à la fois la curiosité des patients voulant avoir des renseignements et des précisions sur leurs maladies, et les exigences scientifiques de la langue de la médecine.

En français, le patient ne maîtrise pas forcément les termes techniques, sauf s'il est initié. Il utilise plutôt les mots « tour de rein » au lieu de *lumbago*, « mal de cœur » pour *nausée*. Les mots techniques sont réservés aux spécialistes du domaine.

Le médecin fait des efforts pour simplifier son discours en utilisant des mots faciles, même si Christian Balliu (2010) évoque le caractère scientifique « indissociable de la démarche médicale ». Il soutient l'idée que le vocabulaire médical est par essence monosémique et univoque et non pas ambigu.

Mais, il y a toujours une façon d'exprimer les notions scientifiques, entre autres en médecine, d'une façon simplifiée parce que le médecin reste en contact permanent avec ses patient ; du coup, il doit être en mesure de se faire comprendre. Nous observons ces deux exemples suivants qui illustrent cette notion de variation entre le vocabulaire technique et le vocabulaire courant :

(1) *Le virus du SIDA s'attaque aux lymphocytes, ce qui induit une diminution sensible des cellules T4.*

(2) *Le VIH est lymphotrope, ce qui induit une lymphopénie T4.*

On remarque qu'il n'y a pas de différences sémantiques dans les deux exemples, mais l'exemple (1) est plus ou moins accessible au grand public ; par contre, dans l'exemple (2), il y a eu recours à la concision avec des mots très techniques qui ne sont compris que par les spécialistes de la médecine ou par l'élite.

3. Études applicatives

Dans cette partie, nous nous consacrons à la terminologie médicale et au vocabulaire médical de point de vue de la traduction, de la recherche documentaire et du traitement automatique du vocabulaire de la médecine.

3.1. Traduction

La traduction des textes médicaux était placée au centre des préoccupations des médecins et des chercheurs depuis l'antiquité grecque parce que les souffrances de l'âme et du corps ont toujours été au centre des préoccupations de l'homme.

Aujourd'hui, la traduction des textes médicaux reste une priorité majeure des médecins, des chercheurs et des traducteurs, puisque les motivations sont toujours les mêmes ; comprendre les découvertes, les médicaments et les nouveaux traitements mis en place par les chercheurs pour apporter des solutions et de l'apaisement aux souffrances physiques et morales. L'anglais reste la langue la plus traduite dans le domaine médical, notamment en raison de nombreuses recherches et découvertes du côté anglo-saxon.

3.1.1. Didactique de la traduction médicale

Pour réussir l'enseignement de la traduction médicale, l'enseignant a une tâche très difficile à accomplir. Il s'agit de la motivation des étudiants.

Selon Hannelore Lee-Jahnke (2001), la motivation des étudiants grandit grâce à la connaissance de la matière à enseigner. Cela peut être facilité aussi par la connaissance des étudiants et éventuellement le bon choix des textes à traduire.

L'enseignement interdisciplinaire serait un bon choix, comme ce que l'Université de médecine de Genève a appliqué. Dans cette faculté, les étudiants en traduction travaillent en collaboration avec les étudiants en médecine. Dans ce cas, les étudiants devraient connaître les théories applicables aux textes scientifiques.

Pour réussir l'enseignement de la traduction, notamment en médecine, les enseignants ont toujours recours à la documentation. Elle devrait, selon les pédagogues, faire partie intégrante du parcours des étudiants en traduction, dans leur cursus.

Il semble indispensable d'intégrer l'usage des nouvelles technologies. Il s'agit d'enseigner aux futurs traducteurs comment choisir et classer les documents d'une manière systématique, tout en sachant que chaque mois, environ 4000 revues médicales sont publiées. Afin de maîtriser ces informations, David Jewel (1995) propose une lecture décrite en trois phases. Il s'agit:

- a) de feuilleter rapidement l'ouvrage pour déceler ce qui est intéressant
- b) d'être à la recherche, dans la documentation, de solutions à des problèmes spécifiques.
- c) d'effectuer une lecture ciblée pour acquérir des connaissances approfondies.

Hannelore Lee-Jahnke (2001) ajoute que l'apprentissage de la traduction doit fournir des outils adéquats pour réussir à faire des exercices. Il existe aujourd'hui une banque de données, notamment Data Star qui a beaucoup d'informations dans le domaine biomédical.

3.1.2. Outils pour la traduction médicale

La traduction médicale nécessite l'exploitation des outils pour maîtriser la traduction d'un texte de médecine, notamment la documentation qui semble très importante. Le texte à traduire fournit parfois des explications qui éclairent ses propres zones d'ombre.

La traduction médicale ne se base plus sur une seule source. Il est important de faire des lectures critiques parallèles pour pouvoir accéder à la compréhension parfaite du texte à traduire. Le traducteur doit diversifier ses consultations, en commençant par les documents en langue d'origine, et en restant vigilant par rapport au degré de fiabilité des articles en fonction des points suivants : notoriété de l'auteur, lieu de publication et date de publication.

Concernant la documentation à exploiter en vue de la traduction médicale, nous signalons la base de données Medline qui contient plus de trois millions d'articles. Dans cette base de données, on peut obtenir des informations sous trois formes différentes : sur papier, en ligne sur internet ou sur cédérom. La consultation de la base aide le traducteur à comprendre et à cerner le texte à traduire (Hannelore Lee-Jahnke 2001).

Le dictionnaire reste un outil pédagogique incontournable. Il est nécessaire à chaque fois que l'on n'arrive pas à comprendre des termes médicaux. La répartition de la nomenclature en fonction des catégories grammaticales est tout à fait différente de celle des dictionnaires de langue.

Les dictionnaires médicaux comportent entre 74% et 94% de noms, le nombre d'adjectifs varie entre 4% et 20%, et les autres parties de discours varient entre 1% et 4%. Ces proportions ne sont pas identiques pour tous les dictionnaires médicaux. Par exemple, le

dictionnaire médical **PManu** se distingue des autres dictionnaires médicaux par sa grande proportion d'adjectifs (près de 20%) (M. Rouleau 2001).

3.1.3. Difficultés de la traduction médicale

Les difficultés liées à la traduction des textes médicaux sont à la fois d'ordre lexical, d'ordre terminologique et d'ordre phraséologique, et elles impliquent des contraintes syntaxiques. Les difficultés que rencontrent les traducteurs dans leur travail ont un impact important sur le sens et la compréhension des textes traduits généralement ambigus au départ.

Le professeur Sournia (1874) condamne « l'usage abusif » de l'adjectif car il traduit les ambiguïtés liées à la traduction des locutions d'origine anglo-saxonne telles que *chirurgien infantile*, *blessé grave*. Il signale l'ambiguïté des locutions d'origine anglo-saxonne générées par le calque de la dérivation adjectivale en anglais. Cette pratique est généralisée en français, au point que certaines locutions adjectivales comme *profils comportementaux* et *traitements institutionnels* ne sont plus ambiguës.

Selon F. Maniez (2001), la traduction du nom adjectival est particulièrement difficile pour le traducteur qui se trouve dans la situation de choisir entre deux structures possibles : celle de complément du nom et celle qui utilise l'adjectif formé par dérivation à partir du nom français correspondant. Ce choix est déterminé par l'usage : *cell repository* est traduit par « banque de cellule », *cell wall*, par « paroi cellulaire ». Cet usage est figé dans ces deux exemples de collocations et dépend de facteurs syntaxiques et sémantiques.

En français, la plupart des adjectifs qui se terminent par *-aire*, *-al*, *-eux*, *-ien* et *-ique* n'ont pas d'équivalent en anglais. Ils ne peuvent se traduire en anglais qu'à l'aide de la pré-modification par le nom adjectival. C'est plus compliqué avec les adjectifs composés.

Hannelore Lee-Jahnke (2001) évoque les difficultés de la traduction médicale, liées à la terminologie, aux acronymes et à la phraséologie.

La méconnaissance du grec et du latin rend la pratique de la traduction des textes médicaux difficile. Pour y remédier, Maurice Rouleau (1994) a développé une approche qui consiste à diviser les mots et à analyser leur contenu pour permettre à l'apprenti traducteur de se familiariser avec la langue à traduire.

La langue de la médecine est pleine d'acronymes standardisés et d'acronymes spécifiques aux auteurs. Dans ce cas, le traducteur doit reconnaître à chaque fois les acronymes et les expliquer.

Les difficultés liées à la phraséologie médicale sont dans la plupart des cas inhérentes au genre du texte. Parfois, le traducteur se trouve confronté à une certaine redondance spécifique au domaine médical.

Henri Van Hoof (2001) évoque les difficultés de la traduction des éponymes banalisés (tout terme substantivé : *parkinsonism/ parkinsonisme*, verbalisé : *to pastorize/ pasteuriser*, adjectivé : *parkinsonian/ parkinsonien*) en anglais médical.

En ce qui concerne les substantifs, les difficultés liées à leur traduction peuvent être causées par différents problèmes : soit le français n'a pas d'équivalent éponymique ou d'équivalent banalisé, soit il donne une seule forme alors que l'anglais en a deux : *bartonellosis, bartonellosis/ « bartonellose »*.

Pour la forme verbale en anglais, on forme des verbes en ajoutant le suffixe *ize* à n'importe quel radical ou on verbalise tout simplement les substantifs. Cette pratique est rarement permise sauf dans le cas d'éponymes verbalisés.

3.2. Recherche documentaire

A partir des années 1990, le web est devenu une source majeure d'informations scientifiques. Pour les professionnels de la santé et les chercheurs, l'utilisation du web s'avère difficile à cause de la multiplication des annuaires et des outils de recherche. Les moteurs de recherche généralistes (comme Google) ne permettent jamais d'obtenir des informations scientifiques claires, organisées et surtout exhaustives, notamment dans le domaine de la médecine.

3.2.1. Ressources documentaires

La consultation des documents spécialisés fait partie de la démarche documentaire du chercheur, quelque soit le domaine dans lequel il exerce. Selon J. Delisle (1997), la

compréhension d'un texte peut nécessiter la consultation de diverses sources de documentation.

La compréhension de la langue de la médecine étant difficile pour les non spécialistes, il existe sur le web des documents efficaces qui donnent la possibilité de comprendre un sujet quelconque dans le domaine en question. Pour faciliter l'accès à l'information, par exemple sur une maladie donnée, on propose de déterminer d'abord quel est l'appareil, l'organe, le système ou le tissu que la maladie attaque. Ensuite, on propose la documentation sur l'organe ou l'appareil afin de connaître son anatomie et son fonctionnement, puis les médicaments nécessités par le traitement (A. Francoeur & M. Brisebois 2001).

Sylvie Vandaele (2001) soutient que les bases de données bibliographiques qui sont disponibles depuis quelques années sur le web constituent une ressource privilégiée pour le chercheur. En biomédecine, plusieurs bases de données sont disponibles et chacune a sa spécificité.

Elles donnent les références d'articles et d'ouvrages de différents domaines. Elles sont répertoriées dans des fiches correspondant à des références et comportent différents champs dans lesquels on indique le titre, le nom de l'auteur, etc.

Certaines bases de données contiennent des textes complets, comme dans Medline ; ils sont généralement rédigés en anglais, et certains le sont dans d'autres langues européennes. Chaque base traite un ensemble de sous-domaines biomédicaux particuliers.

Dans le cadre de son travail, le chercheur peut consulter de nombreuses ressources en ligne : dictionnaires bilingues et monolingues, encyclopédies, glossaires. Néanmoins, cette recherche ne pourra pas être complète et exhaustive puisque les ressources du domaine biomédical sont dispersées sur des milliers de sites web, malgré les efforts de certains. Par exemple, l'association d'aide à la recherche sur le cancer a mis sur son site un lexique de cancérologie. D'autres organismes ont fait de même comme l'Office de la langue française qui a diffusé sur son site un inventaire de terminologies. En juillet 2000, l'inventaire regroupait plus de cent liens vers des ouvrages du domaine médical (A. Francoeur & M. Brisebois 2001).

3.2.2. Moteurs de recherche

La recherche de la documentation spécialisée, notamment en médecine, nécessite la consultation de répertoires spécialisés. Ce sont des carnets d'adresse de sites web regroupés à l'intérieur de catégories définies. Les chercheurs les exploitent pour trouver de la documentation de haut niveau dans des domaines spécialisés, notamment en médecine.

Avant de rajouter un site web aux répertoires spécialisés, les documentalistes prennent soin de s'assurer de la validité des informations diffusées sur le site en question. Il s'agit de leur donner de la crédibilité.

Dans les répertoires spécialisés, on ne se contente pas de donner l'adresse des sites, mais on présente aussi une description généralement détaillée de leurs contenus.

En science biomédicale, les sites sont généralement présentés de deux façons. Dans la majorité des cas, on propose la liste des grandes spécialités en médecine, comme la chirurgie et la pédiatrie. La sélection d'une grande spécialité donne une multitude de sous-domaines de la même spécialité. Dans d'autres cas, c'est l'inverse ; on accède d'abord à une liste de catégories de ressources qui donne accès par la suite au répertoire des domaines de spécialisation.

Les portails médicaux proposent à partir de la page d'accueil un nombre de ressources et de services : service du courrier électronique, groupe de discussion, grands titres d'actualité, petites annonces, etc. Les informations diffusées par ce type de portails sont différentes. Cela leur doit un énorme succès.

Les portails constituent un outil important des textes journalistiques dont les thèmes traitent de la médecine comme les brochures destinées aux patients pour s'informer sur leurs maladies.

Philippe Eveillard (2000) mentionne que la qualité de l'information diffusée par les portails médicaux n'est pas meilleure que celle qui est diffusée par les institutions.

Les sites généraux comme *Yahoo*, *La toile du Québec* et *Nomade* proposent eux aussi de la documentation dans le domaine médical. Ils ont une catégorie « médecine » ou « santé » qui permet de consulter des documents médicaux, mais ils sont plutôt destinés au grand public.

Les moteurs de recherches tels que *Google*, *Alta Vista* et *Northern Light* donnent aussi la possibilité d'accéder à des ressources documentaires en médecine. Les sites répertoriés dans ces moteurs de recherches fournissent un contenu qui n'est pas évalué par les spécialistes du domaine ; la qualité de l'information peut donc varier d'un site à un autre.

Les sites web des organisations, associations, institutions et sociétés du secteur médical, tel que le site de *l'Organisation mondiale de la santé*, *l'Institut Pasteur*, *Santé Canada*, etc. est une source d'information d'une grande importance pour le spécialiste et le chercheur dans le domaine de la santé.

3.2.3. Bases de données médicales : l'exemple de MeSH

La recherche d'information dans une collection de documents nécessite leur indexation. Dans le domaine biomédical, le thésaurus MeSH, développé par l'US National Library of Medicine, est l'outil d'indexation de la langue de la médecine. Il se présente sous forme d'arborescences afin de permettre une recherche simple, facile et accessible en matière d'information.

Le MeSH (Medical Subject Heading) est une liste de termes organisée et utilisée pour l'analyse documentaire dans le domaine médical. Il est défini comme le langage MEDLINE utilisé pour indexer de manière standard le vocabulaire médical. L'objectif de MeSH est la classification hiérarchique des termes définis comme concepts pour la réalisation d'un vocabulaire contrôlé. Les termes MeSH d'une référence bibliographique sont définis par des humains. Ils sont organisés de façon hiérarchique à l'intérieur des grandes catégories. Ils fonctionnent de façon hiérarchique dans l'arborescence et ils sont accompagnés de descripteurs qui permettent de préciser le sens du terme.

Ce thésaurus contient :

1. une liste alphabétique comportant une notice pour chaque terme
2. une arborescence des termes classés par ordre hiérarchique
3. un index alphabétique listant les mots qui figurent dans les descripteurs

Les termes MeSH sont classés selon une arborescence qui décrit tout d'abord des notions générales qui englobent plusieurs concepts, pour arriver à des notions plus

spécifiques. Les termes situés en haut de la hiérarchie sont appelés termes « parents » tandis que ceux qui se situent en bas de hiérarchie sont dits termes « enfants ».

Chaque terme MeSH est défini par plusieurs champs lui conférant un statut unique (dont un identifiant MeSH tree number) pour pouvoir naviguer dans l'arborescence.

La base MeSH a une double structuration, alphabétique et hiérarchique. La base contient 22997 descripteurs partagés sur 11 niveaux. Ces niveaux sont des champs d'indexation très élevés (ils sont au sommet de la hiérarchie). Ils contiennent des termes généraux de la médecine comme *anatomie* et *maladies*. Ces descripteurs contiennent des termes spécifiques comme *cœur*, *foie*, *rate*, etc., pour *anatomie*, *maladies métaboliques*, *maladies endocriniennes*, etc., pour *maladies*. L'arborescence contient au total quinze branches hiérarchisées par catégories selon l'ordre alphabétique.

Le mode de recherche des termes fonctionne par classes permettant d'accéder à des sous-classes. Autrement dit, chaque terme « parent » contient des détails qui sont des termes « enfants » dans la hiérarchie, par exemple :

Maladies

Maladies respiratoires

Bronchite

Bronchiolite

Etc.

Chaque terme est caractérisé par des champs, en particulier, par une définition, un numéro NLM (National Library of Medicine) et un synonyme qui permet une recherche bilingue.

Le thésaurus MeSH a été réalisé en anglais et traduit dans plusieurs langues, notamment les langues indo-européennes. Il est en accès libre et gratuit. Il a pour but de faciliter la recherche des mots clés et sujets adéquats lors d'une recherche de ressources documentaires dans le domaine biomédical. Son utilisation est relativement simple. A partir de la page d'accueil, en cliquant sur MeSH, on nous propose trois modes de consultation différents.

1. recherche simple :

Elle est simple et pratique. Pour obtenir des informations sur un terme médical spécifique, il suffit de le saisir dans une case prévue pour cela. Dans le cas de la saisie d'un terme synonyme, MeSH nous renvoie automatiquement à un terme retenu comme un terme standard. Par exemple, la saisie de *jaunisse* aboutit à *ictère* comme réponse. Dans ce mode de recherche, MeSH fournit des réponses en français et propose la traduction en anglais ainsi que des liens vers les références de MEDLINE, CIMEF ou Term Science.

2. recherche guidée par les index

Dans ce cas, on saisit le terme recherché, si l'on n'obtient pas de résultat, on saisit son synonyme dans une autre case. Le terme recherché peut apparaître deux fois ou plus dans l'arborescence, selon sa nature et ses particularités, par exemple, *diabète* peut apparaître à la fois dans la liste des maladies métaboliques et nutritionnelles et dans la liste des maladies endocriniennes.

3. navigation dans l'arborescence

Pour commencer la recherche, il faut ouvrir l'arborescence au niveau d'un terme comme *anatomie* pour accéder aux parties du corps.

MeSH a beaucoup d'avantages pour la recherche documentaire sur le web dans le domaine biomédical, puisqu'il est indexé de manière pertinente et cohérente. La recherche documentaire sur MeSH garantit des résultats de qualité en un temps record.

3.3. TAL

L'avènement de l'informatique a changé les habitudes des chercheurs et des professionnels dans plusieurs domaines. Dans le domaine biomédical, comme dans d'autres domaines scientifiques, on a pris conscience du rôle de l'informatique dans la recherche documentaire, mais également pour avoir des résultats pertinents et concrets, notamment en

traduction automatique et semi-automatique. Il convient d'analyser les descripteurs formalisés de la médecine.

3.3.1. Le vocabulaire de la médecine et le traitement automatique

Le traitement automatique des langues naturelles a connu un essor considérable ces dernières années. La recherche d'information sur le web et la traduction automatique et semi-automatique sont devenues des besoins indispensables des chercheurs, des médecins et des traducteurs dans le domaine médical.

Les linguistes ont démontré que les langues de spécialité telles que la langue de la médecine fonctionnent exactement de la même façon que la langue générale, du point de vue syntaxique et sémantique.

Pour J. Domengeot, P. Le Beux et G. Weil (1994), le traitement automatique et l'intelligence artificielle sont importants dans le domaine biomédical pour des raisons différentes. Il considère qu'ils facilitent premièrement la circulation des documents sur le web afin de favoriser la coopération entre les spécialistes d'autres domaines médicaux. Deuxièmement, ils donnent les outils nécessaires aux praticiens de la santé et aux chercheurs pour la structuration du savoir dans le domaine médical. Troisièmement, ils permettent à la fois une analyse du lexique et du discours médical afin de lever l'ambiguïté.

D'autres applications ont été mises en place grâce au traitement automatique. Il s'agit du résumé automatique des textes et l'indexation automatique de documents. Concernant l'indexation, « *plus que de compréhension, il s'agit de projection des documents traités sur un certain nombre de concepts clefs plus ou moins fixés a priori en vue de retrouver les plus pertinents à partir d'une question posée* » (*ibid*). Grâce au traitement automatique, dans le domaine médical, on génère de l'information à partir de l'extraction des données par l'indexation.

En ce concerne le traitement automatique de la langue générale, les textes et les corpus sont un matériau fondamental. Cyril Grouin, Arnaud Rosier, Olivier Dameron et Pierre Zweigenbaum (2009) considèrent que cela s'applique aussi au traitement automatique des langues spécialisées, notamment à la langue de la médecine, soit pour l'indexation et le

codage automatique, pour la catégorisation des patients, ou pour le repérage des gènes et des protéines.

Les langues spécialisées fonctionnent exactement de la même façon que la langue générale (P. Lerat 1995). Cela sous-entend qu'un corpus spécialisé est suffisamment utile pour étudier les problèmes liés au traitement automatique et mettre au point un système d'analyse (C. Grouin, A. Rosier, O. Dameron & P. Zweigenbaum 2009).

Fiammetta Namer (2005), dans son article « automatiser les définitions des termes médicaux : qu'est ce que le traitement automatique du langage apporte à la théorie morphologique ? » présente un système d'analyse Dérif (dérivation en français) où elle a exploité à la fois les termes du vocabulaire général extraits du TLF (Trésor de la langue française) et les termes spécialisés de la biomédecine, dans lequel elle présente « système d'analyse morphosémantique automatique du français associant aux mots inconnus de la langue médicale construits par dérivation ou composition une définition basée sur le rapport existant entre ces mots et les éléments qui les composent ». Elle montre par ce biais que le traitement automatique de la langue et les théories morphologiques, en particulier celles qui concernent la structure interne du mot, s'enrichissent mutuellement.

3.3.2. Quels outils pour le traitement des textes médicaux ?

La langue de la médecine est une langue de spécialité. Cette dernière est constituée de phrases, constituées de prédicats et d'arguments. Il s'ensuit qu'une langue de spécialité fonctionne comme la langue générale (G. Gross et Mathieu-Colas 2001 & Mejri 2011).

Les trois fonctions primaires (prédicat, argument et actualisateur) et la notion de classes d'objets mise en place par G. Gross, peuvent constituer un descriptif d'analyse des textes dans la perspective du traitement automatique.

Selon G. Gross & M. Mathieu-Colas (2001), il convient de préciser que les mots ne sont pas des unités qui fonctionnent de manière indépendante, mais qu'il faut les étudier dans des phrases car leur sens résulte de leur contexte.

La plupart des mots sont polysémiques, même lorsqu'ils sont dans les domaines spécialisés.

Classer les mots selon qu'ils fonctionnent comme des prédicats, des arguments ou des actualisateurs peut contribuer à réduire la polysémie.

La classification des arguments est une étape importante pour le traitement automatique, puisqu'elle permet de réduire la polysémie et de générer toutes les phrases possibles pour le même emploi prédicatif.

La notion de prédicat ne concerne pas seulement les verbes ; elle concerne d'autres parties du discours. Une même racine prédicative peut avoir trois réalisations morphologiques différentes (*il désire s'en aller, il a le désir de s'en aller, il est désireux de s'en aller*). La description des unités lexicales consiste à élaborer des dictionnaires de prédicats et d'arguments.

Pour l'apport informatique, la langue de la médecine a connu de nombreux travaux notamment ceux de Pierre Zweigenbaum et Asma Ben Abacha. Leurs recherches sont motivées par l'évolution et l'enrichissement de plus en plus rapide de la langue de la médecine. Ils considèrent que le vocabulaire médical double en volume tous les cinq ans, voire tous les deux ans. Avec l'augmentation des termes médicaux et d'information en médecine, les chercheurs et les médecins ont besoin d'outils leur permettant la recherche efficace, rapide et précise tel que les systèmes de question-réponse.

Pour ce faire, Pierre Zweigenbaum et Asma Ben Abacha (2011), ont proposé une approche hybride afin de détecter les relations sémantiques dans les articles et les résumés indexés par Medline. Cette méthode s'intéresse en particulier aux relations sémantiques qui relient deux des principaux termes de la médecine, il s'agit de *maladie* et de *traitement*. Ce choix s'explique par la fréquence qui est relativement élevée de ces deux types d'entités dans les corpus et les textes médicaux. Leur méthodologie consiste à décrire le corpus afin de faciliter la reconnaissance des relations qui lient les deux entités de mots *maladie* et *traitement* puis ils combinent une méthode à base de patrons et une méthode statistique.

Pour la description du corpus, Pierre Zweigenbaum et Asma Ben Abacha ont exploité le corpus extrait de Medline (2001). Ce corpus va être annoté avec plusieurs types de relations sémantiques entre *maladie* et *traitement*. Le processus d'annotation comporte trois étapes successives :

- a) la reconnaissance des entités médicales

- b) l'identification des relations sémantiques existant entre les deux entités (maladie et traitement)
- c) la transformation des annotations en RDF (Resource Description Framework)

La reconnaissance des entités médicales se fait par le traitement des textes qui consiste en découpage des phrases et l'identification des syntagmes nominaux, la reconnaissance des termes médicaux et leurs types sémantiques et la classification des résultats selon les types sémantiques.

3.3.3. Les principales difficultés

Les difficultés liées au traitement automatique des langues sont principalement linguistiques. Il s'agit notamment de l'interprétation des mots et des expressions. Le dilemme du traitement automatique est le fait de mettre en place des outils linguistiques qui servent à éliminer les zones d'ombre en relation avec les aspects de référence, d'inférence, de présupposition et d'implication intervenant dans le dialogue homme-machine (J. Domengeot, P. Le Beux et G. Weil 1994).

Le vocabulaire de la médecine comporte un grand nombre de séquences figées. Certaines sont totalement opaques du point de vue sémantique comme (*pomme d'Adam*), d'autres le sont moins (*tendon d'Achille* et *dent de sagesse*) et d'autres encore sont totalement transparentes (*cuir chevelu*). Ces locutions n'ont pas le même degré de figement (S. Mejri 2005).

La compréhension des énoncés semble compliquée dans le dialogue entre les humains à cause du phénomène de sous-entendu et de présupposition. Elle devient encore plus compliquée quand il s'agit de dialogue homme-machine, à cause de la polysémie des verbes.

Si les bases de connaissances sont formées d'unités linguistiques dont les signifiés existent dans l'esprit des gens, en médecine, les difficultés semblent inévitables à cause de l'hétérogénéité du domaine de définition des termes. J. Domengeot, P. Le Beux et G. Weil (1994) notent qu'« il n'existe pas de consensus établi pour les définitions des symptômes, des syndromes, des maladies et des traitements. L'utilisation largement répandue des synonymes, des éponymes, des acronymes ou des noms propres complique largement les procédures de communication ou de compréhension entre observateurs ou effecteurs médicaux. Pour

permettre une communication efficace et dépourvue d'ambiguïté, un minimum de standardisation est nécessaire ». Cela serait encore plus compliqué pour le dialogue homme-machine qui nécessite une analyse fine, détaillée et rigoureuse des énoncés linguistiques. Certains énoncés qui paraissent évidents pour un humain semblent poser beaucoup de problème d'ambiguïté pour la machine. Le verbe *suivre* porte trois sens différents selon les exemples suivants : *L'enfant suit sa mère* (il marche derrière elle). *Le médecin suit Paul* (il le soigne). *Le policier suit le malfaiteur* (il le poursuit).

François Yvon (2006) souligne les difficultés liées au traitement automatique des langues sur plusieurs niveaux linguistiques. Ces difficultés concernent principalement le niveau lexical et le niveau syntactico-sémantique.

L'une des difficultés de l'analyse syntaxique dans une perspective de traitement automatique est l'ambiguïté. Ambiguïté lexicale : par exemple, *traité* peut renvoyer au participe passé du verbe *traiter* comme dans *avoir traité un patient* ou à un nom comme dans *traité de Rome* et *traité de Versailles*. L'ambiguïté grandit avec la précision de la description lexicale. Il est plus ou moins facile de désambiguïser les verbes dans les phrases.

M. Graitson et G. Dunham (1976) évoquent le problème lié à l'écriture des accents de certains ordinateurs anglo-saxons qui n'ont pas de représentation pour des graphèmes du français tels que l'accent grave, l'accent aigu, l'accent circonflexe, le tréma et la cédille. Ces derniers ne sont pas négligeables, ils sont même indispensables pour le traitement du texte parce qu'ils véhiculent des informations pertinentes. Pour illustrer cela, les exemples ne manquent pas comme la distinction entre le *à* préposition et *a* qui est une forme verbale, et *ou* pour lequel la présence ou l'absence de l'accent grave distingue entre *où* qui est pronom et *ou* qui est conjonction.

La présence de l'accent grave et de l'accent aigu permet la sous-catégorisation des suffixes en français notamment dans le vocabulaire médical. Leur absence rendrait ambiguë environ une trentaine de suffixes nominaux sans compter les adjectifs et les participes passés. Par exemple, le suffixe « *ère* » permet de distinguer la catégorie nominale (*ulcère*, *artère*, *ictère*, etc.) de la catégorie des adjectifs qui se terminent par « *éré* » (*macéré*, *régénéré*, *altéré*, etc.).

La sous-catégorisation des affixes n'est qu'une petite tâche qui permet à chaque fois de lever l'ambiguïté. L'accent aigu est souvent l'outil qui permet de distinguer les noms, les

adjectifs et les participes passés (*côte/ côté, greffe/ greffé*). L'accent aigu distingue aussi les adjectifs ou les participes passés masculin en *-é* qui s'écrivent au féminin en *-ée* (*vitaminé / vitaminée, atrophié / atrophiée*) des adjectifs et participes passés qui s'écrivent au féminin en *-e* et au masculin par la suppression de *-e* comme (*utérine / utérin, unie / uni*).

M. Graitson et G. Dunham (1976) évoquent aussi les ambiguïtés sémantiques qui peuvent être provoquées par l'ajout ou la suppression de l'accent aigu. Par exemple, les préfixes *ite* dans *bronchite* et *ité* dans *instabilité* (les deux mots sont considérés comme des noms dans un dictionnaire électronique). Mais *ite* est un suffixe rajouté en médecine à des bases savantes pour désigner une inflammation (*otite, bronchite, gingivite, conjonctivite, etc.*), et de sert de marque pour repérer les maladies inflammatoires dans un dictionnaire électronique médical.

Conclusion

Nous constatons à travers ce chapitre sur l'état de l'art de la langue de la médecine que ce domaine a bénéficié d'une prise en charge sous plusieurs angles, notamment du point de vue morphologique, sémantique et applicatif. Ce dernier concerne l'analyse et la description de la langue de la médecine pour la traduction, la recherche documentaire et le traitement automatique. Néanmoins, les études sur le traitement automatique de la langue de la médecine sont peu nombreuses. Notre objectif est de décrire et d'analyser le vocabulaire médical afin d'élaborer une interface d'interrogation, en exploitant les dictionnaires électroniques. Le but principal est de d'effectuer des recherches ciblées sur la terminologie médicale. Les recherches seront axées sur plusieurs modes de recherche, notamment du point de vue morphologique, sémantique et combinatoire.

CHAPITRE 3

MÉTHODOLOGIE

Introduction

Dans ce chapitre, consacré à la méthodologie, nous présentons les outils méthodologiques que nous utilisons dans l'élaboration de la thèse. Nous traitons trois questions fondamentales pour la réussite de cette étude. Il s'agit notamment du cadre théorique où nous reprenons et nous expliquons la théorie des trois fonctions primaires, notamment la fonction prédicative, la fonction argumentale et la fonction actualisatrice, dans le cadre de la phrase simple, ainsi que ses principes d'analyse. Nous expliquons ensuite la méthodologie suivie dans l'élaboration des dictionnaires. Pour ce faire, nous reprenons la méthodologie établie par P. A. Buvet (2009), en l'occurrence, celle relative aux descripteurs de définition et aux descripteurs de condition. Ensuite, nous évoquons la constitution des corpus et leurs types ainsi que les ressources exploitées pour leur réalisation. Enfin, nous abordons la question relative aux données, à leur structuration et à leur description.

1. Cadre théorique : modèle de données

Nous exposons ici le modèle de données du laboratoire LDI dans la perspective du traitement automatique des langues, notamment, les outils d'analyse, la théorie des trois fonctions primaires ainsi que la notion d'emploi.

1.1. Outils d'analyse

L'analyse des faits de langue est basée sur l'analyse phrastique. La phrase est considérée comme l'unité minimale d'analyse. Elle est toujours constituée d'un prédicat et souvent d'arguments et d'actualisateurs dont chacun a un emploi spécifique selon sa nature et sa position dans un énoncé.

1.1.1. La phrase élémentaire

En grammaire, la phrase est considérée comme l'unité privilégiée de l'analyse grammaticale (C. Marchello-Nizia 1979). Néanmoins, sa définition suscite quelques interrogations.

M. Galliot (1966) associe la notion de phrase à celle de proposition. Il rajoute que la phrase est dite simple lorsque elle est constituée d'une seule proposition, et complexe quand elle est constituée d'au moins de deux propositions. Pour A. Hamon (1964), « la phrase est un ensemble de mots qui a un sens complet et qui va d'un point à un autre. Elle peut être très brève, même réduite à un seul mot... »

Mais suffit-il de réduire la notion de phrase à une proposition ou à une suite de mots délimités par des points ?

L'avènement du traitement automatique des langues a suscité de nouvelles réflexions sur la notion de phrase. Il n'a pas seulement défini, mais il a aussi éclairé et délimité la notion de phrase. Le besoin de traitement automatique a fait en sorte qu'on définisse la phrase en tant que prédicat suivi éventuellement des arguments et des actualisateurs.

1.1.2. Le postulat

L'analyse des éléments linguistiques est fondée sur les trois fonctions primaires (la fonction prédicative, la fonction argumentale et la fonction actualisatrice) (S. Mejri 2009). Les trois fonctions primaires permettent de catégoriser les éléments linguistiques sur le plan syntactico-sémantique et d'expliquer le rôle qu'ils jouent dans la construction d'une phrase (P. A. Buvet 2011).

La fonction prédicative et la fonction argumentale rendent compte du contenu syntactico-sémantique de l'énoncé ; la fonction prédicative spécifie quel est l'élément qui joue le rôle du prédicat, c'est-à-dire l'élément qui fait la relation dans la phrase. La fonction argumentale stipule quels sont les éléments sélectionnés par le prédicat qui sont mis en relation avec ce dernier. Leur rôle est important pour le sens du prédicat.

La fonction actualisatrice permet d'insérer la phrase dans une situation d'énonciation. Elle prend en charge la catégorie du temps, de l'espace et de la personne, autrement dit, c'est l'insertion de la structure prédicat-argument dans une situation d'énonciation.

1.1.3. Principes d'analyse

La description du lexique consiste à analyser un grand nombre de données hétérogènes selon les mêmes principes développés par P. A Buvet (2009). Les principes d'analyse sont des règles méthodologiques qui permettent de décrire les unités lexicales à partir du postulat fondamental de la théorie. Ils sont nécessaires à l'élaboration de taxonomies suffisamment cohérentes pour faire apparaître des régularités et expliquer le fonctionnement des unités lexicales. Il y a six principes d'analyse ; leur ordre d'apparition n'est pas indifférent car ils s'impliquent successivement.

Le premier principe d'analyse s'énonce comme : « pas de métalangue hors de la langue ». Le deuxième principe stipule que « la phrase simple est le cadre d'analyse des unités lexicales ». Le troisième principe précise que « l'étude du lexique s'effectue conjointement sur les plans morphologique, syntaxique et sémantique » Le quatrième principe repose sur « le souci d'exhaustivité et de systémativité ». Le cinquième principe réside dans « la nécessité de l'homogénéité ». Le sixième et dernier principe porte sur 'l'analyse ascendante' (P. A. Buvet 2009).

1.2. La théorie des trois fonctions primaires

Cette théorie permet d'effectuer la description des éléments linguistiques dans le cadre de la phrase. Nous étudions successivement les notions de fonction prédicative, de fonction argumentale et de fonction actualisatrice.

1.2.1. Fonction prédicative

La notion de phrase repose sur le prédicat. Autrement dit, il n'y a pas de phrase sans prédicat. Il est considéré comme l'élément essentiel de la phrase. Selon Z. Harris (1976), toute phrase est constituée d'un opérateur linguistique et de ses arguments.

La notion de prédicat n'est pas forcément associée au verbe (G. Gross & M. Mathieu-Colas 2001). Un prédicat peut être verbal comme *panser* dans *l'infirmier panse la blessure*, nominal comme *blessure* dans *Paul a une blessure au genou* ou adjectival comme *blessé* dans *Paul est blessé*. La même racine prédicative *bless* peut fournir plusieurs prédicats avec des morphologies différentes (I. Sfar 2007).

Dans le vocabulaire médical, on considère comme prédicat tous les noms de maladies qui apparaissent dans les constructions avec le verbe *avoir* qui est considéré comme verbe support (P. A. Buvet 1998) :

Il a une amygdalite

Elle a un kyste aux ovaires

Mejri (2011)⁶ considère qu'un prédicat concerne toutes les parties du discours ainsi que la forme phrastique. Ainsi, on peut considérer que *fièvre de cheval* fonctionne comme prédicat dans l'énoncé suivant avec le verbe *avoir* :

Il a une fièvre de cheval

L'introduction des verbes appropriés comme *soigner*, *disséquer*, *diagnostiquer*, *cautériser*, *amputer*, *greffer*, etc., change l'emploi des unités linguistiques dans la phrase. Ce procédé appelé paraphrase transforme les noms de maladie qui sont initialement des prédicats nominaux en arguments, et les verbes appropriés deviennent des prédicats appropriés.

Il diagnostique l'adéphasie

Le kyste d'ovaires comprime les veines du bas-ventre

Il a atténué sa fièvre de cheval

1.2.2. Fonction argumentale

Le traitement linguistique comprend le classement et l'analyse des arguments, d'autant plus qu'ils ont un rôle indispensable du point de vue syntaxique et sémantique. Syntaxique parce qu'ils fonctionnent autour du prédicat qui fait la relation dans la phrase, et sémantique parce que ce sont les arguments qui donnent le sens au prédicat lorsqu'il est verbal.

Les substantifs du français ne sont pas syntaxiquement homogènes (P. A. Buvet 1998). Certains peuvent fonctionner, tantôt comme des prédicats et tantôt comme des arguments, alors que d'autres sont toujours des arguments. Les noms comme *main* (partie_corps), *sirop* (médicament) et *scalpel* (instrument_médical) ne fonctionnent jamais comme prédicats puisqu'ils sont des noms élémentaires. On les considère comme des arguments. Par contre, les

⁶ Dans ses cours dispensés à l'université Pris 13 dans le cadre du séminaire sur le figement lexical.

noms comme *sida* (maladie), *chimiothérapie* (soins) et *ablation* (opération_chirurgicale) peuvent s'interpréter soit comme des prédicats, soit comme des arguments.

Cette différence conditionne l'élaboration de deux dictionnaires électroniques, le premier sera consacré aux prédicats (PRED_DIC), le deuxième sera consacré aux arguments (ARGU_DIC), dotés des entrées et des descripteurs qui leur sont associés (P. A. Buvet 2009) auxquels s'ajoute le dictionnaire des humains du domaine de la médecine (HUM_DIC)⁷ et le dictionnaire encyclopédique du vocabulaire médical (ENCY_DIC).

1.2.3. Fonction actualisatrice

Les actualisateurs sont des éléments linguistiques, autres que les prédicats et les arguments, constituants de la phrase. Leur rôle et l'instanciation du prédicat et de ses arguments (P. A. Buvet 2009). Autrement dit, ils permettent l'insertion d'une structure prédicat-argument dans une situation d'énonciation. Les actualisateurs prennent en charge la personne, le genre, la modalité, le temps, l'aspect, etc. Nous verrons leurs types et leur fonctionnement dans la sous-section consacrée aux niveaux linguistiques, *cf. infra*.

1.3. La notion d'emploi

Nous évoquons, premièrement, les niveaux linguistiques d'analyse phrastique, notamment le niveau logico-sémantique, le niveau énonciatif et le niveau interprétatif. Deuxièmement, nous abordons la question des descripteurs de définition et des descripteurs de condition suivant la méthodologie appliquée aux dictionnaires, mise en place par P. A. Buvet (2009) dans le cadre du traitement automatique des langues.

1.3.1. Niveaux linguistiques

L'analyse syntactico-sémantique de la phrase se fait sur trois niveaux hiérarchiques. Le niveau logico-sémantique, le niveau énonciatif et le niveau interprétatif. Dans le premier cas, il s'agit de rendre compte du prédicat et des arguments constituant la phrase, dans le

⁷ Nous consacrons un dictionnaire aux noms d'humains médicaux. Ces derniers fonctionnent soit comme des prédicats soit comme des arguments.

deuxième, de ses actualisateurs qui inscrivent la phrase dans une situation de communication et dans le troisième, il s'agit du « sens situationnel », c'est-à-dire, du sens que peut avoir une phrase lorsqu'elle est insérée dans le discours (P. A. Buvet 1998)

On considère que dans toute phrase, il ya une relation logique entre le prédicat et ses éventuels arguments (Z. Harris 1976). Ainsi, on peut la schématiser par la chaîne suivante :

$$P = \text{Prédicat} (\text{Arg0} + \text{Arg1} + \text{Arg2} + \text{ArgN})$$

Dans les exemples suivants, formés uniquement de prédicats et des arguments, nous montrons qu'il existe une relation logique naturelle entre le prédicat et les arguments :

Paul amputer jambe

Paul prendre médicament

Médecin soigner malade

Les prédicats *amputer*, *prendre* et *soigner* entretiennent une relation logique avec les arguments qu'ils ont sélectionnés dans chaque exemple. D'ailleurs, la substitution de ces arguments entraîne une modification et une altération du sens.

**Paul amputer médicament*

**Paul prendre médecin*

**Malade soigner médecin*

À travers cette analyse, nous constatons que les prédicats sont des opérateurs qui entretiennent des relations logiques avec les entités (arguments).

Le niveau énonciatif est pris en charge par les actualisateurs. Ces derniers sont des unités lexicales et des unités grammaticales qui servent à insérer les phrases dans des situations d'énonciation. Les verbes sont actualisés par la conjugaison : *Paul est amputé d'une jambe*, les substantifs prédicatifs sont actualisés par les verbes supports : *le médecin procure des soins au malade*, les substantifs argumentaux par les déterminants *Paul est amputé d'une jambe* (P. A. Buvet 2009), et les adjectifs, par le verbe être : *Paul est malade*.

1.3.2. Descripteurs de définition

Les descripteurs de définition constituent la microstructure du dictionnaire. Il y a quatre types de spécification : a) la racine prédicative ; b) la classe sémantique ; c) le trait sémantique ; d) l'aspect inhérent. Les descripteurs de définition indiquent les propriétés sémantiques de chaque emploi des entrées (P. A. Buvet 2009).

Le tableau ci-dessous récapitule les différents descripteurs des deux premières parties du dictionnaire (PRED_DIC). La première colonne est subdivisée en deux sous-colonnes : la première sous-colonne comporte la forme lemmatisée de l'emploi, la deuxième un éventuel indice qui permet de dissocier les emplois qui ont la même forme. La seconde colonne est subdivisée en quatre sous-colonnes correspondant aux différents types de descripteurs de définition.

entrée		définition			
emploi	i n d i c e	racine prédicative	classe	type	aspect inhérent
<i>consulter</i>	1	consult-	EXAMEN_MÉDICAL	act	duratif perfectif
<i>consulter</i>	2	consult-	LECTURE	act	duratif perfectif
<i>diagnostic</i>	1	diagnosti-	DIAGNOSTIC	act	duratif perfectif
<i>diagnostic</i>	2	diagnosti-	DIAGNOSTIC	éta	abouti
<i>diagnostiqué</i>		diagnosti-	DIAGNOSTIC	éta	abouti

<i>diagnostiquer</i>		diagnosti-	DIAGNOSTIC	act	duratif perfectif
<i>cireux</i>		cir-	SYMPTOME	éta	provisoire
<i>fiévreux</i>		fièvr-	SYMPTOME	éta	provisoire
<i>fièvre</i>		fièvr-	SYMPTOME	éta	provisoire

- i) Un emploi prédicatif avec une seule forme morphologique prédicat et une seule racine. Son lemme (*cireux*) porte une seule information métalinguistique (symptôme).
- ii) Les emplois prédicatifs polysémiques comme les deux emplois de *consulter* qui ont la même racine (*consult-*) mais qui appartiennent à des classes différentes (procedure_médicale et lecture).
- iii) Les emplois prédicatifs qui se rapportent à un même prédicat. Ils ont des différences morphosyntaxiques ; le verbe *diagnostiquer* et les deux emplois du substantif *diagnostic* et l'adjectif *diagnostiqué* ont la même spécification de racine (*diagnosti-*) et la même spécification de classe (diagnostic).

Les interprétations des emplois sont le résultat de tous les descripteurs de définition, cf. *supra*.

- i) L'emploi de *cireux* s'interprète comme une propriété transitoire appropriée aux humains, du trait *état*, de l'aspect inhérent *provisoire* qui fonctionne avec un seul argument (partie_corps).

- ii) l'interprétation des deux emplois de *consulter* est proche : ils ont le même trait *action*, le même aspect inhérent *duratif perfectif* et un nom d'humain comme sujet ; ils se distinguent par la nature de leurs compléments, cf. *supra*.
- iii) Les emplois verbal, nominal et adjectival prédicatifs formés à partir de la même racine *diagnosti-* se rapportent tous au même prédicat mais ils s'interprètent différemment :

L'emploi verbal et l'emploi nominal ont des caractéristiques communes: ils ont le même trait sémantique *action*, le même aspect inhérent *duratif perfectif* et la même structure argumentale (un nom d'humain) en position sujet, un nom de maladie en position de complément direct et un nom d'humain en position de complément indirect. L'emploi adjectival a les caractéristiques suivantes : le trait sémantique *état*, l'aspect inhérent *abouti* et la structure argumentale prend un nom de maladie en position sujet.

1.3.3. Descripteurs de condition

Les descripteurs de conditions comprennent les trois sortes de spécification suivantes :

- e) la construction ; f) la distribution morphosyntaxique; g) la distribution syntactico-sémantique.

Les descripteurs de condition permettent de vérifier la structuration des arguments par rapport à l'emploi prédicatif, le nombre d'arguments que porte chaque prédicat ainsi que le caractère facultatif ou obligatoire de l'argument en position complément (P. A. Buvet 2009).

- Les emplois *cireux* et *diagnostiqué* ont la construction avec un support être **X0 être A** en tant qu'adjectifs prédicatifs qui n'acceptent qu'un sujet (**X0**).

- Les emplois du verbe *consulter* ont la construction **X0 V X1** en tant que verbes prédicatifs qui prennent (**X0**) en position de sujet et obligatoirement un (**X1**) complément d'objet direct.
- L'emploi *diagnostiquer* a la construction **X0 V X1 (prép X2)** en tant que verbe prédicatif qui accepte un sujet (**X0**) et obligatoirement un complément d'objet direct (**X1**).
- L'emploi *douleur* a la construction **X0 Vsup DET N (prép X1)** en tant que nom prédicatif qui accepte un sujet et éventuellement un complément circonstanciel.
- L'emploi *diagnostic* a la construction **X0 Vsup DET N** en tant que nom prédicatif qui n'accepte qu'un sujet **X0**.
- Le deuxième emploi de *diagnostic* a la construction **X0 faire DET N sur X1** en tant que nom prédicatif qui accepte un sujet (**X0**) et obligatoirement un complément d'objet indirect (**sur X1**).

Nous constatons dans notre corpus que ce genre de noms prédicatifs se combine souvent avec les verbes supports les plus canoniques : **X0 avoir DET N** pour *handicap*, **X0 faire DET N** pour *diagnostic* et **X0 ressentir DET N** pour *douleur*.

Le tableau ci-dessous rend compte de l'utilisation des descripteurs du type f) à propos des exemples étudiés.

Les descripteurs de la catégorie g) ont un rôle important dans la spécification des classes sémantiques et des sous-classes sémantiques, voire des hyperclasses des substantifs

des groupes nominaux en position **X0**, **X1** voire **X2** ou **X3**. L'hyperclasse (humain) caractérise les arguments nominaux qui occupent la position **X0** dans les constructions des emplois *consulter1*, *consulter2*, *diagnostic1*, *diagnostiquer*, *cireux* ; la classe (maladie) caractérise les arguments nominaux qui occupent la position **X0** dans les constructions des emplois *diagnostic2* et *diagnostiqué*.

Le tableau ci-dessous permet de récapituler les différents descripteurs de la première et de la troisième partie du dictionnaire PRED_DIC. La première colonne est subdivisée en deux sous-colonnes. La seconde colonne est subdivisée en trois sous-colonnes correspondant aux différents types de descripteurs de conditions.

entrée		conditions		
emploi	i n d i c e	construction	distribution 1	distribution 2
<i>cireux</i>		X0 être A	X0=GN	X0 = PARTIE_CORPS
<i>consulter</i>	1	X0 V X1	X0=GN X1=GN	X0 = HUMAIN X1 = HUMAIN_MÉDECIN
<i>consulter</i>	2	X0 V X1	X0=GN X1=GN	X0 = HUMAIN X1 = TEXTE
<i>diagnostiquer</i>		X0 V X1	X0=GN	X0 = HUMAIN

			X1=GN	X1 = MALADIE
<i>diagnostic</i>	1	X0 faire DET N de X1	X0=GN	X0 = HUMAIN X1 = MALADIE
<i>diagnostic</i>	2	X0 avoir DET N	X0=GN	X0 = MALADIE
<i>diagnostiqué</i>		X0 être A	X0=GN	X0 = MALADIE

2. Corpus : collection de données

Nous précisons quelles étaient nos sources d'information pour la constitution du corpus. Nous précisons ensuite les différents types de corpus et quelles sont les sources exploitées pour son élaboration.

2.1. Constitution du corpus

Pour constituer notre corpus, nous avons consulté des dictionnaires, des ouvrages spécialisés et les informations disponibles sur le web.

2.1.1. Dictionnaires

Nous avons utilisé *Le trésor de la langue française*. La consultation de ce dictionnaire est d'une grande utilité pour les études sur les langues de spécialités car cet ouvrage permet à la fois de collecter tous les termes d'un domaine particulier et d'obtenir tous les emplois possibles de chaque mot. Nous avons exploité également le dictionnaire *Le Petit Robert* pour chercher les emplois potentiels des verbes, des substantifs et des adjectifs recensés. Nous avons également exploité le *Dictionnaire illustré des termes de médecine* de Garnier Delamare. Ce dictionnaire est d'une grande utilité, surtout quand il s'agit de vérifier l'orthographe des mots et leurs définitions. Il fournit surtout les informations

extralinguistiques comme la localisation de certaines maladies et des parties du corps ainsi que leurs descriptions détaillées

2.1.2. Web

Le dictionnaire informatisé de la médecine a été très utile dans la recherche des noms de parties de corps et de tout ce qui relève de l'anatomie, ainsi que les noms de médicaments. On a aussi exploité des moteurs de recherche sur le Web comme Google et des sites de vente de matériel médical pour lister les instruments médicaux. Nous disposons surtout du *Trésor de la langue française informatisé (TLFi)*, dont l'utilisation est pratique avec ces modes de recherche, notamment la recherche assistée et la recherche complexe. Nous avons également utilisé *Doctissimo*, portail médical grand public sur la santé et le bien-être. Le dictionnaire médical *Doctissimo* contient énormément de définitions de termes médicaux comme les noms de maladies, les parties du corps, les médicaments, les symptômes, etc.

2.1.3. Littérature

La santé de l'être humain fait partie de ses préoccupations majeures dans la vie. Désormais, la santé fait partie des sujets les plus traités par les média comme la presse écrite. Sa large diffusion fait que des livres sont entièrement dédiés à l'étude et à l'analyse du vocabulaire médical.

Pour constituer notre corpus, nous avons consulté beaucoup d'articles de presse⁸. Les revues sur la santé sont abondantes, soit pour la diffusion de la recherche médicale entre professionnels, soit pour vulgariser quelques phénomènes médicaux compliqués afin de les expliquer au grand public.

Nous avons consulté quelques livres et articles destinés à l'étude des spécificités du vocabulaire de la médecine comme le livre de Joseph Ghazi (1985) sur la structure, le fonctionnement et l'apprentissage du vocabulaire médical.

⁸Notamment avec l'éclatement de l'affaire *Médiator*, des prothèses mammaires *PIP* ou bien de la pilule *Diane* 35.

2.2. Types de corpus

Pour analyser le vocabulaire médical, nous avons constitué trois types de corpus différents : un corpus spécialisé, un corpus non spécialisé et un corpus moyennement spécialisé. Nous montrons leurs spécificités ci-dessous.

2.2.1. Les termes spécialisés

Pour alimenter nos dictionnaires de mots spécialisés, nous avons collecté plusieurs termes sur la santé. Ces termes proviennent des textes des magazines sur la santé publique. Il s'agit de textes écrits et publiés par les médecins et les chercheurs en médecine. Ils sont écrits dans une langue soutenue, technique et spécialisée. Ce corpus concerne les communications ou les discours soutenus et pointus entre les médecins et les chercheurs.

Dans ce genre de revue, les chercheurs traitent des sujets variés. Ils n'écrivent pas uniquement sur les maladies qui menacent l'être humain mais aussi sur les nouveaux traitements mis en place ainsi que leurs effets indésirables, les situations sanitaires et les systèmes sanitaires en Europe, et surtout en Afrique et dans les pays asiatiques et latino-américains où les problèmes de santé sont considérables et la bonne santé des personnes reste un défi majeur pour les populations et pour les gouvernements.

Les textes spécialisés sont difficilement accessibles au public non initié parce qu'ils sont pleins de termes techniques et hautement spécialisés, notamment ceux qui proviennent du grec et du latin.

2.2.2. Les termes non spécialisés

Le corpus non spécialisé est présenté dans une langue de vulgarisation, dont le niveau n'est pas élevé. On utilise souvent des termes courants et des termes de la langue générale pour désigner des phénomènes compliqués que les gens ne connaissent qu'à travers leurs expériences en tant que patients. L'ignorance des patients vis-à-vis des termes spécialisés et techniques les incite à creuser dans leur vocabulaire général afin de trouver des substituts. Par exemple, on utilise *bleu* pour évoquer *hématome* et *ecchymose*. L'ignorance de ces derniers fait qu'on a cherché un terme dans la langue générale qui est «bleu», couleur provoquée

justement par les *ecchymoses* et les *hématomes* sur la peau. C'est la même chose pour les verbes *enlever* et *donner* que les patients ont dû utiliser par méconnaissance du substantif *ablation* et du verbe *prescrire*. Le terme « foie gras » est utilisé à la place de la *stéatose*. De même pour le *bruit des intestins* (borborygme) et *extraction* (avulsion).

2.2.3. Les termes moyennement spécialisés

Le corpus moyennement spécialisé est à mi-chemin entre le corpus spécialisé et le corpus non spécialisé. Nous l'avons constitué par la collecte de termes parus dans la presse sur la santé en général et sur les maladies et les médicaments en particulier.

Dans ce cas, les textes ne sont pas rédigés dans une langue soutenue mais standard, puisqu'ils sont destinés au grand public qui n'a pas forcément assez de connaissances dans le domaine médical. Les textes contiennent plutôt des termes courants et peu techniques. On utilise plutôt *grippe porcine* au lieu *H1N1*, *grippe aviaire* à la place *H5N1* ou *cancer* à la place de *néoplasme*.

Ce genre de texte ne contient pas beaucoup d'abréviations et de sigles, comme l'atteste notre corpus. Cela facilite la compréhension de la terminologie médicale aux non initiés.

2.3. Sources

Les sources de notre corpus sont multiples et variées. Nous avons exploité des bases de données, les forums sur la santé, les revues spécialisées et les revues de vulgarisation pour la collecte des entrées des dictionnaires que nous avons élaborés.

2.3.1. Bases de données

Le vocabulaire médical collecté pour l'élaboration de nos dictionnaires PRED_DIC, ARGU_DIC et HUM_DIC provient essentiellement des bases de données comme l'arborescence MeSH. Elle fonctionne par domaines et spécialités biomédicales comme *anatomie*, *maladies* et *organismes*, etc.

Le portail médical Doctissimo contient beaucoup de termes relatifs à la santé comme la liste des maladies et des médicaments.

L'encyclopédie *wikipédia* fournit plusieurs listes de maladies classées par spécialités et organes du corps touchés par les maladies en question, par exemple :

- *Liste des maladies dermatologiques*
- *Liste des maladies cancérigènes*
- *Liste des maladies et troubles oculaires*
- *Liste des maladies génétiques (à gène identifié, à gène non identifié)*
- *Liste des maladies infectieuses*
- *Liste des maladies vulvo-vaginales*
- *Liste des maladies mentales*
- *Liste des troubles de l'humeur*
- *Liste des troubles de la personnalité*
- *Liste de maladies rares*

2.3.2. Forums

Nous avons utilisé les forums pour regrouper les termes non spécialisés, comme *Doctissimo*, portail médical grand public sur la santé et le bien-être où les patients se rencontrent pour parler éventuellement de leurs maladies, soit pour obtenir des renseignements de personnes souffrant déjà de la même maladie, soit pour informer les malades et leur apporter un réconfort moral, après avoir éprouvé une maladie quelconque ou subi une opération chirurgicale. Les patients qui ne connaissent pas bien leurs maladies essaient parfois de s'informer sur d'éventuelles complications ou sur les risques encourus.

Nous avons extrait aussi des mots du site internet *www.e-santé.fr* dans son forum réservé au grand public, dont les objectifs sont les mêmes que ceux de *Doctissimo*.

2.3.3. Revues

La revue *santé publique* est une grande source de termes spécialisés exploités dans ce travail de recherche. C'est une revue bimestrielle et multidisciplinaire écrite en français. La revue santé publique est soutenue par la *Société française de santé publique*. Elle est destinée aux chercheurs, aux professeurs et aux professionnels du domaine de la santé. Sa vocation

principale est de soutenir la recherche en santé publique, de favoriser la diffusion des connaissances et du savoir dans le domaine et de faciliter les échanges de connaissances et d'expériences entre professionnels.

3. Dictionnaire : structuration et description des données

Tout dictionnaire doit être jugé sur l'ensemble de ses entrées et sur les informations que comporte chaque entrée (N. Campos Martin 2009). Nous montrons ici comment sont structurés les dictionnaires en général et les données de notre corpus en particulier.

3.1. Structuration

A la différence des dictionnaires papiers, les dictionnaires électroniques ont une structuration et une organisation particulière qui répondent à des besoins spécifiques, notamment concernant le TAL.

3.1.1. Dictionnaires électroniques

La description du lexique dans les dictionnaires est faite d'une façon relativement exhaustive. Nous avons élaboré un dictionnaire des emplois prédicatifs (PRED_DIC) qui contient des verbes et des substantifs, un dictionnaire des arguments élémentaires (ARGU_DIC) qui est constitué uniquement de substantifs ayant un emploi argumental, un dictionnaire des noms d'humains qui fonctionnent comme des prédicats pour certains et comme arguments pour d'autres, (HUM_DIC) et un dictionnaire encyclopédique qui regroupe les termes des trois premiers dictionnaires (ENCY_DIC).

Pour la sélection des unités lexicales, on notera que la langue de la médecine, ainsi que tous les textes spécialisés est constituée de trois types de vocabulaires :

- des mots de la langue générale qu'on utilise dans la langue de spécialité, et qui occupent une place importante dans les textes spécialisés.

- des mots spécifiques à la langue de spécialité : *hospitaliser, soigner, cautériser, anesthésier, panser, guérir, clinique, hôpital, vaccin, etc.*

- des mots de la langue générale, mais qui portent un sens particulier dans la langue spécialisée : *prendre* un médicament, *circulation* du sang, la douleur *se réveille*, le patient *se remet*, douleur *aigüe*.

À partir des dictionnaires que nous avons élaborés, nous avons constitué des dictionnaires plurifonctionnels. C'est-à-dire qu'ils fournissent à la fois des informations linguistique (morphologie, syntaxe et sémantique) et des informations médicales. Ces dernières sont disponibles dans le dictionnaire encyclopédique qui regroupe les termes de tous les dictionnaires avec les champs suivants : entrée, classe sémantique, domaine et sous-domaine. Il permet aux utilisateurs de l'interface d'effectuer des recherches particulières sur les termes médicaux.

3.1.2. Macrostructure

La constitution du corpus a été la première étape de notre travail. Nous avons établi différentes listes d'unités lexicales du domaine de la médecine à partir des sources d'information présentées ci-dessus. Nous avons collecté :

a) des noms prédictifs

- des noms de maladies comme *diabète*
- des noms d'examen médicaux comme *biopsie*
- des noms de professions comme *aide-soignant*
- etc.

b) des noms argumentaux

- des noms de partie du corps comme *pied*
- des noms d'appareils comme *stéthoscope*
- des noms de lieux comme *clinique*
- etc.

c) des verbes prédicatifs

- *diagnostiquer*
- *guérir*
- *soigner*
- *intuber*
- *amputer*
- etc.

d) des adjectifs prédicatifs

- *cancéreux*
- *diabétique*
- *sidatique*
- etc.

Une fois nos listes constituées, nous avons étudié chaque unité lexicale en contexte à l'aide du moteur de recherche Google pour préciser leurs conditions d'emplois dans les textes. Nous nous sommes particulièrement attaché à décrire la polysémie et la polymorphie prédicatives. La polysémie est une caractéristique du vocabulaire de la langue générale. Le verbe *traiter* change de sens selon qu'il est relié à des arguments : *le médecin traite un patient* qui est équivalent à *soigner*. *Luc traite un sujet*, qui équivaut à *aborder un sujet*. Le prédicat peut avoir plusieurs formes morphologiques différentes (verbale, nominale et adjectivale) : *le*

médecin soigne Luc, Luc est soigné par le médecin, Luc a subi des soins. Le prédicat autonome n'a qu'une seule forme morphologique.

3.1.3. Microstructure

L'ensemble des dictionnaires que nous avons élaborés contient toutes sortes d'entrées. Ici, il faut tenir compte de l'ensemble des informations que contient chaque entrée du dictionnaire.

Les entrées du dictionnaire des prédicats comportent des informations qui les définissent et qui conditionnent leurs emplois dans des énoncés corrects (P.A. Buvet 2009). Elles sont d'ordre sémantique, syntaxique et morphologique.

Les entrées du dictionnaire des arguments comportent des informations d'ordre sémantique et morphologique.

Les entrées du dictionnaire des humains sont définies par des informations relatives à leur morphologie, leur classe sémantique, leur domaine et leur sous-domaine.

3.2. PRED_DIC

Nous indiquons les principales spécificités du dictionnaire PRED_DIC, ensuite nous présentons les spécificités des verbes, des noms et des adjectifs qui constituent le dictionnaire.

3.2.1. Particularités du dictionnaire

Nous présentons sous forme de tableaux les différentes données chiffrées relatives à la constitution de notre dictionnaire (PRED_DIC). Nous détaillons ensuite les différents descripteurs utilisés.

<i>champ</i>	<i>quantité</i>
<i>entrée</i>	<i>11300</i>
<i>racine</i>	<i>1608</i>
<i>trait</i>	<i>3</i>
<i>classe</i>	<i>34</i>
<i>aspect inhérent</i>	<i>4</i>
<i>construction</i>	<i>25</i>
<i>X0 (morphosyntaxe)</i>	<i>11300</i>
<i>PREPI</i>	<i>2316</i>
<i>X1 (morphosyntaxe)</i>	<i>2392</i>
<i>X0 (sémantique)</i>	<i>11300</i>
<i>X1 (sémantique)</i>	<i>2392</i>
<i>Mots simples</i>	<i>258</i>
<i>Mots composés</i>	<i>9679</i>
<i>Mots dérivés</i>	<i>1360</i>
<i>Sigles</i>	<i>02</i>
<i>Morphologies</i>	<i>342</i>
<i>Domaines</i>	<i>11</i>
<i>Sous-domaines</i>	<i>123</i>

Les classes sémantiques qui caractérisent les mots du dictionnaire sont :

- ADDICTION (*abus de cannabis*)
- ANESTHÉSIE (*anesthésie locale*)
- CONSULTATION (*consulter*)
- CRISE (*crise épileptique*)
- DIAGNOSTIC (*diagnostic précoce*)
- ÉPIDÉMIE (*épidémie*)
- Etc.

Les types sémantiques qui caractérisent les mots du dictionnaire sont :

- état (*anémie*)
- événement (*lèpre*)
- action (*soigner*)

Les aspects inhérents qui caractérisent les mots du dictionnaire sont :

- abouti (*guérir*)
- duratif perfectif (*opérer*)
- provisoire (*grippe*)
- permanent (*diabète*)

Les constructions qui caractérisent les mots du dictionnaire sont :

- X0 avoir DET N
- X0 être V Prép X1
- X0 être A Prép DET N
- X0 V sup DET N Prép X1

Les constructions qui caractérisent les verbes du dictionnaire sont :

- X0 V X1 (*le médecin suit le patient*)
- X0 V (*le médecin exerce*)

Les morphologies qui caractérisent les mots du dictionnaire sont :

- Mot simple
- Mot composé [type : N de N]
- Mot composé [type : N A]
- Etc.

Les domaines qui caractérisent les mots du dictionnaire sont :

- Médecine
- Paramédical
- Pharmacologie
- Etc.

Les sous-domaines qui caractérisent les mots du dictionnaire sont :

- Chirurgie
- Endocrinologie
- Gynécologie
- Etc.

3.2.2. Noms et Adjectifs

Si l'on compare les propriétés morphologiques des noms et adjectifs d'une langue de spécialité et ceux d'une langue générale, les différences sont minimes notamment en ce qui concerne les procédés de formation des mots comme la dérivation et la composition.

Notre corpus est composé à la fois de noms simples comme *scalpel* et *seringue* et de noms composés comme *infection rénale* et *maladie pulmonaire*.

La composition est une juxtaposition de deux mots libres qui permet de former un troisième mot à part entière, et dont le sens ne se laisse pas forcément deviner par celui des deux constituants. On distingue généralement deux types de composition : la composition savante et la composition populaire. Ainsi les noms de maladie relèvent généralement de la composition savante car ils sont d'origine latine ou grecque : *naevus dysplasique*, *noduliter humatoïde*, *ictus amnésique*, *criger-najjare*. Cependant, d'autres noms de maladie relèvent de

la composition populaire dans les tableaux des noms de maladies comme : *grippe aviaire, grippe porcine, trouble panique, etc.*

Certains noms de maladies sont formés par siglaison. Ils sont d'origine phrastique ; le procédé de siglaison est utilisé par souci d'économie. Certains sigles comme SEP (*sclérose en plaques*) sont techniques et réservés aux spécialistes. D'autres comme SIDA (*syndrome d'immunodéficience acquise*) sont largement connus en dehors du domaine médical. Le mot SIDA a donné d'autres termes à partir du radical *sida* par exemple les adjectifs *sidatique, sidéen*.

En français, la dérivation consiste à former un nouveau mot à partir d'un autre mot en lui adjoignant un suffixe ou un préfixe. Ainsi, beaucoup d'adjectifs sont formés à partir des noms de maladies *lèpre/lépreux, cancer/cancéreux, diabète/diabétique, hémorragie/hémorragique, boulimie/boulimique, aphasie/aphasique, anorexie/anorexique, etc.*

3.2.3. Verbes

Nous avons distingué trois types de verbes dans la langue de la médecine. Premièrement, des verbes appartenant exclusivement à la langue de la médecine, et qui sont monosémiques, par exemple *hospitaliser, guérir, panser, cautériser, etc.* ; deuxièmement, les verbes de la langue générale qui sont très utilisés en médecine, dont le nombre reste indéterminé ; troisièmement, les verbes de la langue générale qui ont un emploi et un sens spécifique dans la langue de la médecine comme *aller chez le médecin, voir un médecin, faire une anesthésie, prendre un traitement, appliquer un pansement, etc.*

Nous constatons que certains verbes spécialisés sont à cheval entre deux domaines spécialisés différents. Ainsi le verbe *lifter* est utilisé en chirurgie esthétique, où il désigne une opération qui consiste à supprimer les rides du visage ; au tennis, il désigne une manière particulière de frapper la balle avec la raquette.

Le verbe *greffer* est utilisé en médecine dans le sens de procurer des soins, mais aussi en botanique. Il signifie transplanter une branche d'arbre au tronc.

Le verbe *soigner* est utilisé en langue générale comme *soigner son look, soigner sa tenue* pour dire *devenir élégant*, ou *soigner son image* qui signifie *montrer ses qualités*.

3.3. ARGU_DIC, HUM_DIC et ENCY_DIC

Nous présentons ci-dessous les spécificités du dictionnaire des arguments (ARGU_DIC), du dictionnaire des humains (HUM_DIC) et du dictionnaire encyclopédique (ENCY_DIC), ainsi que les principales informations qui distinguent leurs entrées.

3.3.1. ARGU_DIC

<i>champ</i>	<i>quantité</i>
<i>entrée</i>	5580
<i>genre</i>	4
<i>morphologie</i>	189
<i>classe sémantique</i>	47
<i>domaine</i>	9
<i>sous-domaine</i>	68

Les genres qui caractérisent les noms du dictionnaire sont :

- Le masculin (*os*)
- Le féminin (*cellule*)
- Le masculin pluriel (*cônes de Haller*)
- Le féminin pluriel (*cordes vocales*)

Les combinaisons morphologiques qui caractérisent les noms du dictionnaire sont :

- Nom simple (*tête*)
- Nom composé du type : [ANA] (*condyle de la mandibule*)
- Nom composé du type : [N de NAA] (*commisure des tubercules quadrijumeaux postérieurs*)

- Nom composé du type : [NAA de N] (*artériole nasale inférieure de la rétine*)
- Nom composé du type : [N de N] (*colonne de Bertin*)
- Nom composé du type : [NA] (*artère vésicale*)
- Etc.

Les classes sémantiques qui caractérisent les noms du dictionnaire sont :

- VIRUS (*virus de l'influenza type A*)
- PARTIE_CORPS (*main*)
- PARTIE_CORPS_GLANDE (*glande endocrine*)
- PARTIE_CORPS_VAISSEAU (*artère cérébrale moyenne*)
- ÉQUIPEMENT_MÉDICAL (*brancard*)
- APPAREIL_MESURE (*stéthoscope*)
- Etc.

Les domaines qui caractérisent les noms du dictionnaire sont :

- Médecine
- Chirurgie dentaire
- Biologie
- Pharmacologie
- Etc.

Les sous-domaines qui caractérisent les noms du dictionnaire sont :

- Anatomie
- Microbiologie
- Chirurgie
- Etc.

3.3.2. HUM_DIC

<i>champ</i>	<i>quantité</i>
<i>entrée</i>	621
<i>morphologie</i>	30
<i>classe sémantique</i>	33
<i>domaine</i>	5
<i>Sous-domaine</i>	36

Les morphologies qui caractérisent les noms du dictionnaire sont :

- Nom simple (*sourd*)
- Nom composé du type : [N de N] (*médecin de travail*)
- Nom composé du type : [NA] (*médecin urgentiste*)
- Nom composé du type : [NAA] (*corps enseignant dentaire*)
- Nom composé du type : [N de NA] (*personnel de santé auxiliaire*)
- Nom composé du type : [N en N] (*étudiants en dentisterie*)
- Etc.

Les classes sémantiques qui caractérisent les noms du dictionnaire sont :

- HUMAIN (*personne âgée fragile*)
- HUMAIN_MALADE (*hypotendue*)
- HUMAIN_MÉDECIN (*médecin*)
- HUMAIN_ANÉSTHÉSISTE (*anesthésiste*)
- HUMAIN_CHIRURGIEN (*chirurgien*)
- HUMAIN_DENTISTE (*dentiste*)
- Etc.

Les domaines qui caractérisent les noms du dictionnaire sont :

- Médecine
- Chirurgie dentaire
- Biologie
- Pharmacologie
- Etc.

Les sous-domaines qui caractérisent les noms du dictionnaire sont :

- Analyse médicale
- Cancérologie
- Cardiologie
- Chirurgie
- Chirurgie dentaire
- Dermatologie
- Etc.

3.3.3. ENCY_DIC

<i>champ</i>	<i>quantité</i>
<i>entrée</i>	<i>17449</i>
<i>classe sémantique</i>	<i>123</i>
<i>domaine</i>	<i>17</i>
<i>sous-domaine</i>	<i>171</i>

Les classes sémantiques qui caractérisent les noms du dictionnaire sont :

- MALADIE (*hépatite*)
- HUMAIN_MALADE (*hypotendue*)
- PARTIE_CORPS (*main*)
- Etc.

Les domaines qui caractérisent les noms du dictionnaire sont :

- Médecine
- Biologie
- Pharmacologie
- Etc.

Les sous-domaines qui caractérisent les noms du dictionnaire sont :

- Analyse médicale
- Cancérologie
- Cardiologie
- Chirurgie
- Chirurgie dentaire
- Dermatologie
- Etc.

Conclusion

Notre méthodologie est fondée sur l'analyse de la phrase, du point de vue des trois fonctions primaires. Cette théorie permet de rendre compte de chaque élément de la phrase. Elle est la clé de voûte de notre analyse et de la description des entrées des dictionnaires que nous avons élaborés.

CHAPITRE 4

LES EMPLOIS PRÉDICATIFS DE LA LANGUE DE LA MÉDECINE

Introduction

Ce chapitre est consacré à l'analyse des prédicats de la langue de la médecine, quelque soit leur forme : les prédicats nominaux (*maladies, examens médicaux, opérations chirurgicales, etc.*), les prédicats verbaux (*soigner, traiter, hospitaliser, etc.*) et les prédicats adjectivaux (*diabétique, anorexique, cancéreux, etc.*).

Les prédicats sont étudiés sous trois angles différents. Nous étudions leurs propriétés morphologiques, leurs propriétés sémantiques et enfin leurs propriétés combinatoires. Nous classons, ensuite, les prédicats de la langue de la médecine en classes pour montrer que chacune d'elles possède des caractéristiques spécifiques et des critères sémantiques définitoires, notamment des verbes appropriés à chaque classe.

Nous faisons une description normalisée des prédicats où le sens de ses derniers est imputable à des descripteurs de définition, c'est-à-dire, les propriétés sémantiques des emplois prédictifs (racine, classe, trait et aspect inhérent), et les descripteurs de conditions (distribution et construction) (P. A. Buvet 2009).

1. Propriétés morphologiques

Premièrement, nous abordons le phénomène de la dérivation affixale et de la dérivation impropre. Deuxièmement, nous traitons des questions qui se rapportent à la composition savante, à la composition non savante et à la composition mixte.

1.1. Dérivation

Nous étudions le phénomène de la dérivation suffixale, préfixale et de la dérivation impropre du vocabulaire de la médecine dans le but de préciser les régularités et les nuances apportées par les affixes du point de vue de la signification.

L'étude des affixes donne des indications sémantiques. L'identification des suffixes ou des préfixes permet de catégoriser certaines classes comme dans le cas du suffixe *ite* ajouté à une base savante pour signifier une inflammation, *dermatite*, *gingivite*, *otite*, *hépatite*, *vaginite*, etc. Cette analyse permet d'identifier ces maladies inflammatoires.

La dérivation est un procédé linguistique qui consiste à former des mots nouveaux par ajout d'affixes à un mot appelé base ou racine. Autrement dit, la dérivation est la combinaison d'éléments linguistiques de statuts différents ; il s'agit de la base et des affixes : *guérir*, *guérison*. Lorsqu'on obtient un nouveau mot par ajout simultané de suffixes et de préfixes, on parle de parasyntèse, comme *intubation* qui résulte de l'ajout simultané du préfixe *in* et du suffixe *-tion* (J. Ghazi 1985). La dérivation impropre se fait sans modification de la forme des termes, mais par changement de catégorie grammaticale comme dans *il est malade* (adjectif), *j'ai vu le malade* (nom).

1.1.1. Dérivation suffixale

La dérivation suffixale est très fréquente en langue de spécialité, notamment dans le vocabulaire médical, comme l'attestent nos dictionnaires. De nombreux termes de la médecine sont des dérivés suffixaux. On ajoute des suffixes à des bases pour former de nouveaux mots. Certains suffixes appartiennent à la langue générale. Ils sont ajoutés à des termes de la langue de la médecine. Ce faisant, ils contribuent à la formation des termes médicaux spécialisés. Par exemple : *ophtalmologiste*, *oculiste*, *dermatologiste*, etc.

D'autres suffixes sont spécifiques à la médecine. Dans tous les cas de figure, les suffixes sont porteurs de sens. Par exemple, certains d'entre eux signifient « inflammation » comme *ite*, ou « maladie » comme *ose*. Dans le tableau qui suit, nous illustrons différents cas de figures avec pour chaque suffixe un exemple.

<i>suffixe</i>	<i>exemples</i>
<i>-tomie</i>	<i>colpotomie, épisiotomie, ovariectomie</i>
<i>-ite</i>	<i>otite, dermatite</i>
<i>-algie</i>	<i>gastralgie,</i>
<i>-ome</i>	<i>adénome</i>
<i>-ectomie</i>	<i>ovariectomie</i>
<i>-émie</i>	<i>glycémie</i>
<i>-ose</i>	<i>arthrose</i>
<i>-tion</i>	<i>luxation, auscultation</i>
<i>-isme</i>	<i>arthritisme, vaginisme</i>
<i>-ance</i>	<i>corticorésistance</i>
<i>-ase</i>	<i>aldolase</i>

La classification des suffixes significatifs spécialisés permet d'établir des classes sémantiques du vocabulaire médical. Clairement, si l'on ne prend que l'exemple des maladies, elles se distinguent en partie par leurs suffixes. Le suffixe *-ite* s'applique aux maladies inflammatoires, *-ose* s'applique aux maladies chroniques non inflammatoires. Le suffixe *-ome* est souvent ajouté à des bases savantes pour désigner *tumeur*, et *-algie* s'ajoute à des bases savantes pour désigner *douleur*.

Nous trouvons aussi des suffixes de la langue générale tels que *-age* (sevrage), *-tion* (luxation), *-isme* (corticorésistance), etc.

1.1.2. Dérivation préfixale

L'observation des dictionnaires a montré que les préfixes servant à former les termes de la médecine sont en majorité des préfixes de la langue générale comme le *a* privatif. Nous illustrons cela dans le tableau des préfixes ci-dessous.

<i>préfixes</i>	<i>exemples</i>
<i>a-</i>	<i>aménorrhée (absence de règles)</i>
<i>ab-</i>	<i>abaxial (loin de l'axe d'un organe)</i>
<i>acro-</i>	<i>acrocéphalie (malformation de la tête)</i>
<i>hyper-</i>	<i>hypertension</i>
<i>hypo-</i>	<i>hypotension</i>
<i>dys-</i>	<i>dysfonctionnement</i>
<i>anti-</i>	<i>antibiotique</i>
<i>para-</i>	<i>paramédical</i>
<i>dé-</i>	<i>désintoxication, déshydratation</i>
<i>ré-</i>	<i>rééducation</i>
<i>poly-</i>	<i>polyarthrite</i>

Contrairement aux suffixes, les préfixes ont souvent une signification non spécialisée comme *a-* et *ab-*. Ils sont ajoutés à des racines savantes ou populaires ;

- Les termes auxquels on rajoute le *a* désignent la privation ou l'absence comme *anomalie* et *aménorrhée*, etc. Le *a* privatif est ajouté parfois à des bases savantes pour désigner des états pathologiques comme *aboulie* et *anaphrodisie*, etc.

- Le préfixe *ab-* signifie l'éloignement comme l'attestent, par exemple les termes : *absence* et *abaxial*.
- *Acro-* signifie *élevé* ou *extrême*, d'où l'exemple *acrocéphalie* qui désigne la *malformation de la tête*.
- *Hyper-* et *hypo-* signifient respectivement *hausse* et *baisse*. On les utilise pour désigner la hausse et la baisse du taux de glycémie dans le sang. Ils désignent également la hausse et la baisse de la tension artérielle.
- Le préfixe *anti-* désigne l'action contraire, par exemple *anti-inflammatoire* et *antihistaminique*, etc.⁹
- Le préfixe *dys-* désigne dans le vocabulaire médical une difficulté ou une gêne d'un organe ou d'un système quelconque, par exemple *dyspnée* et *dyscrasie*, etc. En langue générale, il est ajouté au substantif *fonctionnement* (dysfonctionnement) afin d'évoquer une anomalie du fonctionnement.
- Le préfixe *para-* signifie *à côté de*. Il est utilisé en médecine (*paramédical*) pour désigner le domaine ou le personnel qui accompagnent les médecins dans leurs pratiques.
- Le préfixe *dé-*, signifie un changement ou une modification. Il peut avoir une valeur positive ou une valeur négative, par exemple *désintoxication* (valeur positive), *déshydratation* (valeur négative).
- Le préfixe *ré-* désigne la répétition des actions (vacciner, revacciner), ou des événements (tomber malade, retomber malade).
- Le préfixe *poly-* indique la diversification, *polyarthrite* signifie « inflammation qui touche simultanément plusieurs articulations ».

⁹ Anti-inflammatoire et antihistaminique relèvent de la parasynthèse parce qu'ils comportent un préfixe et un suffixe.

En conclusion, nous dirons que les préfixes et les suffixes sont importants dans le classement sémantique des prédicats de médecine. Ils sont néanmoins différents : les suffixes désignent souvent des concepts spécialisés alors que servent plutôt à former des unités lexicales appartenant à la langue générale.

1.1.3. Dérivation impropre

La dérivation impropre consiste à créer un nouveau terme uniquement par changement de catégorie grammaticale. Les catégories grammaticales les plus exploitées par la dérivation impropre sont celles du nom et de l'adjectif (J. Ghazi 1985). Par exemple, les adjectifs *addisonien*, *aliéné*, *anorexique*, *anémique*, *aphasique*, etc., sont utilisés comme des noms : *un addisonien*, *un aliéné*, *un anorexique*, *un anémique*, *un aphasique*, etc. Ces différents substantifs servent à désigner les personnes qui ont la maladie spécifiée sous une forme adjectivale.

La distinction entre ces deux catégories grammaticales est imputable au contexte. Les adjectifs du français sont habituellement actualisés par le verbe être :

Paul est addisonien

Paul est aliéné

Paul est anorexique

Paul est anémique

Paul est aphasique

Etc.

Tandis que les noms sont en partie actualisés par les déterminants :

L'addisonien consulte un médecin

L'aliéné consulte un médecin

L'anorexique consulte un médecin

L'anémique consulte un médecin

L'aphasique consulte un médecin

Dans le discours, les locuteurs ont tendance à appliquer cette méthode à tous les adjectifs d'état de la langue de la médecine. Néanmoins, le TLF (Trésor de la Langue Française) ne généralise pas cette pratique. Selon ce dictionnaire, beaucoup d'adjectifs en médecine ne fonctionnent pas comme des substantifs : *comateux, délirant, allergique, éruptif*, etc.

En résumé, certains adjectifs de la langue de la médecine peuvent fonctionner comme des noms avec une détermination appropriée (P. A. Buvet 2009).

Paul est un grand délirant

1.2. Composition

L'enseignement du vocabulaire de la médecine, notamment aux étudiants dans les facultés de médecine et de traduction consiste à présenter les significations des termes sans évoquer les questions de morphologie, et particulièrement les procédés de formation des mots. Pour un linguiste, la signification est importante, mais la morphologie est tout aussi primordiale (Lee-Jahnke 2001).

Pour décrire de façon précise les locutions nominales, on s'intéresse principalement à leur construction interne. Les locutions nominales ont des constructions différentes et variées. On distingue généralement trois sortes de mots composés : les composés savants, les composés non savants et les composés mixtes.

La diversité du vocabulaire médical et la liberté de formation des noms composés ont donné lieu à un grand nombre de locutions nominales. Le besoin de dénommer les maladies, les parties du corps, les organes, les examens médicaux et les opérations chirurgicales, etc. a donné le champ libre aux formations de locutions nominales dans le domaine de la médecine (G. Gross 1996).

1.2.1. La composition savante

Les mots composés appelés « mots composés savants » résultent de la juxtaposition de deux mots ou plus d'origine grecque ou latine.

Concernant leurs constructions internes, il existe plusieurs types de locutions nominales. Ainsi, parmi les noms de spécialités en médecine : *traumatologie*, *neurologie*, *dermatologie*, *urologie*, *pathologie*, etc. La composition savante est très fréquente. Il s'agit de mots composés de type NN (nom + nom), c'est-à-dire formés de deux radicaux, autrement dit, chacun des exemples est formé de deux termes significatifs et indépendants, *traumato* « blessure » et *logie* « science » dans *traumatologie*, et *pathos* « maladies », *logie* « science » dans *pathologie*, etc.

D'autres mots composés savants sont formés de deux mots autonomes, significatifs et séparés par un blanc, par exemple *fusarium langsethiae*, (mot d'origine latine donné à un genre de champignon).

Un autre procédé utilisé dans la formation des mots composés savants donne lieu aux composés du type NA (nom + adjectif) qui sont très fréquents. On est en présence de deux mots autonomes ; le nom et l'adjectif relationnel postposé qui s'unissent pour former une unité polylexicale, par exemple : *fusarium poae*, *fusarium solani*, *fusarium tricinctum*, (champignon), *arthrite psoriasique* (maladie), *artériopathies chronique* ; ce dernier exemple est constitué à la fois d'un mot composé uni qui paraît comme un mot simple (*artériopathie*, qui signifie maladie des artères).

On trouve aussi d'autres types de locutions nominales plus complexes du point de vue morphologique comme :

- Les locutions nominales de type N de N (nom de nom) *aphasie de Broca*, *aphasie de Wernicke*, *ataxie de Friedrich*, *atrésies des choanes*.
- Les locutions nominales de type N de NAA (nom de nom adjectif adjectif) *stomates de drainage lymphatique péritonéal*.
- Les locutions nominales de type NA-A-A (nom adjectif-adjectif-adjectif) *atrophies alivo-ponto-cérébelleuses*.
- Les locutions nominales de type NAA non A (nom adjectif adjectif non adjectif) *coma hyperosmolaire hyperglycémique non cétosique*. L'utilisation de la négation dans ce type de locution est due au fait qu'il existe un type particulier de *coma* qui ne partage pas toutes les caractéristiques des autres types de *coma*.

1.2.2. La composition non savante

La langue de la médecine comporte également de nombreux mots composés non savants, c'est-à-dire formés à partir des mots du français de tous les jours.

On constate que les mots composés non savants sont nettement plus nombreux que les composés savants ; cela semble dû à la croissance forte et rapide du vocabulaire médical qui implique une certaine liberté dans la formation des mots composés. Les progrès de la médecine ces dernières décennies ne sont pas restés sans impact sur le vocabulaire médical (J-P. Boissel & F. Dürr 2004).

Nous relevons dans cette partie les différents types de constructions constituant notre corpus de médecine.

Les noms composés de type NA (nom adjectif) sont nombreux, notamment dans la dénomination des maladies, en rapport avec les organes affectés comme *carie dentaire*, *atrésie intestinale*, *atrésie pulmonaire*, *inflammation cutanée*, etc. L'adjectif de relation peut évoquer l'origine de la maladie : *grippe aviaire*, *grippe porcine*, etc., ou toute autre motivation.

Ce type de formation de noms composés est très productif. La structure N de N (nom de nom) permet également la formation de noms composés en médecine pour la dénomination de maladies, d'exams, ou d'instruments, en rapport avec une partie du corps ou une substance, comme *abcès du foie*, *abcès du poumon*, *analyseur de graisse*, *bouffée de chaleur*, *kyste d'ovaires*, etc.

Pour les locutions nominales de type N de NA (nom de nom adjectif), on trouve par exemple des noms de locaux, (*laboratoire d'analyses médicales*), de maladies, (*asthme du cèdre rouge*), ou d'instruments médicaux (*analyseur de composition corporelle*).

Les noms composés de type NA de N (nom adjectif de nom) sont fréquents dans le vocabulaire de l'anatomie. Ils sont utilisés généralement dans la dénomination d'une partie du corps, tout en ayant une caractéristique spécifique comme *bord inférieur du foie*.

Il y a d'autres types de mots composés plus complexes qu'on emploie pour la dénomination de parties du corps ayant plusieurs caractéristiques illustrées par des adjectifs. Il s'agit de noms composés du type NAAA (nom adjectif adjectif adjectif) *bronche segmentaire*

apicale supérieure, du type N Prép-A de NAA (nom préposition-adjectif de nom adjectif adjectif) *bourse sous-tendineuse du muscle grand rond*, et du type NAA (nom adjectif adjectif) *espace perforé postérieur*.

1.2.3. La composition mixte

Certains noms composés du vocabulaire de la langue de la médecine que nous avons relevés dans notre dictionnaire sont constitués à la fois des mots savants et des mots de la langue française. On les appelle les mots composés mixtes.

Comme pour les noms composés précédents, nous présentons les types de formations de mots composés mixtes.

Les noms composés du type NA (nom + adjectif) sont les plus fréquents. Les substantifs *hypertension artérielle* et *hypotension pulmonaire* sont formés des préfixes grecs *hypo* (au dessous) et *hyper* (au dessus). Le nom *pneumopathie obstructive* est formé d'un mot composé de *pneumo* (poumon) et de *pathos* (maladie) et d'un adjectif non savant.

La structure du type NA de N (nom adjectif de nom) est celle des noms composés désignant des maladies ayant une certaine caractéristique et touchant un organe donné, *hypertrophie bénigne de la prostate*.

Nous signalons qu'il y a des noms composés du type NA Chiffre-Chiffre (nom adjectif chiffre-chiffre), *chromosome humain 13-15*, *chromosome humain 16-18*, et du type NA de NA (nom adjectif du nom adjectif), comme *apophyse stéroïde du troisième métacarpien*.

Il y a d'autres noms composés moins fréquents du type N de N (nom de nom) *manque d'insuline*, et du type N de NA *atteinte des nerfs abducens*.

1.3. Autres mots construits

Il y a d'autres modes de formation de mots dans la langue médicale. Il s'agit notamment de la siglaison, de l'acronymie et de l'ellipse.

1.3.1. Siglaison

La siglaison est considérée comme l'une des caractéristiques du vocabulaire scientifique et technique (J. Ghazi 1985). Ce phénomène concerne le vocabulaire médical car il est très utilisé dans les textes médicaux (J-C. Sournia 1997). Les praticiens de la santé utilisent les sigles vraisemblablement dans le but d'être précis et concis. Pour A. Martinet (1970) « *l'évolution linguistique peut être conçue comme régie par l'antinomie permanente entre les besoins communicatifs de l'homme et sa tendance à réduire au minimum son activité mentale et physique* ». L'observation de notre corpus de médecine montre clairement l'usage des sigles dans la langue de la médecine, qui sont autant de réductions des noms composés.

La siglaison concerne tout le vocabulaire de la médecine. Nous nous contentons de donner quelques exemples.

Sigles désignant des examens médicaux :

ECG électrocardiogramme

ERG électrorétinogramme

EMG électromyogramme

Sigles désignant des maladies :

RAA rhumatisme articulaire aigu

SEP sclérose en plaques

IVG insuffisance ventriculaire gauche

Sigles des accidents :

AVC accident vasculaire cérébral

AIT accident ischémique transitoire

Sigles des opérations :

IVG interruption volontaire de grossesse

AAC anastomose artérioveineuse chirurgicale

Nous constatons qu'il y a des sigles qui sont ambigus et qui prêtent à confusion. IVG est l'abréviation de deux noms composés différents, notamment *insuffisance ventriculaire gauche* et *interruption volontaire de grossesse*.

1.3.2. Acronymie

Un sigle peut devenir un acronyme. L'acronyme est la formation d'un terme à partir de la réduction d'un groupe de mots par l'abréviation d'une ou plusieurs des premières lettres des mots, à l'instar du phénomène de la siglaison présenté ci-dessus, mais le résultat est une suite de syllabes qui se prononce comme un mot simple (D. Zolondek 1991).

L'acronyme peut devenir une source de dérivation. Cela permet, à partir de sa base de créer d'autres mots appelés mots dérivés. J. Ghazi (1985) évoque leur rôle dans la participation à la création des néologismes lexicaux : *sida*, *sidaction*, *sidatique*.

Par exemple, le *syndrome d'immunodéficience acquise* donne le sigle *SIDA*, qui est à la base de la formation de *sidaïque* et *sidéen* pour parler d'une personne atteinte du *sida*. *Bacille Calmette-Guérin* est réduit par abréviation à *BCG* qui est à l'origine de *bécégite* qui est le résultat obtenu après la vaccination contre le *BCG*.

Certains sigles ne donnent pas des acronymes parce qu'ils ne sont pas lexicalisés, par exemple : *AAA* (*anévrisme de l'aorte abdominale*), *ACD* (*acidocétose diabétique*), *AAH* (*allocation aux adultes handicapés*), etc.

1.3.3. Troncation et mots valises

La troncation est considérée comme une abréviation d'un terme polysyllabique par suppression d'une ou plusieurs syllabes (Le Nouveau Petit Robert 2007). La langue de la médecine recourt très souvent à la troncation, comme l'attestent les forums sur la santé. Ce genre d'abréviation fait partie des mots de la communication quotidienne dans les milieux hospitaliers. L'abréviation permet aux praticiens d'être concis dans leurs échanges professionnels. On utilise *radio* pour *radiographie*, *néo* à la place de *néoplasie*, *valence-gramme* pour *val*. En général, on utilise la troncation dans la dénomination des spécialités de la médecine telles que *cardio* (cardiologie), *traumato* (traumatologie), *ophtalmo* (ophtalmologie), etc.

On utilise aussi les mots dits valises dans le vocabulaire médical. Nous avons tiré quelques exemples de nos corpus spécialisés et non spécialisés. Par exemple, pour les spécialités médicales on dit *anapathe* pour *anatomie-pathologie*, *chirdent* pour *chirurgie dentaire*, etc.

2. Propriétés sémantiques

La partie sur les propriétés sémantiques sera consacrée, entre autres, aux racines prédicatives du vocabulaire médical sous ses trois formes (nominale, verbale et adjectivale), aux types et aux classes sémantiques des prédicats, et enfin à la notion d'emploi prédicatif, en l'occurrence, les emplois nominaux, verbaux et adjectivaux.

2.1. Racines prédicatives

Tout contenu propositionnel peut se rapporter à une structure prédicat argument, le prédicat étant le constituant principal de la phrase (Z. Harris 1976, M. Gross 1981 & X. Blanco & P. A. Buvet 2009). La notion de prédicat concerne toutes les parties du discours, et pas uniquement le verbe malgré la place importante que la grammaire traditionnelle lui a attribuée dans la phrase (G. Gross & M. Mathieu-Colas 2001).

La notion de racine prédicative concerne les emplois prédicatifs qui partagent la même racine et le même domaine d'arguments (I. Sfar 2007).

À partir de la même racine prédicative, on peut avoir jusqu'à trois prédicats différents du point de vue de la catégorie grammaticale. Par exemple, de la racine prédicative *a(i)m*, on peut obtenir trois réalisations morphologiques différentes :

Paul aime Léa

Paul est amoureux de Léa

Paul éprouve de l'amour pour Léa

Les langues de spécialité ne sont pas des nomenclatures ayant une syntaxe particulière (P. Lerat 1995 & G. Gross 2002). La langue de la médecine est analysée avec les mêmes outils que ceux utilisés pour la langue générale.

Il y a trois types de racines prédicatives différents dans le vocabulaire médical, en l'occurrence, les racines relatives à un seul emploi prédicatif, qu'il soit verbal, nominal ou adjectival, les racines qui ont deux emplois prédicatifs relatifs aux formes verbales et adjectivales, et enfin les racines qui ont trois emplois prédicatifs, c'est à dire l'emploi verbal, l'emploi nominal et l'emploi adjectival.

2.1.1. Prédicats à forme unique

On entend par prédicat à forme unique, un emploi prédicatif qui résulte d'une formation morphologique unique à partir d'une racine prédicative quelconque, comme le prédicat nominal *crise*, qui possède une seule forme.

Cette catégorie de prédicats concerne, généralement dans le vocabulaire médical, la catégorie des noms. Il s'agit de prédicats comme *infarctus*, *crise*, *variole*, *ablation*, etc.

Nous constatons que les prédicats à forme unique ne sont pas fréquents dans notre dictionnaire.

2.1.2. Prédicats à forme nominale et adjectivale

Il n'y a pas de règles homogènes qui régissent la formation de noms et d'adjectifs à partir d'une même racine. La formation de certains des noms de spécialité en médecine montre des variations d'un nom à un autre, par exemple :

Ophthalmologie, ophthalmologue, ophthalmologiste

Radiologie, radiologue, radiologiste

*Cardiologie, cardiologue, *cardiologiste*

*Cancérologie, oncologue, *oncologiste*

En médecine, on a tendance à préférer les noms qui se terminent par *-logue* au détriment de ceux qui se terminent par *-logiste*. Ainsi le remplacement du substantif *ophtalmologue* par celui d'*ophtalmologiste* n'a aucune justification linguistique (G. Petit 2008).

Les adjectifs du vocabulaire médical peuvent se construire différemment, à l'aide des suffixe *-tique*, (*pharmaceutique, épidémique, hépatique*), *-aire* (*dentaire, articulaire*), *-al* (*nasal, ventral*), etc.

Généralement, chaque racine sélectionne quelques suffixes, d'où le blocage de toutes les autres substitutions par d'autres suffixes :

Dent : *dentaire, dental, *dentique*

Pharmacie : *pharmaceutique, *pharmaceutaire*

Sida : *sidaïque, *sidal*

Certains adjectifs sont formés à partir des bases savantes. Dans ce cas, il n'y a presque aucune ressemblance entre les substantifs et les adjectifs relatifs aux mêmes parties du corps, par exemple :

Gastrique est relatif à *estomac*,

Cutané est relatif à *peau*

Ophthalmique est relatif à *œil*

2.1.3. Prédicats à forme nominale et verbale

Dans le vocabulaire médical, les noms et les verbes partagent souvent les mêmes racines prédicatives. Beaucoup de racines prédicatives permettent de former des verbes par ajout d'affixes à des noms (désinences verbales).

Étant donné qu'il est l'une des classes grammaticales qui joue un rôle majeur dans la construction de la phrase, le verbe met en relation les autres éléments de la phrase lorsqu'il correspond à un prédicat. Son fonctionnement dans la phrase est exactement identique dans la langue générale et dans la langue spécialisée.

Pronostic, pronostiquer

Opération, opérer

Examen, examiner

Soins, soigner

Suture, suturer

Ce phénomène linguistique est extrêmement fréquent dans la langue de la médecine, d'après les données de notre corpus.

2.1.4. Prédicats à forme verbale, nominale et adjectivale

La polymorphie prédicative concerne ici trois parties de discours fonctionnant comme des prédicats dans les phrases. Il s'agit de la forme verbale, de la forme nominale et de la forme adjectivale, construites à partir d'une racine unique. Ainsi la racine prédicative *anesthési-* donne lieu à trois emplois prédicatifs : l'emploi nominal (*anesthésie*), l'emploi verbal (*anesthésier*) et l'emploi adjectival (*anesthésiant*). Il est important de noter que la morphologie joue un rôle majeur dans la sémantique de ces variantes. Le substantif est désigné par le formant *-e*, le verbe, par le formant *-er*, et l'adjectif par le formant *-ant* qui signifie l'effet de l'anesthésie.

Dans le tableau ci-dessous, nous donnons de nombreux exemples qui illustrent la variation des formes des racines prédicatives.

<i>Emploi nominal</i>	<i>Emploi verbal</i>	<i>Emploi adjectival</i>
<i>nécrose</i>	<i>nécroser</i>	<i>nécrosique</i>
<i>cancer</i>	<i>cancériser</i>	<i>cancéreux</i>
<i>anesthésie</i>	<i>anesthésier</i>	<i>anesthésiant</i>
<i>traumatisme</i>	<i>traumatiser</i>	<i>traumatisant</i>
<i>blessure</i>	<i>blessier</i>	<i>blessant</i>
<i>intubation</i>	<i>intuber</i>	<i>intubé</i>
<i>soins</i>	<i>soigner</i>	<i>soignant</i>

En guise de conclusion, nous dirons que la polymorphie prédicative de type nom, verbe et adjectif est un phénomène qui concerne de nombreux prédicats. Ce mécanisme dépend du génie de la langue française et de son mode de fonctionnement. Le résultat de cette variation morphologique a un impact direct sur de nombreux prédicats qui concernent notamment les emplois nominaux et les emplois adjectivaux. Les exemples cités ci-dessus attestent clairement cette variation morphologique. Ils sont construits par dérivation, mais leurs suffixes diffèrent d'un nom à un autre et d'un adjectif à un autre. Ainsi, la racine prédicative *anesthési* porte un *-e*, la racine prédicative *thraumat* porte *-isme*, la racine prédicative *bless* porte *-ure*, et la racine prédicative *intub* porte *-ation*, etc. nous avons fait le même constat chez les adjectifs. Ils sont variables du point de vue morphologique. Certaines racines prédicatives se terminent par *-ant* comme *bless*, *anesthési* et *traumatis*, etc. D'autres racines prédicatives se terminent par *ique* comme *névros*, ou par *-eux* comme *cancer*.

2.2. Types et classes sémantiques

Les types sémantiques caractérisent les emplois prédicatifs des prédicats. On dénombre dans les prédicats trois types : états, actions et événements. Cette classification a un intérêt particulier du point de vue sémantique et syntaxique (P. A. Buvet 2009). Le même prédicat a une syntaxe différente selon qu'il est du type *état*: *Paul a la grippe porcine*, ou *événement* : *Il y a eu la grippe porcine au Mexique*.

Du point de vue syntaxique, les emplois prédicatifs du type *événement* ont comme verbes supports *avoir lieu*, *se produire*, etc., tandis que les prédicats du type *état* sélectionnent des verbes supports différents tels que *avoir* (*avoir une tuberculose*), *attraper* (*attraper une syphilis congénitale*), *contracter* (*contracter une hépatite*), *développer* (*développer un cancer*).

Du point de vue sémantique, le même prédicat *grippe* a deux emplois différents selon qu'il est du type *état* (*Paul a la grippe porcine*), dans ce cas, on parle de l'état pathologique de Paul, ou du type *événement* (*Il y a eu la grippe porcine au Mexique*), dans lequel on parle du phénomène de la *grippe porcine* qui est apparue au Mexique.

La constitution d'une classe de prédicats s'effectue sur la base de la synonymie ou de l'hyponymie (P. A. Buvet & A. Grezka 2007). Il est tenu compte également de leurs domaines d'arguments pour traiter l'ambiguïté des prédicats.

2.2.1. Type état/action/événement

Les emplois prédicatifs du domaine de la médecine sont des emplois verbaux, nominaux ou adjectivaux. Ils sont du type état, événement ou action.

Dans le dictionnaire des emplois prédicatifs, chaque entrée est associée à un type sémantique.

Le type *état* caractérise les noms désignant les états pathologiques comme *dépression*, *anémie*, *anorexie*, *fièvre*, etc. Dans le dictionnaire, nous avons comptabilisé 8379 emplois de type état.

Le type *événement* correspond plutôt aux verbes événementiels tels que *se blesser, se rétablir, se remettre, contaminer*, aux inflammations comme *otite, gingivite, conjonctivite*, aux maladies contagieuses *grippe aviaire*, ainsi qu'aux épidémies qui ont forcément une lecture événementielle. Dans le dictionnaire, nous avons comptabilisé 537 emplois de type événement.

Le type *action* caractérise tous les verbes et les noms relatifs à des examens médicaux, des soins comme *panser* et *soigner*, etc., des prescriptions comme *prescrire, donner, interdire*, etc. et les opérations chirurgicales comme *opérer, greffer, opération, intervention*, etc. Dans le dictionnaire, nous avons comptabilisé 2862 emplois de type action.

Une même forme prédicative peut donner lieu à des emplois différents pour ce qui est du type *état*.

Paul a la grippe porcine (état)

Il y a eu plusieurs cas de grippe porcine en France (événement)

La syntaxe du nom prédicatif change selon qu'il est compatible avec le type *état* ou le type *événement*.

Les emplois prédicatifs du type *état* se combinent avec des verbes supports comme *avoir, attraper et contracter*. Ceux du type *événement* se combinent, par exemple, avec le verbe *avoir lieu*.

Dans les exemples ci-dessous, les emplois prédicatifs définis par la même racine n'ont pas le même type, selon qu'il s'agit du verbe ou du nom.

Un élève malade a contaminé ses camarades de classe (action)

La contamination se fait par des aliments souillés (événement)

Nous remarquons que le prédicat de type action est exprimé par un verbe, et le prédicat de type événement est exprimé par un substantif.

2.2.2. Classes sémantiques

L'attribution d'une classe sémantique aux prédicats verbaux s'effectue sur la base de la synonymie. Elle prend en compte le domaine d'arguments qui joue un rôle prépondérant dans la réduction de l'ambiguïté des verbes. Par exemple, le verbe *traiter* s'interprète différemment selon qu'il s'applique à un « patient » (*traiter un patient* équivaut à *soigner un patient*), à une « affaire » (*traiter une affaire* équivaut à *régler une affaire, négocier une affaire*), à un « individu » (*traiter quelqu'un* équivaut à *se conduire envers quelqu'un d'une manière ou d'une autre*), ou à un « sujet » (*traiter un sujet* équivaut à *étudier un sujet*) (G. Gross & M. Mathieu-Colas 2001).

La classe sémantique est définie à partir d'un prédicat exemplaire. Il s'agit du prédicat le plus représentatif de la classe, les autres prédicats sont souvent considérés comme périphériques (P. A. Buvet & A. Grezka 2007). Le prédicat représentatif de la classe est présenté sous forme de racine prédicative. Par exemple, le prédicat représentatif de la classe <traumatisme> est représenté par sa racine prédicative *traumatis-* parce qu'elle peut engendrer trois emplois prédicatifs appartenant soit à la catégorie des verbes (*traumatiser*), soit à la catégorie des noms (*traumatisme*) ou à la catégorie des adjectifs (*traumatisant*), tout en sachant que dans ces trois cas, la racine *traumatis-* donne lieu à un prédicat. Les prédicats autonomes sont présentés sous une seule et unique forme morphologique (verbale, nominale ou adjectivale).

Lorsque les prédicats n'ont pas de formes nominales, la classe est représentée par l'anaphore infidèle (G. Kleiber 2001). Par exemple, la classe (soins) est constituée de prédicats comme *examiner, traiter, consulter, suivre*. À travers ces exemples de la même classe, on peut construire des phrases permettant de reprendre la classe des <soins>. Par exemple : *Le médecin (examine + traite + consulte + suit) le patient.*

Les classes sémantiques qui caractérisent le dictionnaire des prédicats sont les suivantes :

<i>Classe sémantique</i>	<i>Nombre</i>	<i>Exemple</i>
<i>ACCIDENT</i>	<i>10</i>	<i>(se blesser)</i>
<i>ACCOUCHEMENT</i>	<i>04</i>	<i>(accouchement prématuré)</i>
<i>ADDICTION</i>	<i>24</i>	<i>(abus de cannabis)</i>
<i>ADJECTIF_MALADIE</i>	<i>274</i>	<i>(anémique)</i>
<i>ANESTHÉSIE</i>	<i>13</i>	<i>(anesthésie locale)</i>
<i>BRÛLURE</i>	<i>08</i>	<i>(brûlure de premier degré)</i>
<i>CICATRICE</i>	<i>08</i>	<i>(cicatrice de la conjonctive)</i>
<i>CONSULTATION</i>	<i>30</i>	<i>(consultation)</i>
<i>CRISE</i>	<i>35</i>	<i>(crise épileptique)</i>
<i>DIAGNOSTIC</i>	<i>25</i>	<i>(diagnostic par imagerie)</i>
<i>ÉPIDÉMIE</i>	<i>02</i>	<i>(épidémie)</i>
<i>EXAMEN_MÉDICAL</i>	<i>1168</i>	<i>(échographie)</i>
<i>FRACTURE</i>	<i>34</i>	<i>(fracture du crâne)</i>
<i>GROSSESSE</i>	<i>12</i>	<i>(grossesse compliquée)</i>
<i>INFECTION</i>	<i>101</i>	<i>(infecter)</i>
<i>INFLAMMATION</i>	<i>792</i>	<i>(proctites ulcéreuses)</i>
<i>INTOXICATION</i>	<i>28</i>	<i>(intoxication par hallucinogène)</i>
<i>LÉSION</i>	<i>38</i>	<i>(infarctus de myocarde)</i>

<i>MALADIE</i>	4961	<i>(hypertrophie des amygdales)</i>
<i>MALADIE_CHRONIQUE</i>	315	<i>(oncocytome rénale)</i>
<i>MALADIE_CONTAGIEUSE</i>	318	<i>(typhus)</i>
<i>MALNUTRITION</i>	03	<i>(malnutrition du fœtus)</i>
<i>OPÉRATION_CHIRURGICALE</i>	1114	<i>(transplantation rénale)</i>
<i>PARALYSIE</i>	55	<i>(paralysie faciale)</i>
<i>PRESCRIPTION</i>	06	<i>(prescrire)</i>
<i>SOINS</i>	93	<i>(panse)</i>
<i>STÉRILISATION</i>	04	<i>(stérilisation)</i>
<i>SYMPTÔME</i>	624	<i>(migraine commune)</i>
<i>SYNDROME</i>	1182	<i>(syndrome de Rendu)</i>
<i>TRAITEMENT</i>	04	<i>(traitement médical)</i>
<i>TRANSFERT</i>	01	<i>(transférer)</i>
<i>TRAUMATISME</i>	10	<i>(traumatisme acoustique)</i>

2.3.3. Aspect inhérent

D'un côté, il y a l'aspect textuel, c'est-à-dire que c'est dans le contexte qu'on trouve les repères temporels des mots. L'aspect indique la façon dont le procès se déroule du point de vue de son temps interne. Il exprime le début du procès (inchoatif: *il commence à travailler*), le déroulement du procès (progressif : *il est en train de travailler*), l'évolution du procès (continuatif : *il continue à travailler*), ou l'achèvement du procès (cessatif : *il fini son travail*) (A. Grezka & P. A. Buvet 2007).

D'un autre côté, il y a l'aspect inhérent à chaque terme. P. A. Buvet (2007) a dénombré cinq types d'aspect inhérent. Ils contribuent à une meilleure description des

emplois prédicatifs. Ils caractérisent différemment les types de prédicats (état, action et événement).

Les emplois prédicatifs du type *action* et *événement* se caractérisent par :

L'aspect intemporel (*le sida est une maladie transmissible par le sang*)

L'aspect duratif imperfectif (*ce chirurgien exerce à Val de Grâce*)

L'aspect duratif perfectif (*ce chirurgien opère le patient*)

L'aspect ponctuel (*Paul a fait un accident vasculaire cérébral*)

Les emplois prédicatifs du type *état* se caractérisent par :

L'aspect permanent (*Paul a le diabète*)

L'aspect provisoire (*Paul a la grippe*)

2.3. Les emplois prédicatifs

Nous analysons ici d'une façon détaillée des emplois nominaux, des emplois verbaux et des emplois adjectivaux.

2.3.1. Les emplois nominaux

Les substantifs du vocabulaire médical ne sont pas homogènes du point de vue syntaxique et sémantique. Certains peuvent fonctionner, soit comme des prédicats, soit comme des arguments. Cette distinction repose sur la place qu'ils occupent dans la phrase. D'autres sont toujours des arguments, comme les noms d'instruments médicaux (*scalpel*), de parties du corps (*main*), de micro-organismes (*virus*), etc. qui sont des noms élémentaires, cf. chapitre 5.

Les noms comme *cancer* <maladies>, *trabéculéctomie* <opérations chirurgicales>, (*panser*) <soins> figurent tantôt en position de prédicat, tantôt en position d'argument.

Pour ne pas confondre substantifs prédicatifs et argumentaux, il nous semble important de préciser que les noms de maladies fonctionnent en tant que prédicats uniquement dans les constructions en *avoir* (P. A. Buvet 1998).

Paul a une leucémie

Paul a un abcès sur la cuisse

Les mêmes noms de maladies sont des arguments lorsqu'ils se combinent avec des verbes comme *guérir, soigner, diagnostiquer*, etc.

Par exemple :

Le médecin diagnostique une bronchite chez Paul

Paul traite son abcès

Les noms prédicatifs de notre corpus sont l'ensemble des substantifs du vocabulaire médical pouvant fonctionner en tant que prédicats. Cette catégorie regroupe les noms de maladies ainsi que tous les termes qui désignent les souffrances physiques et psychiques que ressentent les humains, et tous les noms désignant les actes médicaux qui visent à les soulager.

Dans notre démarche, nous ferons un classement ontologique de la classe maladie et nous présenterons la syntaxe de chaque sous classe, c'est-à-dire les critères définitoires qui en justifient la création.

a. La classe maladie

Les noms de maladie se combinent avec des verbes et des adjectifs appropriés. Nous commençons d'abord par la catégorie des verbes compatibles avec les noms de maladies. Par exemple :

Paul a la gale

Paul guérit son infection urinaire

Le médecin diagnostique un cancer chez Paul

Paul soigner son carcinome hépatocellulaire

Nous constatons que les verbes *avoir*, *guérir*, *diagnostiquer* et *soigner* se combinent avec tous les noms de maladie, qu'elles soient contagieuses, chroniques, infectieuses ou inflammatoires.

Néanmoins, les maladies contagieuses, chroniques, infectieuses et inflammatoires se combinent avec les verbes appropriés à chaque classe, selon ses caractéristiques.

En ce qui concerne les adjectifs, nous préférons les étudier au cas par cas, parce que chaque sous-classe de maladie se combine avec des adjectifs spécifiques qui répondent à leurs caractéristiques.

b. Les sous-classes maladies

Nous avons présenté dans les exemples ci-dessus une série de verbes susceptible de fonctionner avec les maladies en position complément X1 et les humains en position sujet X0. Mais nous constatons que tous les verbes du vocabulaire de la médecine ne sont pas applicables à toutes les maladies.

b. 1. Maladies diverses

Les noms des maladies diverses regroupent les maladies passagères non contagieuses. Ils se combinent avec les verbes *avoir*, *guérir*, *soigner*, *traiter*, *diagnostiquer*, etc.

Paul a la migraine

Paul traite son aphte

Le médecin diagnostique une jaunisse chez Paul

Paul soigne sa jaunisse

Paul a guéri de sa jaunisse

Les noms de maladies diverses ne se combinent pas avec les verbes qui désignent certaines caractéristiques spécifiques des maladies chroniques et des maladies contagieuses.

b. 2. Maladies contagieuses

Les maladies contagieuses se combinent avec les verbes qui désignent la transmission de ces maladies de sujets malades à des sujets sains.

*Paul a attrapé (*le diabète + *le cancer + la grippe + l'hépatite + le sida)*

*Paul a contracté (*le diabète + *le cancer + la grippe + l'hépatite + le sida)*

*Paul a chopé (*le diabète + *le cancer + la grippe + l'hépatite + le sida)*

*(*le diabète + *le cancer + la grippe + l'hépatite + le sida) a contaminé beaucoup de personnes dans le monde.*

*(*le diabète + *le cancer + la grippe + l'hépatite + le sida) se transmet d'un sujet malade à un sujet sain.*

Néanmoins, cette sous-classe admet les verbes qui s'appliquent aux maladies au sens large (*avoir, souffrir de, soigner, guérir, diagnostiquer, etc.*).

b. 3. Maladies chroniques

La sous-classe *maladies chroniques* se distingue par une syntaxe particulière. Elle dénote une liste de verbes et d'adjectifs qui la différencient des autres sous-classes de la médecine.

La tétanie chronique hypophysaire + la lèpre (dure longtemps + se développe progressivement + s'installe définitivement + évolue lentement)

La sclérose en plaque (dure longtemps + se développe progressivement + s'installe définitivement + évolue lentement)

La polyarthrite chronique juvénile (dure longtemps + se développe progressivement + s'installe définitivement + évolue lentement)

L'arthrose (dure longtemps + se développe progressivement + s'installe définitivement + évolue lentement)

Le diabète (dure longtemps + se développe progressivement + s'installe définitivement + évolue lentement)

Cette combinaison propre aux noms de maladies chroniques s'explique par leurs spécificités, en l'occurrence, la durée dans le temps et l'évolution.

Les noms de la sous-classe *maladies chroniques* peuvent avoir une combinatoire commune avec les noms de maladies passagères et non transmissibles et les noms des maladies contagieuses

Paul (a + traite + soigne) (sa + la) lèpre

Paul (a + traite + soigne) (sa + la) grippe aviaire

Paul (a + traite + soigne) (son + l) ictère

Il reste quelques particularités liées aux maladies chroniques non guérissables. On ne peut pas les combiner avec le verbe *guérir*. Par exemple :

*Le médecin diagnostique, + traite +*guérit (La tétanie chronique hypophysaire + la sclérose en plaques + La polyarthrite chronique juvénile + L'arthrose + Le diabète)*

D'autres noms de maladies partagent à la fois les caractéristiques des noms de maladies chroniques et des noms de maladies contagieuses à l'instar de *sida* et *d'hépatite chronique*.

Paul a (attrapé + contracté) le sida

Paul a (attrapé + contracté) une hépatite chronique

Le sida + l'hépatite chronique sont des maladies qui (durent longtemps + se développent progressivement + s'installent définitivement + évoluent lentement)

Paul a (attrapé + contracté + chopé) le sida

Paul a (attrapé + contracté + chopé) une hépatite chronique

b. 4. Inflammation

Dans la langue de la médecine, les noms d'inflammations n'ont pas de syntaxe spécifique. Les substantifs ont les mêmes propriétés que les noms présentés ci-dessus. Néanmoins, il y a des différences selon les types d'inflammations.

Les noms d'inflammations d'origine virale et bactérienne désignent les inflammations contagieuses, par exemple *les oreillons*.

L'enfant a (attrapé + contracté + chopé) les oreillons

Les noms d'inflammations chroniques fonctionnent syntaxiquement de la même façon que les maladies chroniques.

La myringite chronique (dure longtemps + se développe progressivement + s'installe définitivement + évolue lentement)

La myosite orbitaire chronique (dure longtemps + se développe progressivement + s'installe définitivement + évolue lentement)

b. 5. Symptôme

Le symptôme est une manifestation spontanée d'une maladie qui permet de la découvrir et de la dévoiler. Il est intimement lié à la maladie elle-même. Ils ont une combinatoire spécifique. Les noms de symptômes évoquent leurs spécificités et les caractéristiques de la maladie à laquelle ils sont liés.

Les douleurs thoraciques sont un symptôme tardif

Les douleurs thoraciques sont récurrentes

La fatigue est un symptôme tardif

La fatigue persiste

La toux est un symptôme précoce

Paul a une toux sèche

Les noms de symptômes partagent quelques spécificités avec les noms de maladie, comme l'illustrent les exemples ci-dessous :

L'atrophie se manifeste par la perte du volume des organes

L'atrophie est un symptôme inquiétant

L'atrophie est un symptôme alarmant

L'atrophie est un symptôme

c. La classe attaques

Les noms d'attaques sont définis par plusieurs verbes. Par exemple les noms de *crise*, *de traumatisme* et de *paralysie* sont différents du point de vue syntaxique. Même s'il est parfois difficile de définir ces noms avec exactitude, on observe des différences entre les noms.

Les exemples suivants montrent les particularités syntaxiques des noms d'attaques, entre autres, leur combinaison avec des verbes appropriés à cette classe.

Il fait une crise d'asthme

Les crises d'asthme se manifestent chez Léa

Les crises d'asthme chronique se répètent

Les crises d'asthme l'étouffent

Les crises d'asthme l'empêchent de respirer

La paralysie l'a cloué à son lit

<crise>, <traumatisme> et <paralysie> partagent certains verbes avec les maladies comme *avoir*, *soigner*, *etc.* par exemple :

Paul a des crises d'asthme

Paul soigne ses crises d'asthme

Paul a guéri de ses crises d'asthme

d. La classe accidents

Les noms d'accidents que nous avons relevés dans notre dictionnaire sont les noms de *fractures* et les noms de *brûlures*. Ils peuvent avoir une combinatoire commune avec tous les noms de maladie, par exemple :

Paul (a + est victime de + soigne) une fracture de l'os temporal

Paul (a + est victime de + soigne) une brûlure cutanée

Néanmoins, ils ont une combinatoire qui leur est propre et qui les distingue des autres noms de maladies. Par exemple, les noms de brûlures se combinent avec le verbe *exciser* et les noms de fractures se combinent avec le verbe *plâtrer*.

Le médecin excise les brûlures cutanées profondes

Le médecin plâtre la fracture

e. La classe soins

Les soins, en médecine, regroupe tout acte médical pratiqué dans le but de prévenir et ou de guérir une maladie quelconque. Il s'agit de noms comme *stérilisation, diagnostic, opération chirurgicale, consultation, traitement*, etc.

Cette classe est définie par les verbes supports tels que *prodiguer, appliquer, pratiquer, donner*, etc.

Appliquer un pansement

Pratiquer une ablation

Donner un traitement

Prodiguer des soins

Etc.

Certains verbes comme *stérilisation* et *curetage* fonctionnent comme des prédicats dans la construction avec des noms en position X1, par exemple.

Stérilisation de la plaie

Cureter la carie dentaire

Panser une partie du corps

Les noms de soins ont une combinatoire particulière, ils ne se combinent pas avec tous les verbes supports de façon aléatoire. Par exemple le verbe *donner* se combine avec *traitement* et jamais avec *examen médical*.

Le médecin donne un traitement à Paul

**Le médecin donne un examen médical à Paul*

Le nom de soins *examen médical* se combine avec le verbe *prescrire*.

2.3.2. Les emplois verbaux

La langue de la médecine contient des verbes hautement spécialisés comme *cautériser* et *disséquer*. Ces verbes ne posent pas de problèmes de polysémies dans le vocabulaire médical. Les verbes généraux comme *suivre* et *donner* sont polysémiques, leurs sens sont acquis selon le contexte et le domaine d'arguments. Le verbe générique *suivre* prend des sens différents selon ses arguments. Il signifie *soigner* dans *le médecin suit un patient*, *marcher derrière* dans *l'enfant suit sa mère*, et *s'inspirer* dans *l'étudiant suit la méthodologie de son professeur*. Les arguments sont l'outil indispensable pour réduire la polysémie des verbes. Le verbe *donner* signifie *offrir* dans *il donne un cadeau à sa mère* et *prescrire* dans :

Le médecin donne un traitement au patient.

Les verbes de la langue générale sont très employés dans le vocabulaire médical comme *traiter*, *consulter*, *admettre*, etc. L'étude de leurs domaines d'arguments est indispensable pour préciser leurs sens afin d'éviter les ambiguïtés.

Nous examinons ces trois catégories de verbes.

a. Verbes spécialisés

L'étude des prédicats verbaux nécessite la description des arguments en position sujet X0 et en position complément X1

Le médecin hospitalise le patient (X0 = <hum>, X1 = <hum>)

Le médecin stérilise la plaie (X0 = <hum>, X1 = <plaie>)

La plaie s'est infectée (X0 = <hum>)

Malgré le degré de spécialisation des verbes ci-dessus, la classe d'arguments est importante pour la désambiguïsation, notamment dans le cas du verbe *stériliser* qui appartient au répertoire du vocabulaire médical mais qui peut s'appliquer à plusieurs arguments différents.

Le médecin stérilise la plaie (blessure)

Le médecin stérilise le scalpel (instrument)

Le médecin stérilise le bloc opératoire (locatif)

Dans le premier exemple, il signifie *soigner* lorsqu'il se combine avec *plaie*. Dans le deuxième et le troisième exemple, il signifie *nettoyer* parce qu'il se combine avec un instrument et un locatif.

b. Verbes de la langue générale

Les verbes de la langue générale sont très polysémiques ; leur sens varie selon les arguments auxquels ils s'appliquent.

Il suit une thérapie

Il suit les bons conseils

Il suit les cours à la faculté

En observant le verbe *suivre*, on constate qu'il a trois emplois différents, son premier emploi est *soigner*, le deuxième est *accepter les conseils* et le troisième est *étudier*. Ces exemples à partir du verbe *suivre* paraissent simples mais ils montrent les variations des emplois qui suivent les changements d'arguments dans la phrase.

D'autres exemples ci-dessous montrent comment l'emploi du verbe *prendre* change. Il s'agit de la polysémie prédicative justifiée par le changement de la classe d'arguments.

Prendre du paracétamol
 = *ingurgiter du paracétamol*
Prendre le pouls
 = *Mesurer la tension*

Le verbe *prendre* signifie *consommer* ou *ingurgiter* lorsqu'il se combine avec substantif *paracétamol*. Par contre, il prend le sens de *mesurer* quand il s'applique au substantif *pouls*.

c. Verbes de la langue générale très utilisés en médecine

Les verbes de la langue générale sont très employés dans la langue de la médecine mais ils ont une valeur spécifique. Leur fréquence qui est assez élevée dans le discours médical leur donne une empreinte particulière, c'est-à-dire qu'ils appartiennent à la langue générale tout en donnant l'impression aux usagers du discours médical qu'ils font partie des termes spécialisés, comme *admettre*, *consulter* et *traiter*.

Paul est admis à l'hôpital
Le médecin consulte le patient
Le médecin traite le patient

Néanmoins, le changement de contexte, qui génère effectivement le changement d'emploi de ces verbes, montre avec évidence qu'ils font partie du répertoire de la langue générale.

Paul admet ce comportement
Paul consulte sa montre
Paul traite son ami de con

Nous constatons que les verbes signifiaient auparavant *hospitaliser, ausculter et soigner*, et qu'ils signifient dans les exemples ci-dessus *accepter, regarder et insulter*.

2.3.3. Les emplois adjectivaux

Les emplois adjectivaux de la classe *maladie* concernent tous les adjectifs qui désignent le degré et la gravité de la maladie (*curable, incurable, bénigne, maligne, etc.*), son origine (*sérique, bactérienne, virale, etc.*), sa nature (*maladie chronique, contagieuse, inflammatoire, transmissible, etc.*), la catégorie de personne atteinte (*infantile*), l'organe touché (*osseuse, cervicale, cardiaque, pulmonaire, etc.*), etc.

L'adjectif se construit avec l'auxiliaire *être* qui l'actualise. Son schéma d'arguments se réduit à un seul argument (X0) (M. Mathieu-Colas 1998).

Les adjectifs de maladies ne se combinent pas d'une façon aléatoire. La sélection est plus fine quand on classe les maladies par catégories.

Chaque classe de maladies sélectionne les adjectifs qui les caractérisent. Les maladies chroniques admettent par exemple tous les adjectifs qui évoquent la durée, les maladies contagieuses sélectionnent les adjectifs exprimant la contagion, etc.

L'hépatite A est chronique

Le diabète est chronique

Le sida est contagieux

L'hépatite C est contagieuse

Les cancers sont généralement mortels

La lèpre est invalidante

Nous constatons que chaque maladie est discriminée par un adjectif qui peut s'appliquer soit à une seule maladie, soit à toute la classe.

La leucémie est mortelle

**Le cancer de sein est mortel*

Dans le langage médical, les adjectifs sont souvent interprétables à partir de leur contexte. Ainsi l'emploi de l'adjectif *bon* obtient son sens dans le contexte.

*Cette pommade est **bonne** pour les muscles*

= *Cette pommade est **efficace** pour soulager les muscles*

*Cette pommade est **bonne** pour les douleurs musculaires*

= *Cette pommade est **contre** les douleurs musculaires*

Nous observons, dans les exemples ci-dessus, que l'adjectif *bonne* signifie *efficace* dans le premier cas, et qu'il signifie *contre* dans le deuxième cas.

3. Propriétés combinatoires

La syntaxe est importante dans le processus d'analyse linguistique puisqu'elle nous permet de définir tous les emplois prédicatifs possibles d'un prédicat. Autrement dit, il y a autant d'emplois et autant de sens du prédicat qu'il n'y a de constructions différentes.

3.1. Construction

Nous étudions ici les différents cas de constructions syntaxiques des verbes, des noms et des adjectifs prédicatifs. La construction sert à établir une structuration des prédicats selon leurs emplois prédicatifs, le nombre d'arguments auxquels ils s'appliquent et leurs positions dans les phrases.

3.1.1. Construction des verbes prédicatifs

Nous étudions ici toutes les constructions possibles des emplois prédicatifs des verbes de la médecine utilisés dans notre dictionnaire (PRED_DIC).

L'emploi prédicatif de *consulter* a la construction **X0 V X1** (*le médecin consulte le patient*), autrement dit, c'est un verbe prédicatif qui se combine obligatoirement avec un sujet (**X0**) et un complément d'objet direct (**X1**).

Les emplois prédicatifs de *diagnostiquer* ont les deux constructions suivantes : **X0 V X1** (*le médecin diagnostique une maladie*) et **X0 V X1 PREP X2** (*le médecin diagnostique une maladie chez le patient*). Le premier emploi se combine obligatoirement avec un sujet (**X0**) et un complément d'objet direct (**X1**). Le deuxième emploi se combine avec un sujet (**X0**), un complément d'objet direct (**X1**) et un complément d'objet indirect (**X2**).

Les emplois prédicatifs du verbe *se blesser* ont deux constructions. La première construction est **X0 V** (*Paul s'est blessé*), il prend obligatoirement (**X0**) en position de sujet et un verbe pronominal. La deuxième construction est **X0 V PREP X1** (*Paul s'est blessé à la main*), il prend (**X0**) en position de sujet qui se combine avec un verbe prédicatif et un (**X1**) en position de complément d'objet indirect.

3.1.2. Construction des adjectifs prédicatifs

La construction adjectivale est restreinte, sachant que les adjectifs sont toujours actualisés par le verbe *être*.

Les emplois *pâle*, *cireux*, *diagnostiqué*, et *peureux* correspondent à la construction **X0 être A** en tant qu'adjectifs prédicatifs qui n'acceptent qu'un sujet (**X0**).

Paul est pâle,

Le pied du patient est cireux,

Le cancer est diagnostiqué,

Paul est peureux

3.1.3. Construction des noms prédicatifs

Les emplois *diagnostic*, *handicap* et *douleur* ont la construction **X0 V sup DET N** en tant que noms prédicatifs qui n'acceptent qu'un sujet. Cependant, les deux substantifs ne sont compatibles ni avec les mêmes verbes supports, ni avec les mêmes déterminants :

Il s'ensuit que les deux emplois sont spécifiés avec une construction faisant état de leur verbe support le plus canonique : **X0 avoir DET N** pour *handicap*, **X0 avoir DET N** pour *diagnostic* et **X0 ressentir DET N** pour *douleur*.

Le second emploi de *diagnostic* a la construction **X0 faire DET N de X1** en tant que nom prédicatif qui accepte un sujet (**X0**) et obligatoirement un complément d'objet indirect (**de X1**).

Le premier emploi de *peur* a la construction **X0 avoir N (E+ (de) X1)**, le second emploi la construction **X0 avoir N (de) X1** ; autrement dit, si les deux prédicats nominaux sont caractérisés par la même actualisation (ils se combinent avec l'article zéro et le verbe *avoir*), ils diffèrent pour ce qui est du complément : il est facultatif pour le premier emploi, obligatoire pour le second emploi.

Luc a peur (E + de l'opération chirurgicale) (premier emploi)

*Luc a peur (*E+ que son médecin soit déjà parti) (deuxième emploi)*

3.2. Combinatoire interne

Dans la combinatoire interne, on évoque plutôt la distribution syntactico-sémantique et morphologique des noms composés, des collocations et des séquences dites figées. Nous analysons leurs constructions internes, et leur degré d'opacité et de transparence. Nous montrons également la nature des constituants du point de vue de la construction compositionnelle et de la construction figée afin d'établir des typologies spécifiques à la langue de la médecine.

3.2.1. Séquences opaques

Dans cette partie, nous étudions les expressions figées et leur degré d'opacité dans le vocabulaire médical.

Dans notre corpus, les séquences figées sont peu nombreuses, par conséquent, les langues de spécialité, notamment la langue de la médecine comporte très peu de séquences figées opaques. Cette rareté est justifiée par la volonté des spécialistes de la médecine d'être plus clairs et précis dans leurs pratiques langagières, notamment lorsqu'ils s'adressent aux patients qui ne sont pas initiés au vocabulaire médical.

Les exemples suivants illustrent la présence de certaines séquences figées dans la langue de la médecine :

Pomme d'Adam
Culotte de cheval
Poignée d'amour
Etc.

Certes, ces expressions sont figées, elles concernent le vocabulaire médical et on constate qu'elles désignent les parties du corps humain, mais si on les observe bien, on constate également qu'elles tendent beaucoup plus vers la langue générale que vers la langue de la médecine.

Du point de vue morphologique, ces expressions ont une forme *N de N*. Leur construction interne est identique dans les trois exemples.

La lecture compositionnelle de ces expressions montre qu'elles sont constituées soit de deux arguments, soit d'un argument et d'un prédicat.

Pomme (arg) d'Adam (arg)
Culotte (arg) de cheval (arg)
Poignée (arg) d'amour (préd)

Ces locutions nominales apparaissent comme des arguments dans des phrases puisqu'elles désignent les parties du corps. Ces expressions sont formées par un verbe et un nom.

3.2.2. Séquences transparentes

Nous nous intéressons ici à la combinatoire interne de deux types de séquences ; il s'agit des séquences figées et des collocations.

Le procédé de dénomination en langues de spécialité augmente le nombre de mots composés dans le vocabulaire médical. Le dictionnaire des prédicats comporte 9304 mots

composés sur 11300 entrées. Cela présente une proportion de 80% de mots composés dans le dictionnaire.

Il existe des expressions transparentes comme fièvre de cheval. Cette expression est suffisamment transparente parce que *fièvre de cheval* est une *fièvre*, et le mot *cheval* signifie *forte*. Autrement dit, une *fièvre de cheval* est une *forte fièvre*. A l'instar des expressions transparentes, les expressions opaques comme *culotte de cheval*, *tendon d'Achille*, etc., relèvent plutôt de la langue générale. Cela justifie davantage le fait que les langues spécialisées privilégient la transparence et la précision.

Les collocations sont très nombreuses dans le vocabulaire médical, elles servent à dénommer les noms de maladies touchant des organes et des parties du corps particuliers ou des opérations chirurgicales faites sur des parties du corps et des organes précis.

Les morphologies les plus récurrentes des collocations sont tu type *N de N*, (Nom De Nom) comme :

Cancer du sang

Tumeur du foie

Opération du cœur

Et *NA* (Nom Adjectif) comme :

Douleur abdominale

Épilepsie partielle

Fièvre hémorragique

Les noms comme *cancer*, *tumeur*, *opération*, *douleur*, *épilepsie* et *fièvre* (la liste est longue) servent de bases à la création de collocations, en les associant à des collocatifs, ces derniers concernent tous ce qui est relatif à la base de la collocation.

La lecture compositionnelle de ces expressions nous permet de constater que les éléments constituant les collocations sont soit des prédicats soit des arguments. Les séquences sont différentes de ce point de vue.

Cancer du sang (Pred/ Arg)

Tumeur du foie (Pred/ Arg)

Opération du cœur (Pred/ Arg)

Douleur abdominale (Pred/ Arg)

Épilepsie partielle (Pred/ Pred)

Fièvre hémorragique (Pred/ Pred)

Fracture ouverte (Pred/ Pred)

Nous constatons que les collocations, dans notre dictionnaire, sont formées soit d'un prédicat et d'un argument, soit de deux prédicats. Nous n'avons pas recensé d'expressions formées de deux arguments parce que le dictionnaire de prédicats comporte uniquement les prédicats. Parce contre nous pouvons relever cela dans le dictionnaire des arguments (*lame de bistouri*).

3.2.3. Détermination standard

Nous montrons dans cette partie les contraintes sur la détermination standard des prédicats du vocabulaire de la médecine. Nous nous contentons de l'appliquer aux prédicats de la langue de la médecine, tels que les noms de maladies, les noms de soins ainsi que les noms d'examens médicaux.

Avant de donner des exemples sur la détermination standard, nous définissons ce concept. Selon P. A. Buvet (2009), la notion de détermination standard relève essentiellement de celle de signification grammaticale. Dans le cas de la détermination standard, la signification grammaticale est faible : *Le médecin lui prescrit du paracétamol*. Par contre, dans la détermination non standard, la signification grammaticale est forte : *Le médecin lui prescrit (une + deux + trois + plusieurs) boîte (s) de paracétamol*.

Comme tous les noms du français, les noms de maladies nécessitent des déterminants, qui leur servent d'actualisateurs, pour constituer des groupes nominaux.

(1) *Paul a une grippe*

(2) *Paul a un cancer*

Nous constatons que les déterminants utilisés dans ces exemples de noms de maladies n'apportent aucune information supplémentaire sur les maladies respectives *grippe* et *cancer*. Dans ce cas, la signification des déterminants est peu marquée. Elle est grammaticalement faible, il s'agit de la détermination standard (P. A. Buvet 2009).

La détermination standard des prédicats nominaux médicaux est identique pour presque tous les substantifs médicaux. Néanmoins, certains noms n'admettent pas quelques déterminants comme ceux qui signifient le pluriel (*des, les, etc.*). Cette caractéristique est liée à la spécificité intrinsèque des substantifs.

(3) *Prendre le pouls*

(4) *Faire un examen visuel*

(5) *Paul a (le + *les) diabète (s)*

La détermination dans les exemples (3) et (4) relève de la détermination standard. Elle nous donne le minimum d'informations sur *pouls* et *examen visuel*. Dans le premier cas le nom *pouls* est déterminé par le pronom défini *le*, et dans le second, *examen visuel* est déterminé par le pronom indéfini *un*.

Dans l'exemple (5), le nom de maladie *diabète* n'admet pas le pluriel. Cela relève des spécificités extralinguistiques du substantif. C'est-à-dire, un humain ne peut pas avoir plusieurs diabètes à la fois.

3.3. Combinatoire externe

La combinatoire externe dépend du statut syntactico-sémantique du nom composé et de son emploi dans le cadre de la phrase, ainsi que des éléments linguistiques avec lesquels ils se combinent, notamment les verbes et les adjectifs.

3.3.1. Les verbes

Les noms composés ont un statut syntactico-sémantique variable lorsqu'ils se combinent avec les verbes. Autrement dit, ils peuvent occuper la position de prédicat ou

fonctionner comme des arguments, selon les constructions syntaxiques des énoncés dans lesquels ils apparaissent. Leur syntaxe est semblable à celle des noms simples.

Les noms composés ont une fonction argumentale lorsqu'ils se combinent avec des verbes.

Opérer la tumeur de la vessie

Traiter les troubles de l'éveil

Stériliser la plaie pénétrante

Tandis que les noms composés de la classe *maladie* et de la classe *symptôme* ont un emploi prédicatif lorsqu'ils se combinent avec le verbe avoir :

Il a une tumeur cérébrale

Il a une hémorragie postopératoire

Il a des douleurs abdominales

Par contre, ils fonctionnent comme des arguments quand ils se combinent avec certains verbes spécialisés ou généraux comme *diagnostiquer*, *soigner*, *ressentir*, etc.

Il a diagnostiqué une tumeur cérébrale

Il soigne son hépatite chronique

Il ressent des douleurs abdominales

3.2.2. Les adjectifs

L'emploi des noms composés est différent, du point de vue de la combinatoire externe, lorsqu'ils se combinent avec des adjectifs. Ils sont toujours en position argumentale, et les adjectifs fonctionnent comme des prédicats.

La lame de bistouri est rouillée

La salle de soins est nettoyée

Le cancer des os est paralysant

Les adjectifs associés aux noms composés désignant les instruments, les appareils, les outils, les parties du corps, etc. désignent les états ou les caractéristiques de ces derniers. Ils ont comme type sémantique *état*. Ils fonctionnent comme des prédicats.

Lorsque les noms composés qui désignent les maladies et les symptômes, se combinent avec les adjectifs, ils ont un emploi argumental.

Les douleurs abdominales sont récurrentes

La laryngite tuberculeuse est transmissible

L'hépatite est une maladie infectieuse

Pour conclure, nous constatons que la combinatoire des expressions est semblable à celle des noms simples. Ils sont soit des prédicats, soit des arguments dans la phrase, mais c'est leur position qui détermine cette variation d'emploi.

3.3.3. Détermination non standard

On parle de détermination non standard lorsque, du point de vue de la signification grammaticale, les déterminants sont porteurs d'une valeur marquée. Il s'agit d'une signification grammaticale forte (P. A. Buvet 2009).

(1) *Paul a une forte grippe*

(2) *Paul a un début du cancer*

Dans (1) et (2), la signification grammaticale des déterminants est dite forte car ils ne se contentent pas de stipuler la non définitude des noms de maladies *grippe* et *cancer*, mais impliquent des valeurs supplémentaires en terme d'intensité pour *grippe* et d'inchoatif pour *cancer*.

(3) *Paul ressent plusieurs douleurs*

L'actualisation du substantif *douleur* par *plusieurs*, participe également d'une signification grammaticale forte. Le remplacement de *des* par *plusieurs* prend en charge les particularités de *des*, notamment le pluriel et apporte la fréquence comme valeur supplémentaire.

Conclusion

Dans cette partie, nous avons présenté globalement le dictionnaire des prédicats ainsi que les classes sémantiques qui le constituent.

Cette étude s'est focalisée sur l'analyse des emplois prédicatifs de la langue de la médecine considérée comme une langue spécialisée. Nous avons montré que cette langue s'analyse selon les mêmes principes généraux que la langue générale, que ce soit sur le plan morphologique, sémantique et combinatoire.

Nous avons tout d'abord étudié la morphologie des emplois prédicatifs, sachant que certain affixes aident à la reconnaissance automatique. Par exemple, toutes les maladies qui se terminent par le suffixe *ite* signifient *inflammation*. Par la suite nous avons traité la sémantique selon le modèle des classes d'objets. Enfin nous avons analysé la combinatoire des emplois prédicatifs.

Dans le vocabulaire médical, il existe des mots qui fonctionnent comme des prédicats dont l'analyse est prioritaire afin de comprendre la syntaxe de la langue de la médecine, et il existe également la catégorie des noms élémentaires qui ne fonctionnent que comme des arguments. Leur analyse est fondamentale pour le traitement automatique puisque chaque classe a ses caractéristiques et ses critères définitoires, notamment les prédicats appropriés à chaque classe.

CHAPITRE 5

LES NOMS ARGUMENTAUX DU DOMAINE DE LA MÉDECINE

Introduction

Les substantifs étudiés dans cette partie sont uniquement des noms argumentaux. Ils ont comme particularité de ne jamais fonctionner comme des prédicats mais uniquement comme des arguments, contrairement aux noms prédicatifs qui fonctionnent soit comme prédicats, soit comme arguments.

Il y a un nombre important de noms argumentaux dans le vocabulaire médical (5581). On trouve entre autre les noms d'instruments et d'appareils comme *bistouri*, *scalpel*, *toise bébé*, etc., les noms des parties du corps : (*tête*, *main*, *pied*, *peau*, etc.), les noms de médicaments comme *paracétamol*, *extencilline*, *rifabutine*, etc., les noms des locatifs comme *hôpital*, *laboratoire d'analyse médicale*, *pharmacie*, etc., et les noms des parties des lieux comme *salle de soins*, *bloc opératoire*, *salle de réanimation*, etc.

Ce chapitre se divise en trois parties. La première partie sera consacrée à la morphologie des noms élémentaires. Il y a des noms simples comme *bistouri*, *hôpital*, des noms composés comme *toise bébé*, *bloc opératoire* et des noms dérivés comme *front* (frontal). Dans la deuxième partie, nous évoquerons les définitions ontologiques des noms élémentaires et nous analysons les classes sémantiques avec des codifications précises pour chacune d'entre elles. La troisième partie sera consacrée aux propriétés syntaxiques des noms argumentaux du point de vue de leur combinatoire avec des verbes, des adjectifs et des noms.

1. Propriétés morphologiques

Nous abordons la morphologie des termes médicaux qui fonctionnent uniquement comme des arguments selon qu'ils constituent des noms simples, des noms dérivés ou des noms composés.

1.1. Noms simples

Dans la catégorie des noms simples du dictionnaire des arguments, il y a tous les noms qui fonctionnent uniquement comme des arguments dans des phrases. Ils sont dotés du trait *inanimé concret* comme *scalpel, seringue, sirop, dakin, os, tibia, etc.*, ou du trait *animé* comme *virus, champignon, bactérie*.

Les noms simples du vocabulaire médical et du français en général peuvent varier selon le genre (féminin et masculin) et selon le nombre (singulier et pluriel). Dans notre dictionnaire, nous avons relevé les genres et les nombres suivants :

- le masculin singulier : *ventre, bistouri, scalpel, about, abdomen, etc.*
- le féminin singulier : *tête, dent, artère, cellule, etc.*
- le masculin pluriel : *ventres, bistouris, scalpels, abouts, abdomens, etc.*
- le féminin pluriel : *têtes, dents, artères, cellules, etc.*

Selon le genre et le nombre, il y a des noms qui ont une seule forme (*abcès*), des noms qui ont deux formes (*cancer, cancers*) et des noms qui ont quatre formes (*handicapé, handicapée, handicapés, handicapées*).

L'existence de certains mots qui sont à la fois masculin et féminin en médecine peut s'expliquer par leur origine gréco-latine. Il existe en latin et en grec un troisième genre dit neutre, tout en sachant que les mots invariables relevés dans le dictionnaire des arguments sont tous d'origine latine ou grecque.

Pour certains noms, seul le déterminant peut être en mesure de montrer la marque du genre, par exemple : *une dent*. L'absence de marque du féminin (e) dans *dent* ne signifie pas non plus l'appartenance de ce mot au genre masculin. Cela est valable également pour le nombre. Les noms comme *corps* et *virus* s'écrivent de la même manière au singulier et au pluriel. Le *s* de *virus* et *corps* ne signifie pas la marque du pluriel.

La variation observée est à la fois le fruit des règles grammaticales de la langue française qui permet quelques irrégularités telles que celles signalées ci-dessus, et l'héritage du vocabulaire de la médecine des langues anciennes que sont le grec et le latin, dans lesquelles on distingue clairement deux nombres et trois genre.

Après avoir vu les caractéristiques générales des noms simples, nous étudions les noms de la langue de la médecine, selon qu'ils sont propres ou non à la médecine.

1.1.1. Noms spécialisés

Certains termes de la langue de la médecine appartiennent exclusivement au domaine médical. Leur utilisation dans la phrase est univoque, ils ne se prêtent pas à des interprétations ambiguës :

<i>Entrée</i>	<i>Classe</i>	<i>Domaine</i>
<i>bistouri</i>	<i>INSTRUMENT_CHIRURGIE</i>	<i>chirurgie</i>
<i>scalpel</i>	<i>INSTRUMENT_CHIRURGIE</i>	<i>chirurgie</i>
<i>stéthoscope</i>	<i>INSTRUMENT_MESURE</i>	<i>médecine</i>
<i>liquide amniotique</i>	<i>LIQUIDE_ORGANIQUE</i>	<i>médecine</i>
<i>bactérie</i>	<i>MICRO_ORGANISME</i>	<i>médecine</i>
<i>hôpital</i>	<i>LOCATIF</i>	<i>médecine</i>
<i>camisole</i>	<i>VETEMENT</i>	<i>psychiatrie</i>

Le dictionnaire des arguments est constitué, en majorité, des noms élémentaires du domaine biomédical, notamment l'anatomie, la pharmacologie, la microbiologie, les lieux de soins, les instruments médicaux, etc.

Habituellement, l'utilisation des termes dans le cadre de la phrase peut donner des interprétations différentes, selon les termes avec lesquels ils se combinent. Néanmoins, les termes spécialisés ne changent pas de signification et leur syntaxe est limitée. Ils ne changent pas de sens parce qu'ils ne sont pas polysémiques, c'est-à-dire qu'ils portent un seul sens et qu'ils ne s'utilisent que dans le domaine spécialisé de la médecine. Leur syntaxe est limitée

parce que, par exemple sur l'axe syntagmatique, ils ne s'accordent pas avec beaucoup de verbes, mais uniquement avec des verbes appropriés:

Le liquide amniotique se renouvelle

**Le liquide amniotique se boit*

1.1.2. Noms non spécialisés

Les langues de spécialité ne sont pas constituées uniquement de termes spécialisés et techniques, elles comportent également des termes de la langue générale. Ainsi le dictionnaire des arguments ne comporte pas uniquement des termes appartenant exclusivement à la médecine, mais aussi les termes de la langue générale qui sont très employés dans le discours médical. Dans le tableau ci-dessous, nous donnons quelques exemples qui illustrent cela :

<i>Entrée</i>	<i>Classe</i>	<i>Domaine</i>
<i>main</i>	<i>PARTIE_CORPS</i>	<i>général</i>
<i>pied</i>	<i>PARTIE_CORPS</i>	<i>général</i>
<i>nez</i>	<i>PARTIE_CORPS</i>	<i>général</i>
<i>sang</i>	<i>LIQUIDE_ORGANIQUE</i>	<i>général</i>
<i>ciseau</i>	<i>OUTIL</i>	<i>général</i>
<i>cheveu</i>	<i>PARTIE_CORPS</i>	<i>général</i>
<i>dose</i>	<i>QUANTITE</i>	<i>général</i>

Les noms non spécialisés employés dans le discours médical concernent essentiellement les parties du corps les plus connues et les plus visibles. Ils sont de fait employés fréquemment dans la langue générale. L'utilisation des termes peu connus chez les non initiés comme *ADN*, *artère du nerf médian*, *bâtonnets* et *coques aérobies à Gram négatif*, *Moelle osseuse*, etc., relève plutôt de la langue spécialisée.

Les exemples cités ci-dessus sont très employés dans les communications courantes. Dans le discours médical, ils sont employés dans des contextes différents, même s'ils désignent pratiquement les mêmes choses que dans la langue générale. Le nom d'une partie du corps sert à désigner une partie du corps. Par exemple, en médecine les noms de partie du corps servent à désigner les endroits à soigner ou l'endroit où le patient a mal. En langue générale, ils servent à désigner la même chose mais sans souci de précision et dans un contexte général.

1.1.3. Noms transdisciplinaires

Certains termes appartiennent à plusieurs domaines. Ils sont considérés comme des termes polyspécialisés (G. Gross & M. Mathieu-Colas 2001). Nous faisons figurer dans le tableau ci-dessous des noms élémentaires qui appartiennent à la fois au domaine de la médecine et à un autre domaine ou à plusieurs autres domaines.

<i>Entrée</i>	<i>Classe</i>	<i>Domaine</i>
<i>goyave</i>	<i>FRUIT</i>	<i>médecine</i> <i>botanique</i> <i>alimentation</i>
<i>ADN</i>	<i>ACIDE</i>	<i>médecine</i> <i>criminalité</i>
<i>scanner</i>	<i>APPAREIL</i>	<i>médecine</i> <i>imprimerie</i>
<i>bactérie</i>	<i>MICRO_ORGANISME</i>	<i>médecine</i> <i>biologie</i>
<i>champignon</i>	<i>MICRO_ORGANISME</i>	<i>médecine</i> <i>biologie</i>

Les noms transdisciplinaires appartiennent à des classes spécifiques et à deux ou à plusieurs domaines différents. Cela ne sous entend pas qu'ils sont polysémiques, mais qu'ils s'emploient juste dans des domaines différents. Par exemple, *goyave* appartient à la classe *FRUIT*, mais il est employé dans le domaine de la médecine en tant que médicament, dans le domaine de la botanique en tant que fruit du goyavier et dans le domaine de l'alimentation en tant qu'aliment.

1.2. Noms dérivés

La dérivation fait partie des procédés linguistiques qui sont largement exploités dans la formation des adjectifs à partir des noms élémentaires du vocabulaire médical. Nous étudions trois types de dérivation : la dérivation préfixale, la dérivation suffixale et la dérivation impropre.

1.2.1. Dérivation préfixale

Les noms dérivés construits par préfixation sont le résultat d'une formation par ajout d'un préfixe aux bases (dites radicaux ou racines). Les lexies qui servent de bases à la formation des noms argumentaux sont des substantifs. Certains préfixes sont significatifs. Ils donnent un sens spécifique aux noms.

<i>Préfixe</i>	<i>Sens</i>	<i>Exemple</i>
<i>bi-</i>	<i>deux</i>	<i>bicuspidé</i>
<i>tri</i>	<i>trois</i>	<i>tricuspide</i>
<i>anti-</i>	<i>contre</i>	<i>antivirus</i>
<i>anti-</i>	<i>contre</i>	<i>anti-champignon</i>
<i>anti-</i>	<i>contre</i>	<i>antiprotozoaire</i>
<i>anti-</i>	<i>contre</i>	<i>antiparasite</i>

Dans *bicuspide* le préfixe *bi* signifie *deux*. *Tri* de *tricuspide* désigne le chiffre trois et *anti* dans *antivirus* signifie *contre*. Nous constatons que les préfixes portent un sens qui participe à la formation du sens du nom. Pour connaître le sens du nom dérivé, il faut connaître le sens de la racine et aussi le sens du préfixe associé à cette racine.

1.2.2. Dérivation suffixale

Les noms suffixés sont construits avec les noms élémentaires sous formes de lexies suivis de suffixes. C'est le type de dérivation le plus répandu dans le vocabulaire médical (J. Ghazi 1985).

Nous remarquons qu'il n'y a pas de règles régissant le système dérivationnel des noms sémantiquement apparentés. Autrement dit, les noms de la même classe ne partagent pas forcément les mêmes affixes dans la formation des adjectifs.

Dans le domaine de l'anatomie, des noms comme *front*, *cervelle*, *intestin* peuvent produire des adjectifs de relation par suffixation

- *cervelle* + *al* = *cervical*
- *front* + *al* = *frontal*
- *intestin* + *al* = *intestinal*

Nous notons que le suffixe *al* ne s'applique pas à tous les noms élémentaires :

- *Peau*, *cutané*
- *Cheveu*, *chevelu*
- *Epiderme*, *épidermique*

Ces variations s'expliquent parfois par l'origine du mot, notamment *cutané* qui provient du nom latin *cutis*. Ainsi, le nom *peau* peut difficilement produire un adjectif à partir de sa base à cause des difficultés liées à l'articulation.

Il en va de même pour les noms de la classe (micro-organisme) dont les adjectifs se terminent différemment, par exemple *bactérie*, *bactérien*, *virus*, *viral*.

Ce que l'on peut retenir à partir de ces exemples, c'est que les adjectifs se forment à partir des noms élémentaires et leurs morphologies se fixent dans la langue et se considèrent en tant que formes correctes et inchangeables.

Les adjectifs relationnels qui sont formés à partir des noms élémentaires ne sont jamais prédicats. Ils sont toujours des arguments dans les énoncés.

La dérivation préfixale permet de voir des distinctions dans la morphologie argumentale (*front, frontal*).

En guise de conclusion, nous constatons que les noms élémentaires servent de bases pour la formation des adjectifs relationnels par ajout de suffixes. Ils fonctionnent à leur tour comme des arguments parce qu'ils n'établissent pas une relation entre deux entités. Nous développons ce point dans la partie consacrée à sémantique des arguments.

1.2.3. Troncation

La troncation est une pratique langagière courante dans le discours médical. Nous l'avons constatée dans des forums sur la santé comme *e-santé.fr*. Les praticiens hospitaliers cherchent souvent à utiliser des mots courts dans leurs discussions quotidiennes, parce que la langue de la médecine contient des mots souvent longs et compliqués. La simplification des termes de la médecine recourt à la troncation, qui se fait par la suppression d'une ou plusieurs syllabes d'un terme afin de rendre sa prononciation plus facile. Par exemple :

<i>Mots</i>	<i>Troncation</i>
<i>congé maternité</i>	<i>congé mat.</i>
<i>tambour à compresse</i>	<i>tamb.</i>
<i>stéthoscope</i>	<i>stétho.</i>
<i>Agglutinoscope</i>	<i>agglutino.</i>
<i>prodafalgan</i>	<i>prodaf.</i>

Nous constatons que la troncation, dans le vocabulaire médical, est variable. Pour certains mots dérivés, on supprime juste une ou quelques syllabes, comme *agglutinoscope* (agglutino), *prodafalgan* (prodaf) et *stéthoscope* (stétho). Pour les mots composés, on tronque soit une partie du deuxième mot comme *congé maternité* (congé mat), soit une partie du premier mot et le deuxième mot, par exemple *tambour à compresse* (tamb).

Pour résumer, nous disons que la troncation est la recherche de la simplicité chez les locuteurs, c'est la raison pour laquelle les modalités de l'abréviation sont différentes d'un mot à un autre.

1.3. Noms composés

Nous étudions ici les noms composés qui fonctionnent uniquement comme arguments.

Les dictionnaires des langues de spécialités comportent plus de locutions nominales que de mots simples (M. Meunier-Crespo 1997). Dans le dictionnaire des arguments, il y a trois sortes de noms composés :

- 1- les composés savants formés à partir de noms d'origine grecque ou latine ;
- 2- les noms composés non savants formés à partir de substantifs de la langue française ;
- 3- les noms composés mixtes : ces derniers sont composés de noms français et de noms grecs ou de noms français et de noms latins.

1.3.1. Les composés savants

Les noms composés savants sont formés par construction ou juxtaposition de deux ou plusieurs mots savants (latin + latin/ grec + grec/ latin + grec ou grec + latin). La composition savante ne concerne pas uniquement les noms prédicatifs de la médecine, mais aussi les noms argumentaux. Ils concernent, en majorité, les noms des micro-organismes tels que les virus, les champignons et les bactéries.

Les exemples suivants illustrent certains types de noms composés savants que nous avons recensés dans le dictionnaire des arguments.

<i>Nom composé savant</i>	<i>Type</i>
<i>clostridium botulinum</i>	<i>nom composé de type AN</i>
<i>clostridium butyricum</i>	<i>nom composé de type AN</i>
<i>heliobacterium modesticaldum</i>	<i>nom composé de type AN</i>
<i>enterobacter cloacae</i>	<i>nom composé de type NN</i>
<i>callid herpesvirus 2</i>	<i>nom composé de type NN Chiffre</i>

Nous constatons que si le nombre de noms composés savants du dictionnaire des arguments est conséquent, leur morphologie n'est pas très variée. Nous avons identifié essentiellement trois types : *adjectif-nom*, *nom-nom*, *adjectif-nom-adjectif*.

1.3.2. Composés non savants

La composition nominale non savante est le résultat de la juxtaposition d'au moins deux noms élémentaires du français. Elle s'appelle aussi composition populaire. Elle est le fruit de nombreux travaux de recherche menés en français et de la favorisation de la recherche médicale dans l'hexagone et dans d'autres pays francophones.

La composition non savante semble la plus importante du point de vue du nombre et de la variation dans la formation des noms composés, comme l'illustrent les exemples suivants :

<i>Nom composé non savant</i>	<i>Type</i>
<i>dent de sagesse</i>	<i>nom composé de type N de N</i>
<i>laboratoire d'analyses médicales</i>	<i>nom composé de type N de NA</i>
<i>cul-de-sac de Douglas</i>	<i>nom composé de type N-de-N de N</i>
<i>dentisterie médico-légale</i>	<i>nom composé de type N A-A</i>
<i>dynamomètre pour la mesure de la force musculaire</i>	<i>nom composé de type N prép la N de la NA</i>
<i>médecine fondée sur des faits</i>	<i>nom composé de type NA prep de N</i>
<i>filtre veine cave</i>	<i>nom composé de type NNN</i>
<i>amalgame dentaire</i>	<i>nom composé de type NA</i>

Nous avons recensé des noms composés ayant des formes différentes. Les formes les plus importantes sont N de N (lame de bistouri) et NA (nerf crânien). Il existe d'autres formes plus complexes telles que N de NA, N-de-N de N, NNN, etc. Nous considérons que c'est la catégorie des noms composés la plus diversifiée, probablement en raison du progrès de la médecine qui entraîne l'évolution et l'enrichissement du vocabulaire médical.

1.3.3. Composés mixtes

Les composés mixtes sont formés par des mots du français et des mots du grec ou du latin. Ce type de composition est favorisé par l'évolution qu'a connue la médecine ces dernières décennies. Cela témoigne, à la fois, de la forte présence du grec et du latin dans le vocabulaire médical et des travaux de recherche, dans ce domaine, en langue française.

Nous donnons ci-dessous quelques exemples de noms composés mixtes :

<i>Nom composé mixte</i>	<i>Type</i>
<i>acide gastrique</i>	<i>nom composé de type NA</i>
<i>gélatine hémostatique résorbable</i>	<i>nom composé de type NAA</i>
<i>épithélium pigmenté de l'œil</i>	<i>nom composé de type NA de N</i>
<i>épithélium de la cornée</i>	<i>nom composé de type N de la N</i>
<i>circonvolution du cingulum</i>	<i>nom composé de type N de N</i>
<i>bord de l'acétabulum</i>	<i>nom composé de type N de le N</i>
<i>vis orthopédique</i>	<i>nom composé de type NA</i>

La syntaxe des noms composés de la médecine est extrêmement diversifiée. Quand on évoque la composition nominale, on évoque systématiquement le figement lexical qui nécessite une attention particulière vu ses différentes caractéristiques. Contrairement aux noms composés qui appartiennent plutôt à la composition dite libre, les composés figés ont un degré de figement assez élevé, cela veut dire qu'ils n'acceptent pas certaines transformations comme l'insertion d'autres éléments linguistique ou la transformation de l'un des noms qui les composent en forme adjectivale. Par exemple :

<i>Nom composé figé</i>	<i>Transformation</i>
<i>dent de sagesse</i>	<i>*dent sage</i>
<i>poignée d'amour</i>	<i>*poignée du grand amour</i>
<i>dent de lait</i>	<i>*dent lactique</i>
<i>culotte de cheval</i>	<i>*culotte hippique</i>
<i>pomme d'Adam</i>	<i>*pomme adamite</i>

Les noms composés, en revanche, acceptent les transformations.

<i>Nom composé</i>	<i>Transformation</i>
<i>nerfs crâniens</i>	<i>nerfs du crâne</i>
<i>atrium cardiaque</i>	<i>atrium du cœur</i>
<i>artère rénale</i>	<i>artère du rein</i>
<i>ligament cervical</i>	<i>ligament du cerveau</i>
<i>Hormone thyroïdienne</i>	<i>hormone de la thyroïde</i>

En plus de cette fixité qui les caractérise, les composés figés sont différents du point de vue de l'opacité et de la transparence. Les noms composés sont totalement transparents, comme : *cancer de sein*, *acide gastrique*, tandis que certaines séquences figées sont plus ou moins transparentes: (*dent de sagesse*, *tendon d'Achille*) et d'autres sont complètement opaques (*pomme d'Adam*, *culotte de cheval* ou *poignée d'amour*).

2. Propriétés sémantiques

Nous traitons les noms élémentaires du domaine de la médecine du point de vue sémantique. Nous nous sommes appuyé sur les ontologies médicales afin d'établir un classement des noms argumentaux du vocabulaire médical. Ainsi, nous les avons répartis en trois groupes : premièrement, les noms de nature biologique comme les parties du corps, les micro-organismes et les liquides organiques. Deuxièmement, les noms du matériel médical comme les instruments, les appareils d'observation et les appareils de mesure. Troisièmement, les noms locatifs désignant les lieux ou les parties de lieux.

Nous évoquons les questions relatives aux classes sémantiques des noms argumentaux et les spécificités de chacune d'entre elles, sans pour autant négliger leur syntaxe.

2.1. Noms de nature biologique

Les noms de nature biologique sont des noms argumentaux, c'est-à-dire qu'ils fonctionnent toujours comme arguments et jamais comme prédicats, parce qu'ils désignent des entités. Autrement dit, les noms de nature biologique n'ont pas de fonction prédicative parce que les prédicats servent à établir des relations dans le cadre de la phrase (*ibid*).

Les noms de nature biologique ne sont pas homogènes du point de vue sémantique. Pour rendre compte des emplois argumentaux dans l'élaboration des dictionnaires électroniques, nous devons exploiter les traits sémantiques de tous les arguments de la classe, en contexte.

Par exemple, le verbe *soigner* n'aura pas le même sens selon qu'il a en position de complément un inanimé concret telle qu'une partie du corps (*Paul soigne sa jambe*), ou un abstrait (*Paul soigne son look*).

La description de l'hyperclasse consiste à prendre en compte la classe ou la sous-classe à laquelle le nom appartient, ainsi que son trait sémantique. Par exemple, les noms ayant comme trait (inanimé concret) n'appartiennent pas forcément à la même classe.

Paul plombe sa dent

**Paul plombe son stylo*

Il s'avère que les noms sémantiquement apparentés, comme les parties du corps, n'ont pas nécessairement la même combinatoire.

Il extrait une dent

Il greffe un cœur

**Il extrait un cœur*

**Il greffe une dent*

2.1.1. Parties du corps

La terminologie médicale relative à l'anatomie du corps humain concerne les parties du corps. Une partie du corps peut elle-même être constituée d'autres parties du corps, lesquelles, à leur tour, peuvent être composées d'autres parties du corps. Par exemples une

tête est composée du crâne, des yeux, des oreilles, de la bouche, etc. La bouche, qui fait partie de la tête, est composée des lèvres, de la langue, des dents, etc. Les parties du corps se distinguent. Certaines d'entre elles ont des fonctions mécaniques comme la main, la tête, le bras, etc. D'autres ont des fonctions physiologiques, comme la vulve pulmonaire, la veine cave supérieure, l'oreillette droite, l'aorte et le myocarde, etc.

Dans la description de noms des parties du corps, nous constatons qu'ils ont des particularités remarquables. Ils sont définis soit par holonymie, comme *appareil digestif* et *appareil respiratoire*, soit par méronymie, comme *estomac* et *poumon*. Les holonymes désignent les parties du corps comme étant des ensembles constitués d'autres parties du corps et les méronymes désignent les parties du corps qui composent les noms de la première catégorie, c'est-à-dire les holonymes (P. A. Buvet & I. Sfar à paraître).

Nous avons recensé 4217 noms de parties du corps humain dont 10% fonctionnent comme des holonymes, par exemple *tête*, et 90% comme des méronymes, par exemple *nez* et *lèvre*. Le classement des noms de parties du corps semble très difficile parce que chaque partie du corps se trouve composée de plusieurs autres parties. Cela nous a poussé à affiner notre classement. Nous avons ainsi établi un classement plus détaillé comprenant les noms des *os* (*tibia*), les noms des cellules, (*cellule souche*), les noms des nerfs, (*nerf crânien*), les noms des glandes, (*glande mammaire*), les noms des veines, (*veine faciale postérieure*), etc. Ces exemples montrent que certains noms fonctionnent tantôt comme des holonymes et tantôt comme des méronymes, étant donné qu'une partie du corps peut être composée d'autres parties du corps.

Les parties du corps humain ont des fonctions et des particularités propres à chacune d'entre elles. Cela sous entend qu'elles peuvent être associées à des verbes et à des adjectifs (*ibid*).

Nous présentons ci-dessous certains verbes et adjectifs associés aux noms des parties du corps humains.

Tête/ G : nf/ C : partie du corps/ D : anatomie

La tête bouge (verbe)

Ventre/ G : nm/ C : partie du corps/ D : anatomie

Le ventre gonfle (verbe)

Le ventre est gonflé (adjectif)

Gonflement du ventre (nom)

Cœur/ G : nm/ C : partie du corps/ D : anatomie

Le cœur bat (verbe)

Battement du cœur (nom)

Os/ G : nm/ C : partie du corps/ D : anatomie

L'os se fracture (verbe)

L'os est fracturé (adjectif)

Fracture de l'os (nom)

Muscle/ G : nm/ C : partie du corps/ D : anatomie

Le muscle se déchire (verbe)

Le muscle est froissé (adjectif)

Déchirure du muscle (nom)

2.1.2. Micro-organismes

Selon le Trésor de la Langue Française informatisé (TLFi), les micro-organismes sont des êtres vivants visibles uniquement au microscope. Dans les sciences biomédicales, il existe deux types de micro-organismes : ceux qui sont constitutifs du corps humains, comme l'*ADN* (acide désoxyribonucléique) et l'*ARN* (acide ribonucléique), et ceux qui ne sont pas constituants du corps humain. Ces derniers sont des êtres vivants microscopiques qui vivent dans la nature et qui peuvent affecter l'organisme, tels que les *virus*, les *champignons* et les *bactéries*. Leurs nombres respectifs, dans le dictionnaire des arguments sont : 252, 74 et 270.

Sur le plan linguistique, il s'agit de noms argumentaux car ils sont incompatibles avec la fonction prédicative.

Les noms de micro-organisme ont des propriétés linguistiques remarquables. Ils sont souvent définis par hypéronymie. Ils ont comme sème commun *micro-organisme* et des

sèmes spécifiques qui les différencient, comme *unicellulaire* ou *pluricellulaire* pour champignon, *infectieux* et *contagieux* pour virus et à *structure très simple* pour bactérie.

Les noms de micro-organisme ont d'autres particularités linguistiques. Par exemple les noms de virus sont utilisés pour désigner toutes les maladies dont ils sont à l'origine comme : *virus de sida*, *virus de la grippe*, *virus de la rage*, etc. Ils sont au nombre de 48 et peuvent avoir une fonction prédicative parce qu'ils désignent des maladies. Ils se combinent ainsi avec des verbes supports comme *attraper* et *contracter*.

Du point de vue de leur syntaxe, les noms de micro-organismes peuvent être associés à des verbes et à des adjectifs inhérents (P. A. Buvet à paraître). Cela constitue chez les locuteurs un savoir linguistique préconstruit (S. Mejri & P. Mogorron 2008).

Nous montrons ci-dessous quelques verbes et adjectifs associés aux noms de micro-organisme.

Virus/ G : nm/ C : micro-organisme / D : microbiologie

Le virus prolifère

Le virus est transmissible

Champignon/ G : nm/ C : micro-organisme / D : microbiologie

Les champignons se propagent

Les champignons sont microscopiques

Bactérie/ G : nf/ C : micro-organisme / D : microbiologie

Certaines bactéries résistent à la chaleur

Les bactéries sont résistantes

ADN/ G : nf/ C : micro-organisme / D : microbiologie

On analyse l'ADN pour identifier les gens

L'ADN est analysable

2.1.3. Liquides organiques

Nous avons dénombré 71 noms de liquides organiques dans le dictionnaire des arguments. Pour les analyser, nous avons pris quelques exemples de noms tels que *sang*, *sueur* et *urine*. Ils sont définis respectivement, dans le trésor de la langue française informatisé (TLFi), comme :

- *Sang* : « *liquide organique de couleur rouge...* »
- *Urine* : « *liquide organique transparent...* »
- *Sueur* : « *liquide organique aqueux...* »

Si les noms de liquides organiques ont, au moins, une propriété linguistique commune, c'est-à-dire l'hypéronyme « liquide organique », ils ont des sèmes spécifiques à chaque nom comme par exemple : *rouge* pour *sang*, *transparent* et *odorant* pour *urine* et *aqueux* pour *sueur*. Ils ont également des particularités linguistiques différentes. Ils peuvent correspondre à des holonymes comme *sang*, ou à des méronymes comme *globules rouges*, *globules blancs* et *plasma sanguin*.

Certains liquides organiques sont constitutifs du corps humain comme *sang*, *humeur aqueuse*, *liquide amniotique*, et d'autres sont produits par le corps puis, par la suite, rejetés pour des raisons d'hygiène interne, comme les *urines* et la *sueur*. Les deux ont des propriétés linguistiques différentes : les premiers admettent par exemple le verbe *circuler* (le sang circule à travers les veines) tandis que les seconds admettent plutôt le verbe *rejeter* (le corps rejette les urine).

Les noms de liquides organiques sont identifiés par les propriétés inhérentes aux liquides en général, notamment leur association avec des verbes comme *couler*. Par contre, la plupart des noms de liquides organiques ont une syntaxe particulière, comme l'illustrent les exemples suivants :

Sang/ G : nm/ C : liquide organique/ D : anatomie

Le sang coule

Il saigne

Sueur/ G : nf/ C : liquide organique/ D : anatomie

La sueur coule

Il transpire

Urine/ G : nf/ C : liquide organique / D : urologie

Les urines coulent

Il urine

Dans *il urine, il saigne et il transpire*, l'argument X1 en position complément est inféré.

2.2. Matériel médical

Les noms de matériel médical regroupent l'ensemble des éléments dont les personnels de la santé se servent dans leur métier ainsi que des équipements physiques dont sont dotés les établissements de santé. Ils sont considérés comme des noms d'artefact et ont une fonction argumentale (P. A. Buvet & I. Sfar à paraître).

Le dictionnaire des arguments contient 282 noms de matériel médical. Ils sont divisés en trois grandes classes. Les instruments (*bistouri, ciseau médical*), les appareils de mesure (*tensiomètre, thermomètre, dynamomètre de colin, éthylotest*) et les appareils d'observation (*stéthoscope, otoscope, lampe frontale*). Nous les décrivons ci-dessous.

2.2.1. Instruments

Dans le trésor de la langue française informatisé (TLFi), le nom « instrument » est défini comme étant un « objet fabriqué en vue d'une utilisation particulière pour faire ou créer quelque chose, pour exécuter ou favoriser une opération ».

Les noms d'instruments médicaux peuvent correspondre à des holonymes, par exemple *seringue, bistouri, boîte de chirurgie*, etc., ou des méronymes, comme *lame de bistouri, aiguille, ciseau*, etc. Ils correspondent également à des synonymes, tels que *bistouri* et *scalpel*.

Certains noms d'instruments médicaux portent en eux les mots qui désignent leur appartenance à un domaine ou à un sous domaine particulier. Les noms *ciseau médical*, *pince anatomique* et *pince hémostatique* sont particulièrement remarquables de ce point de vue. Ils sont formés d'un nom et d'un adjectif, et ce dernier désigne le domaine auquel le nom d'instrument appartient.

Puisque les noms d'instruments médicaux désignent des objets fabriqués en vue d'utilisation particulière (*ibid*). Cela sous entend que chaque nom peut être associé à une fonction. Ils se combinent avec des verbes et des adjectif appropriés tels que :

Bistouri/ G : nm/ C : instrument médical/ D : médecine

Il dissèque avec le bistouri

Le bistouri est tranchant

Ciseau médical/ G : nm/ C : instrument médical/ D : médecine

Il coupe avec le ciseau médical

Le ciseau médical est coupant

Pince anatomique/ G : nf/ C : instrument médical/ D : médecine

Il coupe avec la pince anatomique

La pince anatomique est coupante

Boîte à chirurgie/ G : nf/ C : instrument médical/ D : médecine

Il stérilise la boîte à chirurgie

La boîte à chirurgie est stérilisée

2.2.2. Appareils d'observation

Les noms d'appareils d'observation du domaine médical concernent tous les appareils utilisés pour observer un phénomène. Dans le *Petit Robert*, le nom *appareil* est défini comme étant un « assemblage de pièces ou d'organes réunis en un tout pour exécuter un travail,

observer un phénomène, prendre des mesure¹⁰ ». Nous avons dénombré 8 noms d'appareils d'observation dans le dictionnaire des arguments. Ce sont des appareils utilisés par les praticiens : *stéthoscope*, *anmioscope*, *agglutinoscope*, *coloscope*, *microscope*, *otoscope*, *stomatoscope* et *stroboscope*.

Les noms d'appareils d'observation ont des spécificités linguistiques particulièrement remarquables. Ce sont des mots construits, ils comportent tous le suffixe *scope* qui signifie en grec ancien « observer » (J. Ghazi 1985). Les noms d'appareil désignent notamment des appareils qui servent à observer des parties du corps précises, par exemple *stéthoscope* veut dire littéralement « observer la poitrine », *coloscope* « observer le colon » et *stomatoscope* « observer l'intérieur de la bouche », etc.

Les noms d'appareils d'observation correspondent à des holonymes comme *stéthoscope* puisque il s'agit d'un appareil composé d'autres pièces, comme *pavillon*, *lyre* et *amplificateur acoustique*. Ces derniers sont donc des méronymes.

Comme les noms d'appareils d'observation sont fabriqués pour observer un phénomène, ils sont forcément associés à des verbes comme *observer* et *ausculter*. Nous illustrons ce phénomène avec les exemples ci-dessous.

Otoscope/ G : nm/ C : appareil d'observation / D : médecine

Il inspecte le conduit auditif avec un otoscope

Stéthoscope/ G : nm/ C : appareil d'observation / D : médecine

Il ausculte le patient avec un stéthoscope

Ophthalmoscope/ G : nm/ C : appareil d'observation / D : médecine

Il observe l'œil avec un ophthalmoscope

Loupe binoculaire/ G : nm/ C : appareil d'observation / D : médecine

Il observe avec une loupe binoculaire

¹⁰ Nous utilisons également cette définition pour la description des appareils de mesure parce qu'ils partagent des point communs avec les noms d'appareils d'observation.

2.2.3. Appareils de mesure

Les noms d'appareils de mesure sont incompatibles avec la fonction prédicative. Ils fonctionnent toujours comme arguments. En médecine, ils désignent tous les appareils employés dans des fonctions spécifiques. Ils sont en nombre de 23 dans le dictionnaire des arguments. Ils ont quelques particularités linguistiques. Du point de vue sémantique, ils sont définis par un hypéronyme (appareil de mesure) et des sèmes spécifiques qui les distinguent les uns des autres. Ils correspondent aussi à des holonymes, comme *éthylotest* ou à des méronymes comme *ballon d'éthylotest*, ou à des synonymes, comme : *éthylotest*, *éthylomètre* et *alcootest*.

Chacun des noms d'appareils de mesure se combine avec des verbes appropriés comme *souffler*, *utiliser*, *ranger*, *nettoyer*, *mesurer* et *consulter*. Par exemple :

Thermomètre/ G : nm/ C : appareil de mesure/ D : médecine

Il mesure la température avec un thermomètre

Audiomètre/ G : nm/ C : appareil de mesure/ D : médecine

Il utilise un audiomètre

Éthylotest/ G : nm/ C : appareil de mesure/ D : médecine

Il souffle dans un éthylotest

Tensiomètre électronique bras/ G : nm/ C : appareil de mesure/ D : médecine

Il prend le pouls avec un tensiomètre

2.3. Locatifs

Les noms relatifs à des locatifs de la langue de la médecine sont divisés en deux catégories : les lieux et les parties des lieux. La première catégorie regroupe les lieux abritant des établissements de santé comme *hôpital*, *clinique*, *dispensaire*, *cabinet médical*, *cabinet dentaire*, *sanatorium*, *laboratoire d'analyses médicales*, etc. La seconde concerne les noms de parties des lieux comme *salle de soins*, *bloc opératoire*, *salle de réanimation*, etc. Ces

derniers peuvent être composés d'autres parties de lieux par exemple : *zone septique*, *zone aseptique* et *zone de lavage manuel*, qu'on trouve également dans un bloc opératoire.

Nous étudions également l'emploi métonymique entre les noms de lieux et les noms d'humains.

2.3.1. Lieux

Le concept *lieu* est défini dans les dictionnaires¹¹ comme une « portion déterminée de l'espace ».

Les noms de lieux du domaine médical peuvent correspondre à des holonymes comme *hôpital*, *clinique* et *dispensaire*, ces derniers sont des ensembles constitués d'autres lieux appelés parties de lieux. Ils sont considérés comme étant des établissements habilités à recevoir les malades pour leur procurer des soins. Ils correspondent également à des synonymes, par exemple : *hôpital*, *maternité*, *polyclinique*, et *sanatorium*. Dans le dictionnaire des arguments, nous avons 16 noms de lieux appartenant à la classe sémantique (lieu de sante).

Les noms de lieu s'associent à des verbes appropriés aux locatifs comme *aller à*, *se rendre à* et *quitter* et à des adjectifs comme *délabré* et *vétuste*.

Hôpital/ G : nm/ C : lieu de santé / D : médecine

Quitter l'hôpital

Cet hôpital est bien équipé

Clinique/ G : nf/ C : lieu de santé / D : médecine

Se rendre à la clinique

Cette clinique est vétuste

Dispensaire/ G : nm/ C : lieu de santé / D : médecine

Aller au dispensaire

¹¹ Nous avons pris les définitions du TLFi et du Petit Robert 2007, dans lesquels est défini le concept *lieu*.

Ce dispensaire est délabré

Centre de soins/ G : nm/ C : lieu de santé / D : médecine

Visiter un centre de soins

Ce centre de soins est neuf

2.3.2. Parties de lieux

Du point de vue lexicographique, il n'y a pas de différences claires entre les concepts de *lieu* et de *partie de lieu* du domaine de la médecine. Tous les deux sont considérés comme des locatifs. Mais les deux sont étroitement liés.

Les noms de partie de lieu correspondent à des méronymes comme *salle de soins*, *salle de réanimation*, *bloc opératoire* et *salle d'opération*. Les noms de lieux correspondent à des holonymes, par exemple : *hôpital* et *clinique* (*ibidem*).

Les noms de partie de lieux ont une particularité remarquable : on ne peut pas concevoir le nom d'une partie de lieu sans le nom qui dénote le lieu qui la contient. Autrement dit, il n'y a pas de méronyme sans holonyme et vice versa.

Les noms de partie de lieux sont semblables aux noms de lieux du point de vue syntaxique. Ils se combinent avec les mêmes classes de verbes et d'adjectifs.

Salle de soins/ G : nf/ C : partie de lieu/ D : médecine

Aller à la salle de soins

La salle de soins est éclairée

Bloc opératoire/ G : nm/ C : partie de lieu/ D : médecine

Rentrer au bloc opératoire

Le bloc opératoire est spacieux

Salle de réanimation/ G : nf/ C : partie de lieu/ D : médecine

Se rendre à la salle de réanimation

La salle de réanimation est grande

2.3.3. Emplois métonymiques

Les noms de lieux et les noms d'humains appartiennent à deux classes différentes. Ils sont, néanmoins en relation de métonymie (T. Massoussi 2009), c'est-à-dire que les noms de lieux correspondent aux noms d'humains et vice versa. Les exemples suivants montrent qu'ils peuvent se substituer au sein des phrases:

Les médecins sont en grève

L'hôpital est en grève

Les médecins dispensent les soins gratuitement

La clinique dispense les soins gratuitement

Étant donné que les noms d'humains collectifs désignent les occupants des lieux, il y a une relation occupant/occupé. Les locatifs endossent le sens des occupants, c'est-à-dire, des humains. C'est la raison pour laquelle ils partagent le verbe *dispenser* et l'adjectif *en grève*. Cela montre le rapprochement entre la classe <locatif> et la classe <humain collectif>, par le mécanisme de la métonymie. L'emploi métonymique des noms de lieu, notamment *hôpital* et *clinique* explique la combinatoire commune car ils impliquent un transfert de classe.

Les parties de lieux s'analysent également comme des anaphores nominales (G. Kleiber 1994 & P. A. Buvet à paraître). Par exemple :

*J'étais admis dans un hôpital, le **toit** était délabré*

*On a construit un hôpital, les **murs** sont en préfabriqué*

La compréhension et la désambiguïsation des mots *toit* et *murs* nécessitent de préciser leurs classes sémantiques et la distribution des prédicats et des arguments avec lesquels ils apparaissent dans les phrases (D. Le Pesant 2002).

En guise de conclusion, nous disons que les noms des établissements de santé peuvent être repris, dans la langue, par une autre catégorie de noms. Il s'agit des noms d'humains avec lesquels ils entretiennent une relation métonymique (T. Massoussi 2009).

3. Propriétés combinatoires

Nous étudions ici la combinatoire des noms argumentaux avec les prédicats sous trois formes différentes : verbale (*disséquer un cadavre*), nominale (*contraction des muscles*) et adjectivale (*l'os est fissuré*). Nous traitons les verbes, les noms et les adjectifs selon qu'ils appartiennent à la langue générale ou à la langue spécialisée et selon leur morphologie (verbes simples et locutions verbales, noms simples et noms composés). Nous étudions ensuite l'actualisation des noms élémentaires.

3.1. Verbes prédicatifs

Les prédicats jouent un rôle important dans la description des noms argumentaux¹² du domaine de la médecine. Qu'ils appartiennent au domaine spécialisé, notamment à la médecine, ou au domaine de la langue générale, les verbes prédicatifs donnent souvent la possibilité de décrire, en extension, toutes les classes d'arguments (*ibid*).

3.1.1. Verbes spécialisés

Les verbes spécialisés de la langue de la médecine ont comme particularité de ne se combiner qu'avec les noms argumentaux médicaux, en particulier, les noms de nature biologique comme les parties du corps, les micro-organismes et les liquides organiques, les noms de matériel médical comme les instruments, les appareils d'observation et les appareils de mesure, ainsi que les noms de lieu.

a- noms de nature biologique

La combinatoire des verbes spécialisés avec les noms de nature biologique impose certaines contraintes. L'élaboration de classes sémantiques répondant au classement de l'ontologie médicale facilite, en quelque sorte, la description syntaxique des noms de nature biologique. Certaines classes de noms peuvent se combiner avec le même verbe, comme *proliférer* (les micro-organismes prolifèrent) et *soigner* (soigner les parties du corps). Mais les verbes génériques, c'est-à-dire qui s'appliquent à toute la classe sont très limités dans notre

¹² Nous étudions la combinatoire des classes des noms argumentaux traités dans la partie sémantique, notamment les noms de nature biologique, les noms de matériel médical et les noms de locatifs.

dictionnaire des prédicats. Nous illustrons cela avec des exemples précis, où nous montrons que chaque nom de partie du corps peut se combiner avec un verbe d'une façon exclusive. Autrement dit chaque nom sélectionne un verbe spécialisé approprié.

Amputer une jambe

**Amputer une tête*

Plâtrer une main

**Plâtrer un rein*

Cautériser le nerf

**Cautériser un œil*

Suturer une plaie

**Suturer une dent*

Bien que tous les noms combinés aux verbes dénotent des parties du corps, nous constatons que les verbes spécialisés *amputer*, *plâtrer*, *cautériser* et *suturer* ne s'appliquent pas tous aux noms de partie du corps.

b- les noms de matériel médical et les noms de lieux

Les noms de matériel médical et les noms de lieux de la langue de la médecine¹³ se combinent avec les verbes spécialisés. Ils ont notamment des particularités communes. Les noms de matériel médical et les noms de lieux médicaux ne fonctionnent qu'avec très peu de verbes spécialisés, par exemple :

Stériliser un scalpel

Stériliser un bloc opératoire

Quand ils se combinent avec les verbes spécialisés, les noms de matériel médical fonctionnent en position d'argument 2 en tant que circonstant, par exemple :

Suturer la plaie avec une agrafe

Inciser un ligament avec un scalpel

¹³ Nous avons mis les noms de matériel médical et les noms de lieux de la langue de la médecine ensemble parce qu'ils ont une combinatoire identique concernant leur fonctionnement avec les verbes spécialisés.

Auscultation d'un patient avec un stéthoscope

Les noms de lieux apparaissent en position de complément d'objet indirect comme :

*Évacuer le patient vers la clinique**Secourir le patient vers l'hôpital***3.1.2. Verbes non spécialisés**

Les verbes non spécialisés sont largement utilisés dans les langues spécialisées. En l'occurrence, dans la langue de la médecine, chaque partie du corps ou presque sélectionne « ses verbes ». Ils s'utilisent dans le domaine spécialisé avec des emplois spécifiques, ou plutôt spécialisés. Les emplois spécialisés sont acquis dans le cadre de la phrase qui est, généralement, constituée d'arguments spécialisés. Les noms élémentaires relatifs aux parties du corps, micro-organismes, médicaments, instruments médicaux, établissements de santé, etc., jouent un rôle primordial dans le sens du verbe, puisque les linguistes, tels que G. Gross et M. Mathieu-Colas, dans le cadre de leurs travaux au sein du laboratoire LDI, s'accordent à dire que c'est la classe sémantique des arguments qui définit l'emploi et le sens du prédicat verbal.

Nous vérifions cela avec des exemples tirés de notre corpus à partir des prédicats verbaux qui se combinent avec les trois classes, en l'occurrence les noms de nature biologique, les noms de matériel médical et les lieux.

*a- noms de nature biologique**Le muscle se froisse**L'os se fissure**Extraire une dent**Les veines gonflent**L'os se fracture**Le ventre gonfle*

Certains verbes spécialisés fonctionnent avec un seul argument. Ils véhiculent un sens qui désigne de façon implicite l'argument 2. Il fait référence à un nom élémentaire, notamment un liquide organique. Autrement dit, l'argument 2 est inféré.

Il transpire (sueur)

Il urine (urine)

Il saigne (sang)

L'usage de ces verbes nécessite la suppression de l'argument en position complément ; le cas contraire est considéré comme un pléonasme.

**Il saigne le sang*

**Il pisse les urines*

**il transpire la sueur*

La seule exception dans le cas où le verbe *saigner* est remplacé par *pisser*, pour exprimer l'abondance du sang qui coule.

Il pisse le sang

b- les noms de matériel médical

Utiliser une seringue

Ranger la pince anatomique

Laver un ouvre bouche

Nettoyer la pince nasale

c- les noms de lieux

Être admis à un sanatorium

Aller dans un dispensaire

Quitter un asile

Visiter un centre de soins

Les noms élémentaires du vocabulaire médical se combinent à la fois avec les verbes prédicatifs généraux et avec les verbes prédicatifs spécialisés. Nous constatons que le problème de l'ambiguïté est rare lorsqu'il s'agit des verbes spécialisés. Ces derniers ne sont pas polysémiques dans les énoncés où ils apparaissent grâce à leur appartenance exclusive à la médecine. Cela signifie qu'ils n'ont qu'un seul sens comme *disséquer, suturer, cautériser et hospitaliser*, etc. Les autres appartiennent à la langue générale comme *casser, fissurer, froisser*, etc. Ils sont largement polysémiques, ils acquièrent leur sens dans le cadre de la phrase.

3.1.3. Actualisation des noms élémentaires lorsqu'ils se combinent avec des verbes prédicatifs

L'actualisation est définie comme le moyen de produire des énoncés bien formés à partir de la structure prédicat-argument. Elle implique des catégories énonciatives (personne, temps, aspect, modalité, etc.) qui permettent l'ancrage des prédicats et des arguments dans un discours, en fonction de la position du locuteur par rapport à ce qu'il énonce. Les actualisateurs participent directement ou indirectement à l'instanciation d'une structure prédicat-argument selon qu'ils sont nécessaires pour construire des phrases simples ou bien qu'ils ne le sont pas (P.A. Buvet 2009).

Nous nous intéressons à la détermination des noms argumentaux du domaine médical, en l'occurrence, les instruments médicaux, les médicaments et les locatifs, lorsqu'ils se combinent avec des verbes prédicatifs.

Le médecin stérilise (la + une + les + plusieurs) lame(s) de bistouri(s)

Le médecin utilise (le + un + les + plusieurs) stéthoscope(s)

Le médecin se sert de (la + une + plusieurs) pince(s) nasale(s)

Les substantifs, *lame de bistouri, stéthoscope et pince nasale*, ne posent pas beaucoup de contraintes du point de vue de la détermination. Ils admettent les déterminants qui expriment la définitude et la non définitude, comme les déterminants *le, un et les*. Le déterminant *plusieurs* dans les exemples ci-dessus a une valeur quantitative qui actualise les arguments nominaux *lame de bistouri, stéthoscope et pince nasale*.

Il a pris (?une + ?deux + un comprimés + deux comprimés) d'aspirine(s)

*Il a pris (*un + *deux + une cuillère à café de + deux cuillères à café de) biocaliptole*

*Il a appliqué (une + ?deux + une fois + deux fois + *un litre de + * une cuillère de)(la) pommade sur sa jambe*

On constate que les noms de médicaments comme *aspirine*, *biocaliptole* et *pommade* n'acceptent pas les mêmes déterminants quantifieurs parce qu'ils correspondent à substances à formes différentes. Les déterminants quantifieurs ne s'appliquent pas forcément à tous les noms de médicaments puisque *aspirine* admet *un comprimé* et pas *une cuillère à café*, et *biocaliptole* admet *une cuillère à café* mais pas *un comprimé*, parce qu'ils sont de natures différentes (solide/ liquide).

Paul a désinfecté (plusieurs fois + une fois + ?une partie de +) la salle (s) de soins

*Paul a séjourné dans (un + plusieurs + *une partie de) hôpital / hôpitaux*

Les noms de parties de lieux de la médecine peuvent être actualisés par les déterminants quantifieurs, comme le cas du nom *salle de soins* avec le prédicat *désinfecter*. Il accepte *plusieurs* et *une fois*. Par contre les noms de lieux comme *hôpital* et *clinique* qui se combinent avec le prédicat *séjourner* admettent *un* comme article indéfini, *plusieurs fois* en tant que quantifieur, mais ils ne se combinent pas avec le partitif *une partie de*.

3.2. Adjectifs prédicatifs

Les noms argumentaux de la médecine se combinent avec des adjectifs spécialisés et des adjectifs non spécialisés. Nous étudions ci-dessous leurs spécificités. Dans cette partie, nous nous focalisons sur les adjectifs qui se combinent notamment avec les noms de nature biologique, parce que c'est la seule catégorie où nous pouvons distinguer la variation entre les adjectifs spécialisés et non spécialisés. Les noms de matériel médical et les noms de lieux n'ont pas cette particularité, excepté l'adjectif *stérilisé*, par exemple : *instrument stérilisé* et *bloc opératoire stérilisé*.

3.2.1. Adjectifs spécialisés

Les adjectifs spécialisés sont propres à la langue de la médecine. Leur emploi s'avère univoque et ne génère pas d'ambiguïtés. Souvent, les noms élémentaires comme les parties du corps se combinent avec des adjectifs relatifs aux maladies qui les affectent.

Cellule pathologique

Cellule cancérigène

Foie intoxiqué

Articulation luxée

Cellule hyperplasiée

Gencive hypertrophique

Les noms de micro-organismes se combinent avec les adjectifs qui signifient leurs caractéristiques intrinsèques.

Bactérie aérobie

Bactérie anaérobie

Agent pathogène

Ces exemples peuvent avoir deux lectures différentes, une lecture compositionnelle et une lecture non compositionnelle. Dans le premier cas, on considère le nom comme un argument, puisque c'est un nom élémentaire, et l'adjectif comme un prédicat. Dans le second cas, la séquence est considérée comme un nom composé. Il constitue alors un argument qui fonctionne avec un verbe prédicatif.

3.2.2. Adjectifs non spécialisés

Les noms élémentaires s'accordent avec les adjectifs prédicatifs, qui relèvent d'un autre niveau de langue. L'étiquette « non spécialisé » ne veut pas dire que les noms se combinent d'une façon fortuite ou avec n'importe quel adjectif. Elle signifie que ces derniers appartiennent au domaine de la langue générale, mais leur sélection est rigoureuse.

Par exemple, chacune des parties du corps, ou du moins, chacune des classes (partie du corps) sélectionne ses propres adjectifs.

Ventre ballonné

Os cassé

Muscle contracté

Certains adjectifs ne se combinent pas uniquement avec un nom de partie du corps mais avec plusieurs noms :

Œil gonflé

Lèvre gonflée

Main gonflée

Pied gonflé

Néanmoins, il est impossible d'associer des adjectifs à certaines parties du corps tels que :

**Os gonflé*

**Os froissé*

**Os déchiré*

Dans les exemples ci-dessus, les adjectifs *gonflé*, *froissé* et *déchiré* ne sont pas appropriés au nom *os*.

Les noms argumentaux peuvent accepter simultanément deux adjectifs, spécialisés ou non spécialisés :

Il a pratiqué l'ablation du foie infecté et nécrosé

Il a arraché la dent jaune cariée

Il a traité son asthme bronchique chronique

3.3. Noms prédicatifs

Les noms élémentaires se combinent également avec les noms prédicatifs spécialisés et les noms prédicatifs non spécialisés, ainsi qu'avec les noms simples et les noms composés.

3.3.1. Noms spécialisés vs noms non spécialisés

Nous traitons, successivement, la combinatoire des noms spécialisés et celle des noms non spécialisés avec les noms argumentaux, en l'occurrence, les noms de nature biologique, les noms de matériel médical et les noms de lieux du domaine médical.

Les noms spécialisés se combinent avec des noms argumentaux qui, eux-mêmes sont des noms spécialisés. Ils apparaissent dans des constructions à compléments :

Il a un herpès sur la lèvre

Il a un hématome au pied

Il a une plaie à l'orteil

Il a un bouton sur la joue

Nous constatons que les noms spécialisés qui se combinent avec les noms de parties du corps désignent tous des maladies, des blessures ou des anomalies.

Les noms non spécialisés sont fréquemment employés avec les noms du domaine médical, pour désigner des caractéristiques physiologiques des parties du corps :

Les battements de cœur

Les contractions des muscles

des opérations chirurgicales ou examens médicaux :

Greffe du cœur

Suture du crâne

ou encore des symptômes :

Chute de tension artérielle

Prise de pouls

Les noms de lieux et les noms de matériel médical se combinent également avec les noms spécialisés et les noms non spécialisés pour désigner des caractéristiques particulières : C'est le cas des noms spécialisés comme :

Hôpital des lépreux

Hôpital des aveugles

Ou des noms non spécialisé comme :

Hôpital de compagnie

Bistouri à usage unique

Certains exemples peuvent être source de paraphrase. Les mêmes concepts peuvent être exprimés avec des formes verbales ou adjectivales :

<i>Forme nominale</i>	<i>Forme verbale</i>	<i>Forme adjectivale</i>
<i>analyse de sang</i>	<i>analyser le sang</i>	<i>analyse sanguine</i>
<i>battement du cœur</i>	<i>le cœur bat</i>	<i>battement cardiaque</i>
<i>contraction du muscle</i>	<i>le muscle se contracte</i>	<i>contraction musculaire</i>

Il n'en va pas de même pour les noms composés qui sont totalement figés comme *dent de sagesse*, pour lesquels la paraphrase est impossible.

Dent de sagesse

**Dent sage*

3.3.2. Noms simples vs noms composés

Les noms de la langue de la médecine se combinent aussi bien avec les noms composés qu'avec les noms simples. Par exemple :

Une analyse de sang

Une analyse médicale de sang

Les différences morphologiques peuvent-elles être contraignantes pour la syntaxe des noms élémentaires ? C'est à quoi nous essayons de répondre dans ce qui suit.

Nous avons abordé la combinatoire des noms élémentaires avec les noms prédicatifs ci-dessus, avec les noms spécialisés et les noms non spécialisés. Cela englobe leur combinatoire

avec les noms prédicatifs simples. Nous y revenons pour la comparer avec celle des noms composés, afin de relever les ressemblances et les différences.

Nous nous intéressons à l'étude de la combinatoire des noms composés avec les noms élémentaires médicaux. Les noms élémentaires se combinent avec des noms prédicatifs composés. Nous abordons dans cette partie les deux cas de figures, c'est-à-dire les noms simples et les noms composés, et nous examinons les relations qu'ils entretiennent.

a) noms simples

Faire des analyses d'urines

Prescrire des examens de selles

Pratiquer une analyse des gaz

b) noms composés

Faire un test d'allergie de la peau

Prescrire une analyse médicale de sang

Pratiquer un examen bactériologique d'urines

Faire un examen parasitologique des selles

Faire un examen médical du liquide céphalo-rachidien

D'après les exemples ci-dessus, nous constatons que les combinaisons des noms simples et des noms composés avec les noms élémentaires ne sont pas différentes. Néanmoins, les noms composés portent des précisions supplémentaires d'ordre sémantique lorsqu'il s'agit, par exemple, d'*analyses du sang* et d'*analyses médicales du sang*. Le mot *médical* précise le domaine, mais il n'est pas indispensable pour comprendre le sens global de l'énoncé. En revanche, dans *analyses d'urines* et *analyses bactériologiques d'urines*, le mot *bactériologique* implique la présence d'un agent pathogène.

Conclusion

L'étude des noms argumentaux du vocabulaire médical a permis de montrer leurs caractéristiques morphologiques et sémantiques ainsi que leur fonctionnement syntaxique. Les résultats de cette analyse donnent un aperçu sur le fonctionnement des noms du domaine médical et sur la fonction argumentale. Prendre en charge la fonction argumentale englobe en même temps la fonction prédicative et la fonction actualisatrice, parce que les noms argumentaux ne sont pas analysables de manière isolée, mais dans le cadre de la phrase.

CHAPITRE 6

LES EMPLOIS DES NOMS D'HUMAINS DU DOMAINE MÉDICAL

Introduction

La description d'une langue doit accorder une place particulière aux noms d'humains car ils ont des emplois spécifiques. En langue générale, comme en langue de spécialité, les noms d'humains sont importants dans la réduction de l'ambiguïté. Certains verbes se combinent uniquement avec des noms d'humains en position sujet, par exemple *penser*, *réfléchir*, *parler*, etc. ; ces verbes excluent néanmoins certains types d'humains :

Mon voisin conduit une voiture

**Mon bébé conduit une voiture*

Ce juge condamne le malfaiteur

**Ce nourrisson condamne le malfaiteur*

Cette remarque s'applique aussi bien aux langues spécialisées qu'à la langue générale. Par exemple, le rapport aux métiers explique les restrictions de sélection que l'on observe avec des verbes comme *soigner*, *opérer*, *disséquer*, *vacciner*, etc.

L'anesthésiste a anesthésié Paul

**L'aide soignant a anesthésié Paul*

Afin de réduire l'ambiguïté des verbes qui se combinent avec les noms d'humains du domaine médical, nous les avons divisés en plusieurs classes et sous-classes, par exemple (humain malade) et (humain médecin), etc., selon le classement ontologique médical.

Puisque le statut des noms d'humains est variable, nous expliquons, avec des exemples, tous les cas de figures où ils fonctionnent comme des prédicats et comme des arguments.

1. Propriétés morphologiques

Cette partie est divisée en trois sous-parties. Dans la première sous-partie, on étudie la formation des noms simples, dans la deuxième sous-partie, la formation des noms dérivés et dans la troisième sous-partie, la formation des mots construits à partir d'autres procédés comme la composition, la troncation et l'ellipse.

1.1. Les noms simples

Dans la catégorie des noms simples, nous étudions les noms spécialisés qui appartiennent exclusivement au vocabulaire médical et les noms non spécialisés, c'est-à-dire ceux qui appartiennent à la langue générale et qu'on utilise dans la langue de la médecine et les noms transdisciplinaires. Ces derniers peuvent appartenir à deux ou à plusieurs domaines spécialisés.

1.1.1. Les noms spécialisés

Les noms simples d'humains utilisés dans le vocabulaire médical servent à désigner soit les humains professionnels comme *médecin* et *infirmier*, soit les humains malades comme *lépreux* et *cancéreux*.

L'utilisation des noms d'humains spécialisés de formes simples dans le vocabulaire médical sert à désigner les personnes du domaine médical. Du point de vue morphologique, ils ne sont ni dérivés d'autres mots, ni composés. Par contre, ils sont utilisés pour la formation des noms dérivés qui ne sont pas forcément des noms d'humains, par exemple *médecin* / *médecine*, *sourd* / *surdité* / *pâle* / *pâleur*, *blessé* / *blessure*, etc., et des noms composés comme *médecin spécialiste* et *médecin interne*, cf. *infra*.

Nous illustrons dans le tableau suivant quelques exemples des noms d'humains spécifiques à la médecine.

<i>Nom simple</i>	<i>Domaine</i>
<i>médecin</i>	<i>médecine</i>
<i>infirmier</i>	<i>médecine</i>
<i>malade</i>	<i>médecine</i>
<i>patient</i>	<i>médecine</i>
<i>pédiatre</i>	<i>médecine</i>

1.1.2. Les noms non spécialisés

Du point de vue morphologique, ils ont les mêmes caractéristiques que les noms spécialisés. Cependant, ils appartiennent à un registre de langue différent. Les noms d'humains non spécialisés utilisés dans le vocabulaire médical appartiennent à la langue générale, mais ils prennent un sens spécifique dans le contexte médical. Par exemple, en dehors du contexte médical, les noms d'humains comme *interne* prennent une toute autre signification, différente du sens de *médecin interne*.

Les exemples des noms d'humains simples sont très rares dans le dictionnaire des humains. Sur 167 noms qui appartiennent à la langue générale, seule une dizaine d'entre eux ont une forme simple, les autres sont dérivés de noms, d'adjectifs ou de verbes. Nous verrons cela dans la partie consacrée à la dérivation et à la composition.

En ce qui concerne les noms de malades et d'handicapés comme *fou*, *folle*, *aveugle* et *sourd*, nous les avons classés dans la liste des noms d'humains non spécialisés parce que leur utilisation peut aussi bien s'étendre à la langue générale qu'à la langue de la médecine. Autrement dit, on utilise beaucoup plus les noms comme *fou*, *aveugle* et *sourd* en société que dans le milieu hospitalier.

<i>Nom simple</i>	<i>Domaine</i>
<i>interne</i>	<i>médecine</i>
<i>fou</i>	<i>médecine</i>
<i>obèse</i>	<i>médecine</i>
<i>aveugle</i>	<i>médecine</i>
<i>sourd</i>	<i>médecine</i>
<i>père</i>	<i>médecine</i>
<i>mère</i>	<i>médecine</i>
<i>adulte</i>	<i>médecine</i>

1.1.3. Les noms transdisciplinaires

Sur le plan morphologique, les noms transdisciplinaires utilisés dans le domaine médical n'ont pas de fonctionnement particulier ni de caractéristiques spécifiques. Leur seule différence réside dans leur appartenance à la fois à deux ou à plusieurs domaines, dont le domaine médical. Nous avons tenu à le préciser, même si cela ne relève pas de la morphologie.

Dans la classification des lexèmes, selon qu'ils appartiennent à la langue générale ou aux langues de spécialité, ces termes appartiennent généralement à deux ou à plusieurs domaines et ne sont pas représentés en langue générale. Ce principe, qui concerne les catégories du discours comme *les verbes*, *les noms* et *les adjectifs*, etc., peu aussi s'appliquer aux noms d'humains.

Les noms comme *docteur* et *professeur* sont à cheval entre l'enseignement supérieur et la médecine. Un *docteur* en enseignement supérieur est une personne ayant soutenu une thèse de doctorat dans une filière particulière, et *docteur* en médecine désigne une personne possédant le titre de docteur en médecine après obtention du diplôme et exercice du métier de médecin ou de chirurgien.

Un *praticien* peut être un clinicien, c'est-à-dire un médecin qui pratique son métier dans une clinique ou un autre établissement de santé. Il peut signifier aussi une personne ou un ouvrier qui dégrossit l'ouvrage de sculpture.

Le nom d'humaine *victime* est à cheval entre le domaine médical et le domaine social. Elle peut être une personne qui subit un accident, dans ce cas, la blessure est physique, ou bien une personne qui subit un préjudice, c'est-à-dire, victime d'une arnaque ou d'une escroquerie.

1.2. Les noms dérivés

La plupart des noms d'humains en médecine sont des dérivés (nous avons relevé 45% de noms d'humains dérivés dans le dictionnaire des humains). Nous étudions ci-dessous les cas de la dérivation suffixale, de la dérivation préfixale et de la dérivation impropre.

1.2.1. Dérivation suffixale

Les noms d'humains du domaine médical formés par dérivation suffixale sont variés. Les suffixes ajoutés aux noms d'humains sont différents d'un cas à l'autre. Nous donnons des exemples ci-dessous pour plusieurs cas de figure permettant d'obtenir un nom d'humain par suffixation.

Certains noms se voient rajouter le suffixe *ier* pour marquer le rapport entre le lieu et les personnes qui y travaillent comme *hôpital*, *hospitalier*. Dans ce cas, c'est le locatif *hôpital* qui sert de base pour la formation du nom d'humain. C'est le même cas pour le nom du lieu *clinique* qui prend le suffixe *ien* pour la formation du nom d'humain *clinicien*.

Les noms de quelques spécialités médicales servent de bases pour la formation de noms d'humains. Par exemple, on rajoute le suffixe *ien* à *chirurgie* pour obtenir le nom *chirurgien*. Mais toutes les spécialités médicales ne prennent pas le suffixe *ien* afin de former les noms des spécialistes qui exercent ces disciplines. Dans *cardiologue*, le nom dérivé provient de la spécialité médicale *cardiologie*, mais par ajout du suffixe *-gue*. C'est la même chose pour *radiologue* qui est formé avec le même suffixe *-gue* à partir de la spécialité *radiologie*. Ce dernier peut aussi accepter le suffixe *-iste* dans le cas de la formation du nom d'humain *radiologiste*. Les possibilités varient selon les spécialités médicales.

On observe la même irrégularité dans les noms de malades formés par suffixation. Certains d'entre eux prennent *eux* pour le masculin et *euse* pour le féminin comme *cancéreux* et *lépreux*, etc., d'autres prennent *ique* comme *diabétique*.

Nous constatons qu'il y a une certaine liberté dans la formation des noms d'humains dans le vocabulaire médical.

1.2.2. Dérivation préfixale

Nous avons relevé quelques exemples de notre dictionnaire qui appartiennent essentiellement au répertoire des noms de malades comme : *hypertendu*, *hypotendu*, *hypoglycémique*, *hyperglycémique*, etc. Leur nombre ne dépasse pas la trentaine.

Ces noms ne sont pas nombreux, mais ils ne sont pas négligeables puisqu'ils portent des préfixes significatifs. Cela peut certainement représenter une avancée considérable pour l'analyse du vocabulaire médical, étant donné que la morphologie des termes joue un rôle essentiel dans la reconnaissance des termes et de leurs significations.

Nous constatons que les noms d'humains formés par préfixation portent généralement les mêmes préfixes que les noms de maladies comme *hypo* et *hyper*. Les noms de malades et les noms de maladies formés avec le même préfixe sont étroitement liés. Autrement dit, leur base est commune comme le cas de *glycémie* pour *hypoglycémie* et *hypoglycémique*, et *tension* dans *hypertension* et *hypertendu*.

1.2.3. Dérivation impropre

La dérivation impropre consiste à former un nouveau mot sans pour autant changer sa morphologie, seule sa catégorie grammaticale qui change. Ainsi des verbes deviennent des noms *diner* (le diner), *manger* (le manger), des adjectifs deviennent des substantifs, comme *il est malheureux* (le malheureux), *il est sportif* (le sportif), etc.

Ce procédé est très fréquent dans la formation des noms d'humains dans la langue de la médecine. Nous avons relevé 274 noms de ce type. Nous avons pris des exemples essentiellement dans les noms de praticiens faisant état de leurs qualités ou de leurs caractéristiques. Nous avons ainsi les noms formés depuis adjectifs:

Il est asthmatique (l'asthmatique a une respiration sifflante)

Il est dépressif (les dépressif sont parfois suicidaires)

Paul est malade (le médecin soigne le malade)

Paul est diabétique (le diabétique suit un régime alimentaire rigoureux)

Paul est sidatique (les sidatiques suivent des traitements lourds)

Les noms de malades qui sont formés par dérivation impropre depuis les adjectifs sont en rapport avec les maladies dont ils souffrent.

Nous constatons également que la formation des noms d'humains par dérivation impropre est généralement faite par la détermination : on rajoute à chaque fois des déterminants aux adjectifs ou aux verbes pour obtenir les noms.

1.3. Autres noms construits

Nous étudions ici d'autres procédés de formation des noms d'humains du vocabulaire médical en français. Il s'agit de la composition, de la troncation et de l'ellipse.

1.3.1. Composition

La dénomination dans les langues spécialisées se base beaucoup sur la composition (P. Lerat 1995). En observant notre dictionnaire, nous avons constaté que le nombre de noms composés du vocabulaire médical est aussi important que le nombre de noms simples (170 noms d'humain composés).

Le sens des mots composés ne signifie pas forcément la même chose que les mots qui le composent. Autrement dit, le sens des mots composés n'est pas toujours déductible du sens des mots qui le composent. Ainsi, *sage femme* ne veut pas dire *femme sage*. Aucun de ces deux termes ne renvoie à la médecine. Cependant, dans la plupart des cas, dans la dénomination des humains dans le vocabulaire médical, les noms sont transparents, c'est-à-dire que nous pouvons déduire le sens du nom composé à travers les mots simples qui contribuent à sa construction comme *médecin généraliste*, *chirurgien dentiste*, *aide soignant*, *élève infirmier*, etc.

Les noms d'humains de la médecine se distinguent, généralement, par les mots qui désignent soit leur fonction, soit leur profession, comme l'attestent les exemples ci-dessus. Néanmoins, certains noms de fonction ne sont pas propres à la médecine : *chef de service*, par exemple, peut aussi appartenir au service éducatif ou au service de police municipale.

Quelques noms composés d'humains sont formés par nécessité de distinguer le féminin du masculin. Le nom de *médecin femme* est ainsi formé pour désigner les femmes qui sont médecins. Traditionnellement, on forme le féminin par ajout de *e* à la fin du nom au masculin. Mais cette option n'est pas possible parce que la forme *médecine* est prise par le nom du domaine (la médecine).

La morphologie des noms d'humains varie d'un nom à un autre. Ils peuvent exister sous plusieurs formes, comme dans les exemples suivants :

<i>Morphologie</i>	<i>Exemple</i>
N de N	<i>médecin de travail</i>
NA	<i>médecin traitant</i>
AN	<i>sage-femme</i>
N à ANN	<i>nourrisson à faible poids naissance</i>
N Adv ANN	<i>nourrisson très faible poids naissance</i>
N de Chiffre N et Adv	<i>adulte de 80 ans et plus</i>
N de NA	<i>adultes d'âge moyen</i>
N en N pour la N	<i>professionnels en éducation pour la santé</i>
N en N	<i>étudiant en dentisterie</i>
N en NA	<i>malades en phase terminale</i>
N Pp de NA	<i>personnes souffrant d'un trouble mental</i>
N Pp en N	<i>adolescent placé en institution</i>
NAA	<i>personne âgée fragile</i>

1.3.2. Troncation

D'une façon générale, les locuteurs cherchent à communiquer plus facilement. La complexité du vocabulaire médical justifie le recours à toutes sortes d'abréviations. Il s'agit ici de la troncation, c'est-à-dire de la suppression d'une ou plusieurs syllabes du terme. Elle

s'effectue généralement sur les mots les plus longs. Ce genre d'abréviation est très fréquent dans les communications quotidiennes dans le milieu hospitalier. C'est ce que nous avons constaté dans des forums sur la santé.

La troncation des noms d'humains en médecine se fait essentiellement par suppression de syllabes à la fin du mot. Ce procédé s'appelle l'apocope, contrairement à l'aphérèse qui consiste à enlever les syllabes en début de mots.

Nous donnons ci-dessous quelques exemples de noms d'humains formés par troncation.

<i>Troncation</i>	<i>Forme complète (humain)</i>
<i>ostéo</i>	<i>ostéopathe</i>
<i>psy</i>	<i>psychiatre</i>
<i>ophtalmo</i>	<i>ophtalmologue</i>
<i>traumato</i>	<i>traumatologue</i>
<i>doc</i>	<i>docteur</i>
<i>gynéco</i>	<i>gynécologue</i>
<i>dermato</i>	<i>dermatologue</i>

Ce procédé sert à former des néologismes de forme dans le domaine médical. Ils peuvent prêter à confusion lorsque les mêmes mots tronqués servent à la formation des spécialités de la médecine que les praticiens exercent. Par exemple :

Troncation	Forme complète (fonction)
<i>ostéo</i>	<i>ostéopathie</i>
<i>psy</i>	<i>psychiatrie</i>
<i>ophtalmo</i>	<i>ophtalmologie</i>
<i>traumato</i>	<i>traumatologie</i>
<i>gynéco</i>	<i>gynécologie</i>
<i>dermato</i>	<i>dermatologie</i>

Néanmoins, dans le cadre de la phrase, les termes tronqués ne peuvent pas être ambigus parce que la structure de la phrase permet de distinguer les humains médecins et les spécialités de la médecine :

Il a consulté un psy

Elle voit un gynéco pour sa grossesse

1.3.3. Ellipse

L'ellipse consiste dans la suppression d'un ou plusieurs constituants du mot composé qui sont normalement indispensables pour la compréhension du mot. En général, les ellipses sont le résultat de la suppression du substantif des séquences formées par des substantifs et des adjectifs. C'est-à-dire les substantifs tombent et les adjectifs restent (G. Skvortsov 1994). Dans ce cas, les termes qui sont initialement des adjectifs dans les séquences formant des noms composés deviennent des substantifs par le maintien des déterminants initiaux.

Le médecin spécialiste → le spécialiste

Le médecin généraliste → le généraliste

Le chirurgien dentiste → le dentiste

Le médecin dentiste → le dentiste

Le chirurgien viscéraliste → *le viscéraliste*

L'ellipse est un phénomène linguistique très fréquent dans le vocabulaire médical. Il concerne également les noms d'humains de la médecine les plus employés. Cela ne semble pas poser de problèmes de compréhension chez les locuteurs, c'est-à-dire que le locuteur ne se trouve pas, à chaque fois, dans une situation d'incompréhension qui l'oblige à rappeler les substantifs tronqués.

L'ellipse s'intègre parfaitement dans le langage de la médecine, d'une façon autonome. Les noms comme *viscéraliste* ne nécessitent même pas de rappeler les noms composés qui sont à l'origine de leur formation. Leur forte fréquence dans la langue fait en sorte que le rappel du substantif est considéré comme une répétition.

2. Propriétés sémantiques

Nous traitons les humains du domaine médical du point de vue de leur emploi prédicatif ou argumental. Nous entamons ensuite l'étude des noms propres ainsi que celle des humains collectifs.

2.1. Humains prédicatifs

Les humains, de façon générale et ceux utilisés dans la langue de la médecine en particulier, ont deux fonctions distinctes dans la phrase. Ils peuvent fonctionner en tant que prédicats et en tant qu'arguments.

2.1.1. La prédicativité des noms d'humains

Nous nous penchons ici sur la question de la prédicativité des noms d'humains dans le vocabulaire médical. En linguistique, la description des noms d'humains est particulière. Ils peuvent fonctionner soit comme des prédicats soit comme des arguments. C'est la raison pour laquelle nous leur avons consacré un dictionnaire à part. Autrement dit, les noms d'humains ont un double statut (prédicatif et argumental). Ils peuvent fonctionner comme prédicats parce qu'ils portent en eux-mêmes la notion de prédicativité, avec le type sémantique *action*. Les noms d'humains de la langue de la médecine désignent notamment leurs professions. Ils peuvent fonctionner aussi comme arguments parce qu'ils désignent des entités, en l'occurrence les humains.

Si nous regardons d'une façon superficielle les noms d'humains en médecine, nous considérons qu'ils n'ont rien de particulier ; un médecin est une personne qui pratique la médecine, un chirurgien est une personne qui pratique la chirurgie et un malade est une personne atteinte d'une maladie quelconque. Or, du point de vue linguistique, notamment dans le cadre de la phrase, les noms d'humain ne sont pas homogènes. Ils peuvent avoir une fonction prédicative ou une fonction argumentale. Dans les exemples suivants :

Il est homéopathe

Il est kinésithérapeute

Les noms comme *homéopathe* et *kinésithérapeute* ont une fonction prédicative, tandis que dans :

Le chirurgien ampute Paul d'une jambe

Le pharmacien prépare les médicaments

Les noms d'humain *chirurgien* et *pharmacien* fonctionnent comme des arguments. Ce double statut nécessite de voir en détail la fonction prédicative et la fonction argumentale des noms d'humains.

2.1.2. Fonction prédicative

Les substantifs humains peuvent fonctionner comme des prédicats (G. Gross 1995). Ce postulat peut s'adapter à la langue de la médecine puisqu'elle fonctionne de la même manière que la langue générale (P. Lerat 1995).

Les noms comme *médecin*, *infirmier* ou *sage femme* sont tellement fréquents dans le langage médical qu'on se sent contraint de les analyser et de voir dans quelles situations ils fonctionnent en tant que prédicats.

L'emploi strictement prédicatif des noms d'humains dépend de leur syntaxe. Les humains prédicatifs apparaissent dans les structures suivantes :

Paul est médecin

Paul est le médecin de la famille

Dans ces énoncés où le nom médecin fonctionne comme prédicat, on évoque le métier que *Paul* exerce en position attribut. L'introduction de verbes comme *consulter* ou *opérer* change la donne. Dans ce cas, les verbes appropriés occupent la place de prédicat et les substantifs deviennent des arguments.

En général, les humains médicaux sont relatifs aux métiers qu'ils exercent. Ainsi on peut paraphraser en introduisant les verbes appropriés *pratiquer* et *exercer* :

Paul pratique la médecine

Paul exerce en tant que infirmier

Nous développerons cela dans la fonction argumentale.

2.1.3. Fonction argumentale

Nous venons de voir que les noms d'humains médicaux peuvent fonctionner comme des prédicats, mais les structures dans lesquelles ils apparaissent ne sont pas nombreuses. Ils peuvent fonctionner aussi comme des arguments.

La fonction argumentale des noms d'humains de la langue de la médecine est très fréquente par rapport à la fonction prédicative. Les noms d'humains apparaissent plus souvent sous la fonction argumentale, d'où la variété des structures phrastiques dans lesquelles les noms d'humains figure en tant qu'arguments. Par exemple :

La structure unaire :

L'aide-soignant exerce

Le bronchique tousse

La structure binaire :

Le neurologue consulte les patients

Le rhumatologue traite Paul

La structure ternaire :

Le médecin prescrit des médicaments à Paul

Le chirurgien opère Paul au genou

Etc.

Nous remarquons que les noms d'humains ci-dessus fonctionnent comme des arguments à cause de l'introduction de verbes prédicatifs qui changent le statut syntaxique des phrases. Les substantifs humains fonctionnent ici comme des arguments. Dans ce cas, nous attribuons à chaque nom d'humain un trait sémantique afin de résoudre les problèmes d'ambiguïté sémantique des énoncés. Ce besoin est majeur parce que les discours spécialisés cherchent toujours la clarté et la précision.

2.2. Les humains argumentaux

Il existe une catégorie d'humains qui fonctionnent toujours comme des arguments. Il s'agit des noms propres comme *Benjamin* et *Bernard*. Ces derniers peuvent avoir des fonctionnements différents en langue. Ils peuvent être attribués à des locatifs, à des objets ou à des phénomènes par éponymie.

2.2.1. Fonctionnement global

L'une des caractéristiques des noms d'humains argumentaux est leur utilité dans la formation des noms dits élémentaires comme les locatifs (*Ambroise paré*, *Robert-Debré*, *Saint-Vincent-de-Paul*, *Avicenne*, *Mustapha Pacha*, *Nadir Mohamed*, *Pasteur*, etc.). Nous montrons ici le fonctionnement linguistique des noms d'humains qui désignent non pas les humains mais les locatifs, par éponymie. Cela a un impact sur l'analyse prédicative.

On s'accorde à dire que les noms d'humains médicaux s'accordent avec les verbes de type :

Consulter, (*Avicenne consultait des patients*)

Traiter, (*Avicenne traitait les maladies infectieuses*)

Diagnostiquer, (*Avicenne diagnostiquait les maladies contagieuses*)

Pratiquer (*Ambroise paré pratiquait la chirurgie*)

Soigner (*Ambroise paré avait soigné le Roi*)

Néanmoins, les verbes cités dans les exemples ci-dessus se combinent aussi avec des locatifs qui portent les noms des humains. On peut ainsi dire :

Cet hôpital soigne une centaine de patient par jour

Cette clinique traite les patients atteints du sida

Cet hôpital diagnostique beaucoup de maladies rares

Cette clinique pratique la chirurgie esthétique

Cet hôpital s'occupe des cancéreux

Cependant, les verbes qui se combinent avec les noms d'humains ne s'accordent pas tous avec les locatifs comme *rencontrer*, *parler*, etc.

**Cet hôpital parle à ses patients*

**Cette clinique rencontre les patients*

2.2.2. Fonctionnement particulier

D'autres noms d'humains ont un fonctionnement particulier. En médecine, il est fréquent de dénommer de nouvelles maladies, de nouveaux signes et symptômes ou de nouvelles techniques de soins par le nom la personne qui les découvre. Le phénomène de l'éponymie est très connu dans le vocabulaire médical. Par exemple :

(1) *La maladie de berger (Jean Berger)*

(2) *Syndrome de Bernard (Claude Bernard)*

(3) *Le signe d'Abadie (Jean Marie Charles Abadie)*

(4) *Tiges d'Auer (Jean Auer)*

(5) *Anneaux de Balbiani (Edouard- Gérard Balbiani)*

(6) *Anneaux de Cabot (Richard Cabot)*

Nous constatons que les noms qui sont formés à partir des noms d'humains fonctionnent soit comme des prédicats, soit comme des arguments. C'est le cas des exemples (1), (2) et (3) qui varient selon la construction des phrases dans lesquelles ils apparaissent et selon la position qu'ils y occupent. Cette variation vient de leur nature sémantique puisqu'ils désignent des noms de maladies.

Pour les exemples (4), (5) et (6), ils fonctionnent en tant qu'arguments dans tous les cas, même s'ils dérivent des noms d'humains comme les trois premiers exemples, parce qu'ils désignent des objets ou des parties du corps.

2.2.3. Les dérivés des noms propres

Nous montrons ici une autre spécificité des noms d'humains dans la formation linguistique, notamment dans le vocabulaire de la médecine. Il s'agit ici de la formation des noms de maladies par exemple à partir des noms d'humains qui ont eu la même fonction, mais sous d'autres formes.

<i>Nom propre</i>	<i>Dérivé</i>
<i>Parkinson</i>	<i>parkinsonisme</i>
<i>Parkinson</i>	<i>parkinsonien</i>
<i>Pasteur</i>	<i>pasteuriser</i>
<i>Pasteur</i>	<i>pasteurisé</i>
<i>Alzheimer</i>	<i>Alzheimer</i>

Nous remarquons que le nom d'humain *parkinson* peut former, par dérivation, le nom de la maladie de parkinson (*parkinsonisme*) et dans un autre cas, le nom de la personne atteinte de la maladie du parkinson (*parkinsonien*).

Pasteur est un nom d'humain qui génère le verbe *pasteuriser* ainsi que l'adjectif *pasteurisé*. Le nom *Pasteur* fonctionne toujours comme un argument et jamais comme un prédicat. Par contre, le verbe et l'adjectif, eux, peuvent fonctionner comme des prédicats.

L'exemple du nom *Alzheimer* est un autre cas de figure puisque, ici, il désigne à la fois le nom d'humain et le nom de la maladie.

2.3. Humains collectifs

Notre analyse des noms d'humains nécessite d'étudier tous les noms désignant des humains, y compris les humains collectifs. Mais les noms qui désignent les humains collectifs

ne sont pas toujours clairs. Certains font référence à la fois aux humains et aux locaux occupés par ces derniers.

2.3.1. Noms non ambigus

Les noms d'humains collectifs non ambigus renvoient sans équivoque aux noms humains dont on parle. Les noms comme *équipe médicale, praticiens hospitaliers, équipe de garde de l'hôpital, internes des hôpitaux, résidents des hôpitaux, personnel médical, personnel d'hôpital, personnel sanitaire, personnel de santé auxiliaire, personnel administratif du service de santé ou personnel bénévole de l'hôpital* désignent des groupes de personnes travaillant comme médecins et employés d'hôpitaux. En aucun cas ces noms renvoient aux autres entités, comme dans le cas de la métonymie où les mêmes noms peuvent renvoyer aux locatifs et aux groupes d'humains occupant ces locaux, par exemple, *hôpital* dans :

L'équipe médicale fait grève

L'hôpital fait grève

La signification est identique dans les deux exemples qu'on considère comme une paraphrase. La seule et unique distinction est la transparence dans *l'équipe médicale fait grève* et l'ambiguïté dans *l'hôpital fait grève* parce que le nom *hôpital* désigne le lieu et les humains occupant ce lieu.

L'analyse linguistique des noms d'humain non ambigus nécessite d'apporter des précisions sur leurs traits syntactico-sémantiques afin d'en réduire l'ambiguïté.

2.3.2. Noms ambigus

Les locatif et les humains sont bel et bien différents puisqu'ils appartiennent à deux classes différentes. Néanmoins, ils partagent des prédicats appropriés (T. Massoussi 2009 : 127).

Les praticiens de ce dispensaire sont en colère

Ce dispensaire est en colère

Les praticiens de ce dispensaire dispensent les soins le dimanche

Ce dispensaire dispense les soins le dimanche

Etant donné que les arguments de la classe (humain) s'accordent avec les prédicats appropriés de la classe (humains collectifs) occupant ces locaux (occupant/occupé), les locatifs endossent le sens des occupants, c'est-à-dire des humains. Ils partagent le même verbe support *dispenser* qui, en premier lieu, s'accorde avec les médecins. Cela montre le rapprochement entre ces deux classes (humains collectifs et locatifs), par le mécanisme de la métonymie, qui sont à la base différentes puisqu'elles appartiennent à deux taxonomies différentes.

2.3.3. Autres métonymies

Il existe d'autres types de métonymies dans le vocabulaire médical, non pas entre l'occupé et l'occupant, c'est-à-dire entre humains et locaux, mais entre les institutions, les locatifs et les humains (comme *service de pédiatrie*, *service de gynécologie*, etc.), et l'institution et les humains (comme *le SAMU*).

En premier lieu, *service de pédiatrie* et *service de gynécologie* sont considérés comme des activités que l'on doit assurer dans un établissement de santé ou des travaux professionnels qu'on exerce dans cet établissement.

Les exemples suivant montrent que le même nom renvoie tantôt à l'institution, tantôt au lieu et tantôt au personnel qui exerce dans cette institution.

(1) *Le service de pédiatrie commence à 9 heures du matin*

(2) *Les enfants malades sont soignés dans le service de pédiatrie*

(3) *Le service de pédiatrie menace de durcir la grève*

L'ambiguïté est apparente dans l'exemple (1) puisque *service de pédiatrie* peut renvoyer à l'institution et aux humains. Dans (2), la préposition *dans* montre clairement le sens du nom *service de pédiatrie* qui désigne un lieu. Et dans (3), le nom *service de pédiatrie* renvoie aux humains parce qu'il s'agit d'une grève, c'est-à-dire d'un mouvement social qui ne peut être effectué que par des humains.

Le SAMU est créé en France en 1972

Le SAMU intervient au secours des victimes d'accident de la route

J'ai vu passé le SAMU

Les exemples sur *SAMU* montrent que ce dernier désigne une institution ou des humains collectifs, et non un lieu.

3. Classifications sémantiques et combinatoires

L'analyse syntactico-sémantique des noms d'humains consiste aussi à voir toutes leurs combinaisons possibles avec les verbes, les adjectifs et les noms qui leur sont appropriés. Il existe des verbes, des adjectifs et des noms appropriés aux noms d'humains génériques, c'est-à-dire ceux de la langue générale, et des verbes, des adjectifs et des noms appropriés aux noms d'humains des domaines spécialisés, par exemple ceux de la médecine dont les noms d'humains sélectionnent ceux qui sont appropriés à chaque classe comme les professions, les fonctions et les malades.

3.1. Profession et fonction

Nous montrons ici comment décrire les noms d'humains du domaine de la médecine, à partir de leur combinatoire avec des verbes, des noms et des adjectifs qui leur sont appropriés. Il s'agit ici des humains qui exercent dans les milieux hospitaliers comme *médecin*, *infirmier*, *aide-soignant*, *sage-femme*, etc.

3.1.1. Noms appropriés

Les noms d'humains du vocabulaire médical admettent des prédicats nominaux.

L'auscultation du médecin

La prescription du médecin

Les soins de l'infirmier

L'opération chirurgicale du chirurgien

Les noms comme *auscultation*, *prescription*, *soins* et *opération chirurgicale* sont des substantifs appartenant au domaine médical. Ils s'appliquent aux noms d'humains de la médecine d'une façon sélective. Cette sélection est imposée aussi par le choix des verbes (nous verrons cela ci-dessous).

La sélection des noms appropriés se fait de façon plus ou moins rigoureuse parce que dans un domaine spécialisé comme la médecine, chaque personne est formée pour exercer et faire les tâches pour lesquelles elle est destinée. Cela se fait d'une façon organisée.

3.1.2. Verbes appropriés

Les verbes appropriés aux noms d'humains exerçant une profession ou une fonction dans le domaine médical sont variés. On a des verbes qui appartiennent à la langue générale comme *consulter*, *admettre*, *voir et suivre*, etc., et des verbes spécialisés comme *soigner*, *cautériser*, *disséquer*, etc. Ces verbes ne s'accordent pas d'une façon aléatoire à tel ou tel nom d'humain. Pour bien illustrer cette partie, nous proposons des exemples de phrases à prédicats dyadiques, c'est-à-dire, ceux qui impliquent deux arguments. Dans ce cas, on utilise les verbes transitifs, c'est-à-dire ceux qui impliquent un complément d'objet.

Le médecin cautérise la plaie
Le chirurgien opère le malade
L'anesthésiste anesthésie le patient
Le médecin stérilise le scalpel
Le patient consulte un médecin
Etc.

Le choix de ces verbes spécialisés implique souvent un argument en position de complément (X1). Ce dernier appartient souvent à la langue de la médecine. Parfois, les noms d'humains médicaux s'accordent avec des verbes de la langue générale, et dans ce cas, ce sont les arguments du domaine de la médecine qui donnent un sens spécialisé aux verbes.

Le médecin suit Paul
Paul a vu son médecin

Le médecin consulte son patient

3.1.3. Adjectifs appropriés

Un médecin peut être *réputé*, un infirmier *bon*, une sage-femme *excellente* et un chirurgien *compétent*. Les classes d'humains du vocabulaire spécialisé, notamment du vocabulaire médical, sélectionnent aussi des adjectifs appropriés. Ce sont des adjectifs relationnels. Ces adjectifs peuvent apparaître en position épithète avec un humain.

Un médecin spécialiste

Un médecin consultant

Un chirurgien hospitalier

Un médecin interne

Qui dit appropriation dit collocation. Dans le cas des collocations, on tend vers le figement avec des degrés moindres. Certaines formes adjectivales sont figées, c'est-à-dire qu'elles ne sont pas des séquences compositionnelles dites libres comme *sage-femme* et *aide-soignant*, etc.

Nous séparons les noms de professions et les noms de fonctions parce qu'ils se distinguent, non seulement dans le cadre du travail et des missions accomplies par les uns et les autres, étant régies par un système de hiérarchie similaire au domaine militaire, mais aussi du point de vue de la syntaxe. C'est ce que nous essayons de montrer dans cette partie avec les noms, les verbes et les adjectifs appropriés aux noms d'humains qui désignent des professions et des fonctions.

Nous observons dans notre dictionnaire que les noms d'humains exprimant des fonctions sélectionnent des adjectifs qui peuvent s'appliquer aux noms d'humains qui expriment les professions dans un contexte bien déterminé comme :

Le chef de service est généraliste

Le directeur de l'hôpital est neurologue

Le stagiaire est homéopathe

Ils peuvent aussi admettre les adjectifs qui appartiennent à la langue générale exprimant des tâches et des responsabilités, non seulement dans le domaine médical et le

milieu hospitalier, mais aussi dans tous les domaines possibles, et particulièrement dans l'administration.

Le chef de service est autoritaire

Le maître assistant est compétent

L'interne est impliqué

Le stagiaire est assisté

Le professeur est expérimenté

3.2. Malades

La classe des humains malades fait partie des deux plus grandes classes d'humains médicaux. Cette catégorie de noms d'humains s'accorde avec les noms, les verbes et les adjectifs appropriés dans le contexte médical.

3.2.1. Noms appropriés

Les noms appropriés aux noms de malades sont des substantifs du vocabulaire de la médecine. Il s'agit, en général, des noms de maladies dont les patients sont atteints, les noms de soins et les noms des traitements qu'ils suivent, etc. On rajoute, par exemple, les noms appropriés aux noms de patients pour préciser leurs maladies afin d'éviter les ambiguïtés. Par exemple :

a) noms de maladies

Le diabète du patient

La grippe aviaire du malade

Le cancer du malade

b) noms de soins et de traitements

La chimiothérapie du cancéreux

La trithérapie du sidaïque

Les soins du blessé

Les noms appropriés aux noms d'humains désignant des maladies ou des soins ne s'appliquent qu'aux noms d'humains du vocabulaire médical. Le cas contraire signifie par inférence que l'humain est malade, comme *les soins de Paul*.

3.2.2. Verbes appropriés

Les humains malades admettent un nombre important de verbes appropriés, dans le contexte médical. Ici, nous ne sélectionnons pas tous les verbes appropriés aux humains de façon générale, mais seulement ceux des humains malades.

Nous avons repéré deux types de verbes qui s'accordent avec les humains malades. Il s'agit de ceux relatifs aux soins et à la sensation douleur provoquée par les maladies.

a) les soins

Le cancéreux consulte son médecin

Le sidéen suit son traitement synergique

Le patient subit un examen médical

Le lépreux se fait soigner

b) la sensation de douleur

Le blessé ressent des douleurs

Le grippé souffre des maux de tête

Le patient supporte ses douleurs

Dans les exemples ci-dessus, les verbes appropriés aux humains malades ne sont pas homogènes du point de vue syntaxique. Certains sont des verbes prédicatifs comme *soigner* et *supporter*, tandis que d'autres sont des verbes supports comme *ressentir* et *souffrir*. Nous

constatons qu'ils appartiennent généralement à la langue générale, contrairement à ceux qui s'accordent avec les noms des praticiens hospitaliers.

3.2.3. Adjectifs appropriés

Les noms d'humains malades se combinent avec des adjectifs spécialisés et des adjectifs de la langue générale. De ce point de vue, la syntaxe des noms d'humains est plus contraignante lorsqu'il s'agit de la combinaison de noms de malades avec les adjectifs appropriés. Par contre, dans le cas de la combinaison des noms d'humains malades avec des adjectifs de la langue générale, le choix est plus conséquent.

La combinaison des noms d'humains malades avec des adjectifs spécialisés :

Le blessé est hémorragique

Le cancéreux est souffrant

Le sidéen est incurable

Le grippé est fiévreux

La combinaison des noms d'humains malades avec les adjectifs de la langue générale :

Le diabétique est vigilant

Le cancéreux est stressé

Le sidéen est fragile

Le grippé est fatigué

Les adjectifs s'accordant aux humains peuvent être des locutions adjectivales comme :

Le diabétique est bien portant

Le cancéreux est en mauvaise santé

Le sidéen est en bonne santé

3.3. Autres noms d'humains

Nous ne pouvons pas négliger tous les autres noms relatifs aux humains dans l'analyse du vocabulaire spécialisé, parce qu'ils peuvent naturellement intégrer ce vocabulaire dans le discours. L'utilisation de ces noms dans le vocabulaire médical leur donne un caractère spécialisé. Autrement dit, l'association des noms de la langue générale aux noms, verbes et adjectifs spécialisés les associent systématiquement au vocabulaire médical.

3.3.1. Noms appropriés

Les noms d'humains génériques s'accordent avec les noms appropriés à la médecine, et ces derniers les intègrent dans le vocabulaire spécialisé. C'est par exemple le cas avec les noms spécialisés comme les noms de maladie : *leucémie*, *mucormyose*, *angine* et *méningite*, etc.

La leucémie de Paul

La mucormyose de Max

L'angine de Léa

La méningite de Luc

Les noms d'humains malades de la langue générale s'accordent également avec des noms appropriés comme les examens médicaux tels que *biopsie d'exérèse*, *échographie*, *test auditif* et *sérodiagnostic de la syphilis* :

La biopsie d'exérèse de Paul

L'échographie de Léa

Le test auditif de Max

Le sérodiagnostic de la syphilis de Luc

Nous avons constaté que l'utilisation des autres noms d'humains qui appartiennent à la langue générale avec des noms appropriés, en l'occurrence les noms de maladies et les noms d'examens médicaux, les insère dans un contexte spécialisé, notamment la médecine. Cette

spécificité ne s'applique pas seulement aux noms appropriés, mais aussi aux verbes et aux adjectifs appropriés.

3.3.2. Verbes appropriés

Nous traitons ici les verbes appropriés aux noms d'humains utilisés en médecine. Il s'agit des noms d'humains de la langue générale utilisés fréquemment dans le vocabulaire médical avec des verbes spécialisés.

La combinaison des noms d'humains avec les verbes appropriés de la langue générale tient compte uniquement de la congruence, c'est-à-dire, de la compatibilité de tel ou tel verbe avec un nom d'humain comme :

Paul lit un roman

Paul déguste une pizza

**Paul déguste un roman*

**Paul lit une pizza*

Nous constatons que dans les deux premiers exemples, les verbes *lire* et *déguster* sont compatibles avec les noms d'humains en position de sujet, et *roman* et *pizza*, en position de complément. Par contre, dans les deux derniers, *déguster un roman* (sauf dans la métaphore), et *lire une pizza*, les verbes ne sont pas compatibles avec les arguments en position de complément.

Pour les verbes appropriés spécialisés, la sélection sur l'axe syntagmatique et l'axe paradigmatique est contraignante, puisque les humains ont des tâches spécifiques à accomplir. La particularité des verbes ici, c'est la définition du statut des humains. Par exemple, dans :

Paul soigne

Paul opère

Paul tousse

Paul gémit

Les verbes *soigner* et *opérer* montrent le statut de *Paul*, puisque ces verbes sont attribués uniquement aux humains habilités à donner des soins et à opérer, en l'occurrence, les professionnels hospitaliers (médecin, infirmier, chirurgien, etc.)

Certains verbes précisent la profession exacte de l'humain comme *opérer* qui s'accorde uniquement avec *chirurgien* et exclut les autres comme *infirmier* et *sage-femme*.

3.3.3. Adjectifs appropriés

Les noms d'humains de la langue générale se combinent avec les adjectifs appropriés aux humains comme *fatigué*, *malade*, *courageux*, etc., mais plus on affine les noms humains en sous-classe comme femme, homme, enfant, bébé, père, mère, etc., plus on sélectionne les adjectifs appropriés à chaque sous-classe. On dit plutôt *femme enceinte*, *père travailleur*, etc. cela relève de la collocation.

Puisque nous décrivons la langue de la médecine, nous évoquons les adjectifs appropriés, c'est-à-dire ceux qui s'appliquent aux noms d'humain de la langue de la médecine. Comme pour les noms et les verbes, l'accord n'est pas aléatoire mais sélectif.

Paul est hospitalisé

Paul est malade

Paul est guéri

Paul est dialysé

Les adjectifs *hospitalisé*, *malade*, *guéri* et *dialysé* des exemples ci-dessus sont des termes du vocabulaire médical, donc ils définissent la classe sémantique de *Paul*, (en l'occurrence malade).

Conclusion

A travers cette description des noms d'humains dans le vocabulaire médical, nous avons montré la variation des noms d'humains du point de vue syntactico-sémantique, c'est-à-dire les humains prédicatifs et les humains argumentaux. Le statut double des noms d'humains fait qu'ils sont particuliers. Ils fonctionnent comme prédicats parce qu'ils portent en eux la notion de prédicativité en raison de leurs rôles sémantiques, et ils fonctionnent comme arguments parce qu'ils désignent les entités du monde. La distinction des classes sémantiques des humains médicaux (professions / fonctions, malades et autres) présuppose une syntaxe propre à chaque catégorie de noms d'humains. Cela est résumé par la notion de noms, verbes et adjectifs appropriés à chacune d'entre elles.

CHAPITRE 7

DICTIONNAIRE POLYFONCTIONNEL

Introduction

Le dictionnaire a toujours été un support pédagogique par excellence. Il en existe trois sortes : le dictionnaire papier, le dictionnaire électronique et le dictionnaire papier numérisé. Nous prenons en compte ici uniquement les deux premières versions.

La différence entre les dictionnaires papiers et les dictionnaires électroniques ne relève pas seulement de la différence des supports et du volume. Les dictionnaires papiers sont destinés à expliquer le bon fonctionnement de la langue ainsi que ses règles. Les dictionnaires électroniques sont conçus pour le traitement automatique. Pour mieux gérer la microstructure des dictionnaires, les lexicographes et les informaticiens ont recours aux bases de données. Elles sont utiles pour le stockage et la gestion des données. Pour ce faire, plusieurs outils informatiques et lexicographiques sont nécessaires. Nous les définissons successivement et nous montrons leur utilité.

1. Structuration des données lexicales

Nous évoquons dans cette section les bases de données utilisables pour le stockage des données. Nous montrons ensuite les motivations de notre choix concernant notre corpus. Nous évoquons enfin les tableaux Excel et CSV et les fichiers XML.

1. 1. Bases de données

Nous commençons par les bases de données, ensuite nous montrons les différentes bases de données existantes puis nous justifions notre choix.

1.1.1. Définition des bases de données

Une base de données est considérée comme une entité informatique qui permet de stocker des données d'une façon structurée. Elle peut être utilisée et exploitée par une ou plusieurs personnes. Mettre des informations dans une base de données implique de les mettre à la disposition des utilisateurs pour des consultations faciles, c'est-à-dire facilement accessibles (H. Korth & F. Silberschatz 1988).

La notion de base de données est généralement associée à celle de réseau pour permettre à plusieurs utilisateurs de la compléter et de la consulter en même temps. Le réseau permet ici de partager les données.

Pour pouvoir contrôler la base de données et les utilisateurs qui l'exploitent, le gestionnaire de la base doit mettre en place un système de gestion de la base de données.

1.1.2. Les différentes bases de données

Il existe trois grands paradigmes pour le stockage des données :

- Les bases de données relationnelles, par exemple celles qui utilisent la technologie SQL (MySQL, PostgreSQL, etc.). Elles permettent de gérer un très grand volume de

données avec de bonnes performances et sont à l'heure actuelle les plus utilisées. Elles fonctionnent dans une structure client-serveur : les données sont stockées de manière opaque par un serveur, lequel les renvoie en fonction des requêtes qui lui sont transmises.

- Les systèmes de bases de données embarqués (SQLite, BSDDB, etc.). Ces bases de données fonctionnent de manière autonome (elles ne sont pas interrogées par l'intermédiaire d'un serveur). Elles permettent de stocker un nombre de données conséquent et sont extrêmement efficaces en termes de réponses aux requêtes.

- Les bases de données fondées sur l'arborescence XML. Elles se sont beaucoup développées ces dernières années. Elles constituent des alternatives crédibles aux bases de données relationnelles : si elles ne sont pas aussi performantes en termes de requêtes comme en termes de volume, elles sont généralement plus simples à exploiter, d'autant plus que les limites techniques que nous avons évoquées ne se font sentir que sur un volume de données très important.

Par ailleurs, certains logiciels, par exemple Microsoft Access ou Open Office base proposent à leurs utilisateurs des interfaces facilitant l'utilisation de ces technologies. (En l'occurrence, MS Access et DD base utilisent tous deux les technologies SQL).

1.1.3. Notre choix : PHP SQL

Pour exploiter nos descriptions de la langue de la médecine, nous avons utilisé le langage de programmation PHP. Nous avons choisi le système SQL. Ce choix est motivé par la possibilité d'intégrer les bases écrites en SQL.

Pour P. Rigaux (2005) : « le recours à MySQL permet de masquer les détails complexes et fastidieux liés à l'utilisation de fichiers. MySQL gère pour vous les fichiers constituant une base de données, prend en charge les fonctionnalités de protection et de sécurité et fournit un ensemble d'interfaces de programmation (dont une avec PHP) facilitant l'accès aux données ». C'est cette facilitation d'accès aux données qui nous a poussés à utiliser PHP MySQL.

SQL offre d'autres avantages intéressants comme :

- Il est très rapide et performant
- Il dispose d'un système de sécurité qui permet de gérer et de contrôler les personnes et les machines qui y accèdent
- Il fonctionne sous Windows et sous Linux
- Il est en libre accès

1.2. Tableaux CSV et Excel

La première étape dans la réalisation d'une base de données est l'utilisation des tableaux Excel qu'il faut, par la suite, convertir en format CSV. Il nous semble important de commencer par la définition des fichiers Excel et CSV, puis d'expliquer comment convertir les fichiers Excel en fichier CSV. Nous évoquerons enfin l'exploitation des fichiers CSV.

1.2.1. Définition des fichiers Excel et CSV

Excel est un tableur développé par Microsoft en 1980. Il intègre des fonctions pour le calcul numérique, la représentation graphique et l'analyse des données. Les versions les plus récentes peuvent gérer des documents de tailles considérables (D. Frye 2007).

Excel présente des données sous forme de tableaux structurés en lignes et colonnes. Dans le même fichier (dit table), il est possible de créer de nouvelles feuilles ayant des relations, c'est-à-dire que l'on peut importer des données d'une feuille à une autre dans le même fichier. Il propose des fonctions qui permettent de ranger, de structurer et d'organiser les données de plusieurs façons. D'autres fonctions permettent de supprimer les doublons, d'imprimer les feuilles, de filtrer les données, et d'analyser les données, etc. (D. Frye 2007).

Le format CSV est un format informatique. Il représente les données tabulaires sous forme de valeurs séparées par des virgules. Le CSV est un fichier texte où chaque valeur séparée par une virgule représente une valeur d'une cellule du tableau.

Le format CSV présente, néanmoins, quelques difficultés dans son exploitation. Elles sont liées, notamment, à l'origine des documents qu'on traite. Par exemple, dans un document anglais. Le format CSV utilise les virgules pour la séparation des colonnes et les points pour la séparation des décimales, tandis que, pour les fichiers français, les colonnes sont séparées

par les points virgules et les décimales par les points. Cette variation s'opère en fonction de l'origine du fichier.

1.2.2. Conversion des fichiers Excel vers CSV

Les formats XLS et CSV sont deux formats différents. Pour faire une interface d'interrogation avec PHP MySQL, il est nécessaire de convertir nos dictionnaires électroniques dans des tableaux Excel (format XLS) ou en format CSV, afin de pouvoir les importer dans la base de données SQL.

Nous montrons dans le tableau ci-dessous un exemple de tableau en format XLS converti en format CSV. Dans ce cas, les données seront présentées, non pas dans un tableau, mais en format texte séparées par des virgules.

Fichier XLS :

<i>Nom élémentaire</i>	<i>Genre</i>	<i>Morphologie</i>	<i>Classe sémantique</i>
<i>virus de l'hépatite A</i>	<i>masculin</i>	<i>Nsimple</i>	<i>virus</i>
<i>main</i>	<i>féminin</i>	<i>Nsimple</i>	<i>partie du corps</i>
<i>sang</i>	<i>masculin</i>	<i>Nsimple</i>	<i>liquide organique</i>
<i>cellules de la moelle osseuse</i>	<i>féminin pluriel</i>	<i>Ncomposé de type [N de la NA]</i>	<i>cellule</i>

Fichier CSV correspondant :

Nom élémentaire, Genre, Morphologie, Classe sémantique

VIRUS de l'hépatite A, Masculin, Nsimple, Virus

Main, Féminin, Nsimple, Partie du corps

Sang, Masculin, Nsimple, Liquide organique

Cellules de la moelle osseuse, Féminin pluriel, Ncomposé [type : N de la NA], Cellule

1.2.3. Exploitation des fichiers CSV

Un fichier CSV peut s'ouvrir avec un tableur ou avec un éditeur de texte comme Bloc note et Notepad++. Il est essentiellement utilisé avec les fichiers Excel. L'ouverture des fichiers Excel sous format CSV est essentielle pour les importer dans les tables SQL.

Le format CSV offre la possibilité de transmettre des données et des informations dans plusieurs applications dans un format simple et lisible.

1.3. Fichiers XML

XML est un langage de description et de structuration de l'information. Il est dérivé de SGML. L'objectif du XML était de permettre à SGML d'être traité et diffusé sur le web, de la même manière que HTML (A. Michard 2000).

1.3.1. Définition des fichiers XML

XML (Extensible Markup Language) est un langage de description, d'échange et de diffusion d'information qui fonctionne avec un système de balisage extensible. Il est caractérisé par l'utilisation des chevrons (<>). Il sert à la diffusion automatique d'information sur le web comme les textes (A. Michard 2001).

Le XML ressemble au langage HTML (Hypertext Markup Language) (est un système de balises). Néanmoins, le XML présente beaucoup d'avantages par rapport à HTML parce qu'il propose un nombre illimité de balises. Autrement dit, il permet de créer de nouvelles balises selon les besoins de l'utilisateur. Cela sous-entend qu'il permet de mieux traiter et diffuser les informations textuelles.

XML est reconnaissable grâce aux balises ouvrante (< >) et aux balises fermantes (</ >). Chaque élément du XML est composé d'une balise ouvrante, d'un contenu et d'une balise fermante (A. Michard 2001).

Exemple :

```
<balise ouvrante>  
contenu
```


</balise fermante>

1.3.2. Exploitation

L'exploitation du XML nécessite, tout d'abord, de connaître ses règles de syntaxe. Il ressemble beaucoup au XHTML.

Puisque XML est un langage de balisage, il faut toujours commencer par une balise d'ouverture (*<balise ouvrante>*) et terminer par une balise de fermeture (*</balise fermante>*).

Chaque balise peut comporter ou non du contenu texte. Dans le cas où elle comporte un contenu, ce dernier doit être inclus entre la balise ouvrante et la balise fermante. Si la balise ne comporte pas de contenu texte, on utilise une forme raccourci (*<balise> </balise>*).

Nous donnons comme exemple une simple lettre au format XML dans le but de le diffuser sur internet.

```
<lettre>
<en-tête>
<nom-emetteur> nom, prénom </nom>
<adresse>
<numero>numéro</numero>
<rue>nom de la rue</rue>
<ville>ville</ville><cp>code postal</cp>
</adresse>
</en-tête>
<date>date</date>
<destinataire>
<nom-destinataire>nom</nom-destinataire>
<adresse>
<numero>numero</numero/>
<rue>nom de la rue</rue>
<ville>ville</ville><cp>code postal</cp>
</adresse>
</destinataire>
<objet>objet</objet>
<texte>texte</texte>
<signature>signature<signature/>
</lettre>
```

On exploite le langage XML pour élaborer des bases de données. Pour les rendre dynamique, on utilise XQuery qui permet d'accéder aux informations structurées en fonction de l'arborescence. L'outil qui permet de naviguer dans l'arborescence est XPATH. Le système de gestion de bases de données que nous utilisons est BaseX :

The screenshot shows the BaseX 6.7.1 interface. The top window displays the XML text:

```
<chat>
  <definition>
    <archisememe>felin</archisememe>
    <seme1>petite taille</seme1>
    <seme2>domestique</seme2>
    <synonyme>chaton</synonyme>
    <traduction>cat</traduction>
  </definition>
</chat>
```

Below the text, a tree diagram visualizes the structure. The root node is 'doc("exercicebasx.xml")', which branches into 'chat'. 'chat' branches into 'definition'. 'definition' branches into five nodes: 'archisememe', 'seme1', 'seme2', 'synonyme', and 'traduction'. Each of these nodes further branches into a single leaf node: 'felin', 'petite taille', 'domestique', 'chaton', and 'cat' respectively.

1.3.3. Les inconvénients du MXL

Il nous semble difficile de parler des inconvénients du XML parce qu'il présente beaucoup d'avantages, notamment en terme de transmission de l'information et de lisibilité. Il suffit, dans tous les cas, de respecter les normes syntaxiques, ce qui est relativement faciles. Le respect de la norme 3WC (World Wide Web Consortium) est de rigueur, premièrement,

afin de garantir l'accessibilité du site web pour les utilisateurs, et deuxièmement, dans le but de le rendre compatible avec les navigateurs web.

Néanmoins, en termes de taille et d'occupation de la mémoire, les fichiers XML sont beaucoup plus volumineux que les autres fichiers binaires. La taille des fichiers a, évidemment, une répercussion sur la transmission des données. Autrement dit, un fichier de grande taille pose beaucoup plus de contraintes et de difficulté qu'un fichier léger. Dans un fichier binaire, on transmet uniquement l'information essentielle ; par contre, dans un fichier XML, l'information est transmise dans une chaîne de caractères (balises) qui prend de la place.

2. Description extensive

Nous évoquons ici la diversité métalinguistique, les sources et le mode de structuration des dictionnaires.

2.1. Diversité métalinguistique et extralinguistique

L'élaboration d'un dictionnaire consiste à prendre des mesures nécessaires notamment en ce qui concerne les informations que comportent les entrées du dictionnaire. Une entrée peut comporter des informations d'ordre synchronique, par exemple le sens d'un mot dans la période actuelle, des informations d'ordre diachronique, comme l'évolution des sens des mots à travers le temps, et des informations d'ordre encyclopédique.

2.1.1. Informations métalinguistiques et diachronie

La définition diachronique d'un mot tient compte de son évolution historique (J. Pruvost 2006). Auparavant, les descriptions des mots se faisaient en diachronie. Chaque mot était étudié depuis sa première attestation. On étudiait ses sens, suivant leurs évolutions, son sens archaïque, s'il en avait, etc. Il s'agissait, tout simplement, de décrire l'évolution d'un système linguistique afin de comprendre les changements qu'il a subis (*ibid*).

Cette tendance diachronique a disparu après l'avènement du structuralisme et de l'étude synchronique de la langue avec Ferdinand de Saussure, mais les dictionnaires diachroniques ont réapparu à la fin du 20^e siècle avec la publication du *Dictionnaire historique de la langue française* (1992), sous la direction d'Alain Rey. On a ainsi assisté au retour de l'intérêt des linguistes et des lexicographes pour les études diachroniques de la langue française. Cela se confirme par la publication d'un autre dictionnaire qui traite l'évolution des mots en français, en l'occurrence, le *Dictionnaire historique de la langue française* (1995) chez Larousse.

2.1.2. Informations métalinguistiques et synchronie

Les définitions et les descriptions des mots peuvent se situer en synchronie (J. Pruvost 2006). Elles porteront dans ce cas sur une période restreinte. Ainsi, les lexicographes décrivent les mots tels qu'ils sont employés à une époque donnée.

Souvent, les dictionnaires de langue, tel que *Le Petit Robert*, donnent les définitions des mots en synchronie, c'est-à-dire en fonction du sens ou des sens employé (s) au cours de la période de l'édition du dictionnaire.

Avec le structuralisme et Ferdinand de Saussure, notamment avec la dichotomie saussurienne « synchronie et diachronie », les linguistes avaient une vision différente dans leurs études linguistiques et surtout leurs descriptions des langues. Ainsi, les descriptions des linguistes synchroniciens se basaient sur les emplois des mots en synchronie, tout en négligeant leurs évolutions historiques (*ibid*). Cette tendance qui consiste à étudier les mots en synchronie a beaucoup influencé les lexicographes, leurs travaux se sont ainsi focalisés sur une dynamique qui porte un grand intérêt à l'étude synchronique de la langue. Le résultat des influences structuralistes devient plus concret avec l'apparition en 1966 du *Dictionnaire du français contemporain* sous la direction de Jean Dubois.

2.1.3. Informations extralinguistiques

Selon la définition du trésor de la langue française, le mot *extralinguistique* désigne tout ce « qui est extérieur à l'acte de communication ». « La réalité extralinguistique dans laquelle baigne le discours, l'entourage général ou particulier que supposent les paroles

prononcées dans chaque cas, la situation, en y comprenant — cas limite — cette situation que crée le discours même au fur et à mesure qu'il se déroule : le contexte » (C. Bally 1952 : 76).

Nous parlerons ici des informations extralinguistiques associées aux définitions des mots dans les dictionnaires, comme le domaine et les informations de nature encyclopédique. Les informations sur le domaine auquel appartient le mot ne sont pas linguistiques, mais demeurent nécessaires pour comprendre le sens. Le verbe *lifter* est ainsi défini dans le *TLFi* :

1. *Lifter*, verbe trans. sports. Frapper une balle, un ballon de façon à leur donner du lift.

2. *Lifter*, verbe trans. Procéder à l'opération du lifting.

Dans la première définition, le domaine auquel le verbe appartient (sport) est précisé. Cela donne plus d'éclairage sur son sens. Certaines définitions comportent aussi le sous-domaine du verbe qui est *tennis*. À l'inverse, dans la deuxième définition, le domaine (chirurgie esthétique) n'est pas mentionné. Le seul indice qui indique qu'il s'agit d'une opération chirurgicale est le mot *opération*.

Des informations de nature encyclopédique sont parfois rajoutées aux définitions des mots dans les dictionnaires encyclopédiques. L'encyclopédie est définie dans le *TLFi* comme l'« ouvrage qui fait le tour de toutes les connaissances humaines ou de tout un domaine de ces connaissances et les expose selon un ordre alphabétique ou thématique ». Par exemple la définition du mot *cœur* comme un « organe central de l'appareil circulatoire ... » est suffisante pour le comprendre, mais dans les dictionnaires encyclopédiques, on ajoute des informations « accessoires » comme *sentiment* et *sentimentalité*.

2.2. Sources

Les sources des dictionnaires sont multiples et variées : elles proviennent à la fois de créations personnelles, de sources lexicographiques et de ressources provenant des corpus.

Cette diversité permet d'alimenter et d'enrichir le dictionnaire de façon à ce qu'il soit exhaustif afin de répondre aux besoins de ses utilisateurs.

Si le nombre d'entrées d'un dictionnaire est déterminé par des raisons économiques, comme le prix du dictionnaire et son volume, le choix des mots est plutôt déterminé par le public visé pour satisfaire ses attentes (G. Matoré 1968). On comprend que chaque dictionnaire se fixe des limites, imposées par l'éditeur.

2.2.1. Sources lexicographiques

Toute entreprise lexicographique nécessite l'utilisation des dictionnaires. Qui dit utilisation des dictionnaires dit utilisation des entrées des dictionnaires selon la macrostructure (nomenclature) et la microstructure (définition des entrées). Les lexicographes utilisent des dictionnaires afin d'élaborer un dictionnaire qui correspond aux normes d'usages.

Dans le Grand Larousse, on parle de « stock lexical ». En se basant sur le Trésor de la langue française, les lexicographes divisent le vocabulaire français en deux sortes de mots, les mots grammaticaux et les mots lexicaux. Les premiers, en nombre limité et qui jouissent d'une fréquence élevée¹⁴, sont nécessaires à la grammaire et aux constructions phrastiques. Les seconds, dont le répertoire est ouvert, sont beaucoup plus nombreux. Leur fréquence varie selon les situations et les domaines¹⁵.

2.2.2. Sources provenant des corpus

Pour élaborer un nouveau dictionnaire ou pour enrichir une nouvelle version d'un dictionnaire, les lexicographes consultent des corpus colossaux pour récupérer les mots de la langue. La mise à jour des dictionnaires est très importante puisque la langue évolue rapidement. Cela impose aux lexicographes d'enregistrer quotidiennement les mots et les sens nouveaux qui font leur entrée dans la langue.

Les corpus utilisés dans l'élaboration et l'enrichissement des dictionnaires sont variés. Les lexicographes utilisent à la fois des corpus généraux, c'est-à-dire qui comportent des

¹⁴ Grand Larousse de la langue française

¹⁵ Par exemple, les termes de la médecine sont utilisés dans le domaine médical plus que dans d'autres domaines.

textes de la langue générale, et des corpus spécialisés. Pour le cas des corpus spécialisés, les lexicographes font appel à des spécialistes de chaque discipline pour donner une définition appropriée et surtout à la portée d'un public non spécialiste¹⁶.

2.2.3. Créations personnelles

Dans cette partie, il nous semble important d'évoquer l'aventure de Paul Robert dont l'entreprise lexicographique a donné naissance à un dictionnaire de langue française de référence, *Le Robert*. Son aventure a commencé par la collecte des termes des agrumes à Orléanville (actuelle ville de Chlef) dans son Algérie natale¹⁷.

Les lexicographes scrutent la presse écrite, les médias, les discours politiques et les spots publicitaires dans le but de repérer les néologismes et les calques. Certainement, les lexicographes ne peuvent pas prendre en compte tous les mots de la langue à cause des raisons d'ordre économique et d'ordre pédagogique (chaque dictionnaire vise un public bien défini). Ainsi certains mots leur échappent. La collection *Le Souffle des mots* a ce titre publié le Dictionnaire insolite des mots oubliés. Elle s'est appuyée sur plusieurs sources comme l'exploitation de la presse, les médias et autres supports pour récupérer, premièrement, les nouveaux mots, deuxièmement, pour montrer que beaucoup de mots qui sont pourtant employés avec des fréquences élevées échappent aux lexicographes.

2.3. Mode de structuration

Chaque type de dictionnaire est structuré d'une façon différente. Nous évoquons, premièrement, la formalisation métalexigraphique des dictionnaires, c'est-à-dire leurs types et leurs méthodes, nous définissons ensuite les dictionnaires coordonnés et non coordonnés ainsi que les dictionnaires généraux et spécialisés.

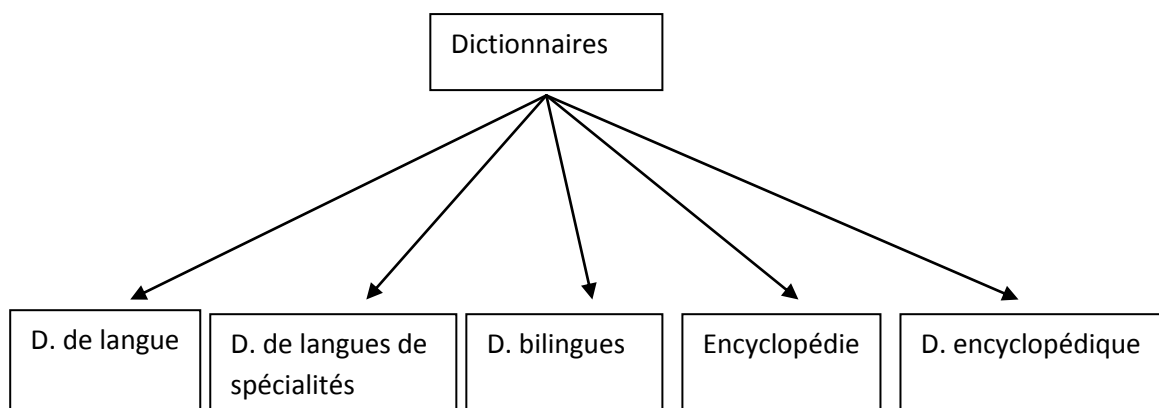
¹⁶ Catalogue Larousse. <http://www.editions-larousse.fr>

¹⁷ Anecdote raconté lors de la 19^e journée des dictionnaires de l'université de Cergy Pontoise par le petit fils de Paul Robert.

2.3.1. Formalisation métalexigraphique

La métalexigraphie est née avec la thèse de Bernard Quemada (1968) sur les dictionnaires du français moderne, 1539-1863. Elle est considérée comme une discipline qui s'intéresse à l'étude des dictionnaires, notamment, du point de vue de leur méthodologie et de leur typologie. Depuis, plusieurs études et publications ont vu le jour (J. Pruvost 2006 : 105). Frank Neveu (2004) la définit dans le *Dictionnaire des sciences du langage* comme « à partir du grec *meta*, ce qui dépasse, ce qui englobe. La métalexigraphie est une discipline dont l'objectif est l'étude des types de dictionnaires de langue et des méthodes qui président à leur construction. Elle ne travaille pas à l'élaboration des dictionnaires, mais fait des dictionnaires, de leur histoire, de leur mode de traitement sémantique du lexique, et des problèmes pratiques résultant du travail lexicographique, son objet de réflexion et de recherche ».

Le principal travail de la métalexigraphie est de classer les dictionnaires par type et de dégager la méthodologie qui correspond à chaque type de dictionnaire. Nous présentons un schéma dans lequel nous faisons le classement des principaux dictionnaires par types et nous définissons la méthodologie de chaque type de dictionnaire.



Un dictionnaire de langue est un ouvrage contenant un ensemble de mots classés par ordre alphabétique et définis avec des exemples et des synonymes.

Un dictionnaire de langue de spécialité est un glossaire qui contient les mots d'une langue spécialisée, définis de la même façon que dans un dictionnaire de langue générale.

Un dictionnaire bilingue est un lexique d'une langue donnée avec un ou plusieurs mots équivalents à chaque mot, dans une autre langue.

Une encyclopédie est un ouvrage qui synthétise toutes les connaissances.

Un dictionnaire encyclopédique est un dictionnaire qui fournit à la fois les définitions des mots et des informations d'ordre encyclopédique.

2.3.2. Dictionnaires coordonnés et non coordonnés

Les dictionnaires coordonnés sont des dictionnaires électroniques destinés au traitement automatique des langues. On parle de deux ou plusieurs dictionnaires coordonnés quand on les élabore en les mettant en corrélation pour des applications de traitement automatique des langues, comme la traduction automatique.

Les dictionnaires coordonnés sont devenus indispensables dans le traitement automatique, vu le nombre important de documents électroniques qui circulent sur le web (J. Sastre Alaiz 2007). Ils sont construits comme tous les dictionnaires. Ils répondent à deux règles de base de tout dictionnaire, notamment, une macrostructure, c'est-à-dire, la description de sa nomenclature ou les entrées du dictionnaire auxquelles correspondent des unités lexicales lemmatisées, et une microstructure, c'est-à-dire que chaque entrée du dictionnaire doit comporter des champs d'informations lexicographiques qui la définissent du point de vue morphologique, syntaxique et sémantique.

Un dictionnaire électronique coordonné montre son utilité pour la traduction automatique, mais il peut être aussi utile « pour des fins sociales comme de faciliter la tâche aux utilisateurs qui décideraient d'entreprendre une recherche d'emploi sur internet ». (*Ibid.*)

L'utilisation classique dans la traduction automatique ne repose pas sur des dictionnaires coordonnés, mais sur des dictionnaires bilingues ou multilingues. Par exemple, la traduction automatique avec *Reverso*, *Google translate* et *Bing* repose sur des ressources terminologiques en plusieurs langues.

2.3.3. Dictionnaires généraux et spécialisés

Il nous semble d'abord, nécessaire d'évoquer la différence fondamentale entre les dictionnaires généraux et les dictionnaires spécialisés. La réponse nous vient de leurs appellations opposées (général # spécialisé). Les dictionnaires généraux traitent la langue générale, tandis que les dictionnaires spécialisés traitent les langues spécialisées comme la langue de la médecine, de l'informatique, de l'aéronautique, etc. Ce qui est proposé aux lecteurs et aux usagers des dictionnaires est différent selon le type de dictionnaire.

Dans un dictionnaire général, le choix des mots se fait de façon intuitive ou sur la fréquence d'emploi (J. Pruvost 2006). Dans les dictionnaires généraux, on trouve les mots de la langue générale. A chaque entrée du dictionnaire sont associées une définition, des synonymes, des antonymes, des exemples d'emplois, et parfois des étymologies. En termes de volume, les dictionnaires généraux contiennent en moyenne 60 000 mots.

Les dictionnaires spécialisés traitent un type particulier de mots de la langue. J. Pruvost (2006) les classe en deux catégories : soit il s'agit de dictionnaires de la langue générale ayant des caractéristiques formelles ou sémantiques (dictionnaires d'orthographe, d'homonymes, des synonymes, d'antonymes, etc.), soit il s'agit de dictionnaires qui traitent les mots de domaines et de spécialités spécifiques.

Un dictionnaire spécialisé est une liste de termes d'un domaine particulier définis de la même façon que les termes du dictionnaire général, sauf qu'il est destiné en priorité à des usagers ciblés. Ainsi le dictionnaire de la médecine vise en priorité les étudiants en médecine, les praticiens dans le domaine de la santé et éventuellement les traducteurs.

Une spécificité des dictionnaires spécialisés réside dans la précision et l'univocité des définitions. A ces dernières sont ajoutées, généralement, des références et des textes d'illustration.

3. Exploitation lexicographique

L'intérêt que portent les linguistes et les lexicographes à la langue n'a fait que contribuer à augmenter le nombre de dictionnaires. Cet outil indispensable pour l'apprentissage des langues existe d'aujourd'hui par milliers. Choisir un dictionnaire est très complexe. Les questions commencent, généralement, par « pour quelle utilisation ? », ou

« comment est-il structuré ? » ou bien « comment présente-t-il les données ? ». Nous essayons de répondre à ces questions afin d'éclairer les zones d'ombre qui entourent les dictionnaires.

3.1. Utilisations diverses

La distinction entre un dictionnaire monolingue, un dictionnaire historique et un dictionnaire plurilingue réside dans le contenu. Qui dit contenus différents dit utilisations et besoins différents.

3.1.1. Dictionnaire monolingue

Un dictionnaire monolingue est un dictionnaire qui traite une seule langue. Il est destiné aux usagers des dictionnaires qui maîtrisent déjà la langue dans laquelle ils cherchent les définitions ou les explications de mots difficiles ou de notions abstraites (B. Quemada 1968). Il est élaboré pour décrire le lexique d'une langue donnée dans le but de donner le sens de chaque mot, de son fonctionnement et de ses emplois. Le dictionnaire monolingue traite aussi les règles d'usage à respecter et les emplois de certains mots.

S'il est différent d'un dictionnaire à l'autre, le lexique du dictionnaire monolingue est, généralement riche ; Le petit Larousse illustré (2012) contient 62 000 mots de la langue générale et des langues spécialisées, Le petit Robert (2010) contient 60 000 mots de la langue générale et des langues spécialisées.

Le dictionnaire à usage monolingue s'oppose au dictionnaire bilingue et au dictionnaire encyclopédique. Le dictionnaire monolingue est conçu pour la connaissance des sens et des usages des mots d'une seule langue. Les explications, les exemples et les définitions d'un mot sont proposés dans la même langue que celle de l'entrée (J. Pruvost 2006). Le deuxième est conçu pour traduire les mots ou donner leurs équivalents dans la langue cible. Enfin l'encyclopédie est conçue pour la connaissance du monde. Le dictionnaire monolingue est composé de toutes sortes de mots, autrement dit, il contient toutes les catégories du discours. Les mots sont classés par ordre alphabétique afin de faciliter leur recherche et leur consultation. Ils sont définis, soit par regroupement, c'est-à-dire donner une seule entrée à plusieurs sens, soit par dégroupement, dans ce cas, chaque sens du même mot contient une entrée.

Les dictionnaires monolingues sont généralement réédités chaque année. Les nouvelles versions ne sont généralement pas différentes des anciennes, mais complétées ou corrigées. La plus grande modification qu'elles subissent est le rajout de mots nouveaux (néologismes lexicaux) et l'ajout de nouveaux sens à des mots déjà existants (néologismes sémantiques).

3.1.2. Dictionnaire historique

Contrairement aux dictionnaires qui décrivent l'usage actuel du lexique, les dictionnaires à usage historique s'intéressent plutôt à l'histoire des mots et à leur évolution.

Plusieurs dictionnaires traitent le lexique du français du point de vue historique comme le grand Robert, le Grand Larousse, le Trésor de la langue française, etc. Leur méthode consiste à dater l'attestation du mot, c'est-à-dire la date de son entrée dans un texte ou dans un dictionnaire (A. Rey 2012). On explique, généralement, les évolutions morphologiques et sémantiques des mots.

Le Robert a consacré un dictionnaire complet à l'évolution historique de la langue française le « dictionnaire historique de la langue française », sous la direction d'Alain Rey, où chaque mot est décrit de façon détaillée, depuis son apparition. Il donne, avec des dates précises, les changements du sens et les transformations morphologiques que subit le mot à travers l'histoire.

Les dictionnaires de langue comme *le Grand Robert*, *le Grand Larousse de la langue française* et *le Trésor de la langue française* rajoutent à leurs entrées la notice *historique* dans laquelle ils donnent la première attestation du mot et son évolution. Cela représente un supplément important et enrichissant pour les dictionnaires généraux.

3.1.3. Dictionnaire plurilingue

Les dictionnaires monolingues, bilingues ou plurilingues se distinguent par le nombre de langues que présentent leurs nomenclatures (J. Pruvost 2006).

J. Pruvost (2006) rappelle que la première vocation des dictionnaires, en Occident, était de traduire les textes sacrés bibliques en langue vernaculaire. Ils ont été des outils nécessaires pour rapprocher les textes de la bible en latin, considérée comme une langue morte, des textes en langue maternelle, tout en sachant que le latin n'était maîtrisé que par les érudits.

Si les dictionnaires monolingues répondent aux questions qu'un usager peut se poser sur le sens des mots rares et difficiles, un dictionnaire bilingue ou plurilingue se concentre sur les emplois des mots les plus usuels (B. Quemada 1968).

L'objectif des lexicographes élaborant les dictionnaires bilingues ou plurilingues est à la fois d'enseigner une langue et de pouvoir donner des mots synonymes ou équivalents afin de servir les traducteurs dans leur travail. Le lexicographe ne se soucie pas de donner une description au mot ni de son histoire ou de son évolution.

Les dictionnaires plurilingues ressemblent par certains côtés aux dictionnaires de synonymes, si ce n'est que les premiers donnent les équivalents des mots dans deux langues, voire trois langues différentes (français/ anglais, français/ espagnol, etc.), alors que les seconds donnent les équivalent des mots dans la même langue (français, anglais, espagnol, etc.) (*ibid*).

Quelles que soit leurs divergences, pour B. Quemada, la méthodologie des dictionnaires bilingues et plurilingues a influencé l'élaboration méthodologique des dictionnaires monolingues. De toute façon, ils sont considérés par les lexicographes comme étant les ancêtres des dictionnaires monolingues, parce que les dictionnaires étaient, avant tout, des livres qui proposaient des équivalents des mots dans d'autres langues. Autrement dit, ils étaient des outils d'aide à la traduction.

3.2. Présentation globale

Nous évoquons ici, quelques dictionnaires ayant pour but de faire une présentation globale du lexique, en l'occurrence, *Morfetik* qui est un dictionnaire morphologique élaboré par M. Mathieu-Colas, le Trésor de la langue française ainsi que les dictionnaires de synonymes.

3.2.1. Morfetik

La base *Morfetik* est un dictionnaire morphologique du français mise en place par M. Mathieu-Colas (2009). Elle compte environ 100 000 lemmes. *Morfetik* inclut un moteur de flexion, des interfaces de consultation et des outils d'exploitation. Elle a pour objectif l'exhaustivité, la précision, la fiabilité et le respect des mots du français. La base *Morfetik* a été élaborée à partir de plusieurs sources lexicographiques riches, variées et différentes tel que Le DELAS (Dictionnaire électronique du LADL), le Petit Robert, le Grand Robert, le Petit Larousse illustré, le Lexis, le Grand Larousse encyclopédique et le Grand Dictionnaire Encyclopédique Larousse (GDEL), le Trésor de la langue française, le Harrap's et le Robert & Collins, des dictionnaires d'argot, des tables de conjugaison, etc. (M. Mathieu-Colas 2009).

Elle comporte cinq catégories réparties dans cinq tables différentes. Il s'agit de la table des noms, la table des verbes, la table des adjectifs, la table des déterminants et pronoms et la table des mots invariables. Si certains mots en français sont invariables comme les déterminants, les pronoms, les prépositions, les adverbes, les conjonctions et les interjections. Il s'agissait d'établir une liste de chaque catégorie de mot. Les tables des noms, des verbes et des adjectifs nécessitent de faire une table pour les lemmes et des sous-tables pour les flexions afin de générer toutes les formes possibles. Par exemple les formes conjuguées pour les verbes, le nombre et le genre pour les noms et les accords pour les adjectifs.

Le système de traitement du français *Morfetik* permet de générer, automatiquement, l'ensemble des formes des mots simples de la langue française, ce qui correspond à environ 760 000 valeurs (L. Tromeur 2011).

3.2.2. TLFi

Le Trésor de la Langue Française informatisé (TLFi) est un dictionnaire élaboré par l'*Atilf* (analyse et traitement automatique de la langue française) qui est une unité mixte de recherche associée au CNRS et à l'université de Nancy 2. Il existe en version papier, imprimé en 16 volumes, entre 1971 et 1994. La version papier et la version électronique sont pratiquement identiques, mais la version électronique propose à l'utilisateur plusieurs systèmes de recherche, contrairement à la version papier qu'on consulte comme un dictionnaire traditionnel.

Le trésor de la langue française informatisé propose au moins trois types de recherches :

- Recherche d'un mot : ce mode de recherche est notamment utilisé dans la version papier et les autres dictionnaires de langue avec plus ou moins de différences, notamment, ceux qui optent pour les définitions, l'étymologie, l'orthographe et la prononciation des mots.
- Recherche assisté : dans la recherche assistée, le dictionnaire met à notre disposition une application qui permet d'affiner la recherche. On peut ainsi entrer un mot et choisir le code comme (adjectif, verbe, article, conjonction, etc.), choisir le type d'emploi (ironique, argotique, littéraire, invariable, etc.) et le domaine (médecine, agronomie, architecture, etc.).
- Recherche complexe : ce mode de recherche est exploité, généralement, dans l'élaboration des corpus. On peut ainsi faire une recherche ciblée telle la recherche de mots utilisés dans le domaine médical par l'académie française avec un emploi soutenu.

Le TLFi est un dictionnaire très riche en termes de vocabulaire et de définitions. Il contient :

- 100 000 mots avec leur histoire ;
- 270 000 définitions ;
- 430 000 exemples ;
- 350 000 000 caractères.

3.2.3. Dictionnaires de synonymes

Les dictionnaires de synonymes donnent les mots synonymes, proches, voisins ou équivalents des entrées du dictionnaire. Ils proposent, souvent, plusieurs synonymes pour chaque mot. Ils sont utiles lorsqu'on cherche le mot le plus proche d'un mot donné.

Certains dictionnaires de synonymes se limitent à un seul terme (B. Quemada 1968). Par exemple *malaria* (paludisme), *malaxer* (pétrir), etc. D'autres multiplient les synonymes et les exemples pour mieux cerner le mot et se rapprocher le plus possible de son sens (*ibid*), par

exemple *maladie* (affection, attaque, atteinte, crise, dérangement, épreuve, incommodité, indisposition, infirmité, etc.). Nous constatons que les synonymes du mot *maladie* sont plus ou moins proches du sens initial (*maladie*). Si l'on prend l'exemple du synonyme *dérangement*, on constate qu'il est loin d'être l'équivalent du mot *maladie* ; c'est la même chose pour *incommodité*, *épreuve* et *indisposition* qui peuvent avoir des sens différents dans d'autres contextes.

Le mot *infirmité* fait partie du vocabulaire médical, mais ne signifie pas exactement *maladie*, en médecine, *infirmité* signifie plutôt *handicap* et *invalidité*

3.3. Présentation ciblée

Les dictionnaires électroniques et les bases de données sont élaborés, généralement, pour donner la possibilité à l'utilisateur de faire des recherches ciblées. C'est ce genre de présentation qui a motivé l'élaboration de nos dictionnaires électroniques de la langue de la médecine, notamment, le dictionnaire des prédicats, le dictionnaire des arguments et le dictionnaire des humains, qui ont servi par la suite à l'élaboration d'une interface d'interrogation des bases de données.

Les présentations ciblées peuvent être multiples et variées, selon les besoins. Par exemple, elles peuvent être d'ordre terminologique, d'ordre historique ou être relatives aux niveaux de langues.

3.3.1. Terminologique

Les présentations terminologiques sont des recherches ciblées. Elles consistent à rechercher des mots bien définis. Elles peuvent porter sur la morphologie des mots, sur leur sémantique, sur leur syntaxe, etc.

Selon Xavier Laurent Salvador et Joëlle Ducos (2011), les travaux sur la terminologie des langues de spécialité ont beaucoup évolué depuis une quinzaine d'années. Ainsi, Michel Mathieu-Colas (2009) a développé une base de données terminologique de français dans le but de la recherche morphologique. Il a recensé plus de 100 000 mots de français tout azimut.

L'objectif de la base de données était de faire uniquement des recherches d'ordre morphologique.

Le principe fondamental de la présentation terminologique des termes stockés dans une base de données est la recherche des termes par domaine d'activité ou par classe ou sous-classe sémantique du même domaine. Ainsi, toute base de données destinée à la transmission des informations doit être structurée de façon à faciliter la recherche documentaire et de permettre aux utilisateurs d'effectuer des recherches terminologiques ciblées.

3.3.2. Historique

L'élaboration des bases de données dans le but de faire des présentations historiques des mots de la langue souhaite apporter, aux linguistes et aux philologues, des résumés fidèles sur l'étymologie et l'évolution des termes à travers le temps. Dans la mesure du possible, des datations figureront devant chaque mot afin de situer, dans le temps, chaque étape que le mot a traversée et les transformations qu'il a subies, avec des repères précis.

Le meilleur exemple pour illustrer le modèle d'une recherche ciblée sur l'histoire de la langue est le projet *Crealscience* de Xavier Laurent Salvador et Joëlle Ducos (2011). Il s'agit d'une élaboration d'un dictionnaire du lexique scientifique du Moyen-âge. Quatre paramètres ont été retenus en compte dans ce cas. Il s'agit de l'étymologie des mots, de leurs lemmes, de leurs définitions et de leurs occurrences.

- L'étymologie des mots est importante. Elle est présente dans beaucoup de dictionnaires de langue comme dans le Trésor de la Langue Française et le Grand Larousse. Elle permet aux linguistes et aux philologues de connaître les origines des mots.
- Les formes lemmatisées des mots posent de véritables problèmes, notamment à cause des variations morphologiques et orthographiques des termes au Moyen-âge.
- Les définitions permettent de « délimiter les traits définitoires du concept et sont au cœur de toute activité lexicographique » (*ibid*).

3.3.3. Par niveaux de langue

Les présentations par niveau de langue consistent à présenter les termes dans l'interface d'interrogation selon que le mot appartienne à un registre soutenu, à un registre courant ou à un registre familier.

Les corpus sont différents du point de vue du niveau de langue. Certains sont hautement spécialisés. Dans ce cas, il s'agit de textes, généralement écrits et publiés par des spécialistes et des chercheurs. Ils sont souvent publiés dans des revues spécialisées.

Les textes spécialisés sont difficilement accessibles au public non instruit parce qu'ils sont pleins de termes techniques et hautement spécialisés, qui proviennent notamment du grec et du latin.

D'autres mots appartiennent au registre courant, c'est-à-dire au corpus moyennement spécialisé, à mi-chemin entre les corpus spécialisé et les corpus non spécialisés. Dans ce cas, les mots du corpus sont rédigés dans une langue standard, puisqu'ils sont destinés au grand public.

Certains mots appartiennent au registre familier, ils apparaissent dans les corpus non spécialisés.

En principe, la présentation se fait, ici, par registre de langue. Ainsi, les usagers font des recherches ciblées, selon les trois registres mentionnés ci-dessus.

Conclusion

Pour conclure ce chapitre dédié aux bases de données et aux dictionnaires, nous évoquons l'association entre l'informatique et les dictionnaires, pour des raisons économiques ou pratique. Aujourd'hui, les linguistes se sont tournés vers l'informatique en tant qu'outil afin de faciliter l'usage des dictionnaires afin que le problème du volume ne se pose plus.

Le meilleur moyen pour stocker et traiter les informations dictionnaires réside dans les bases de données, que ce soit via *PHP*, *Excel*, *XML*, etc. La structuration des bases de données nécessite de prendre en compte les descriptions nécessaires des nomenclatures qui constituent les dictionnaires. Leurs exploitations lexicographiques sont différentes. Elles

peuvent aller des utilisations variées aux présentations ciblées, en passant par les présentations globales comme l'exemple de *Morfetik* ou du *TLFi*. Pour notre base de données, nous optons pour une présentation ciblée afin d'obtenir des informations précises sur le vocabulaire médical, qu'elles soient de niveau morphologique, syntaxique ou sémantique.

CHAPITRE 8

APPLICATION : INTERFACE D'INTERROGATION DU LANGAGE DE LA MÉDECINE

Introduction

Avec toutes les informations qui circulent sur le web et qui sont stockées dans des serveurs, le web est devenu à la fois un champ d'étude, une ressource et une base de données (L. Emirkanian & C. Fouqueré 2003). Si l'ambition du web est de relier les ordinateurs du monde entier afin d'échanger des informations sous forme de documents électroniques, ce but n'est atteignable que par le biais d'une programmation informatique qui s'attache à rendre les corpus visibles et surtout lisibles. Parmi les technologies qui permettent de stocker, de gérer et de diffuser les informations sur le Web, il y a les systèmes de gestion de bases de données. Nous présentons dans ce chapitre les outils informatiques que nous avons exploités pour l'élaboration d'une interface d'interrogation des dictionnaires du vocabulaire médical.

1. La base de données

Nous présentons dans cette partie le langage que nous avons utilisé pour l'élaboration de nos bases de données.

1. 1. Le langage SQL

Ici, nous définissons le langage SQL puis nous expliquons son fonctionnement et enfin nous justifions le choix de ce langage pour notre base de données.

1.1.1. Définition

Le SQL (*Structured Query Language*, en français *langage de requêtes structuré*) est un langage de définition, de traitement et de contrôle de données pour les bases de données relationnelles. Il permet d'exploiter les bases de données en communiquant avec les serveurs qui les stockent. Le langage SQL est élaboré suite à la création de plusieurs modèles relationnels comme *IBM Sequel*, *IBM Sequel/2*, *IBM System/Ret IBM DB2*. Il existe en plusieurs versions comme *SQL 86*, *SQL 89* et *SQL 2*, (*Wikipédia*).

Le langage SQL permet de créer des tables dans lesquelles on peut manipuler, supprimer ou rajouter des données. Il permet de gérer les fichiers d'une base de données et de prendre en charge les fonctionnalités en rapport avec la sécurité et la confidentialité des données et il donne des possibilités qui facilitent l'accès aux données stockées et structurées dans la base (P. Rigaux 2005).

Nous avons exploité, pour la gestion de notre base sur les prédicats, les arguments et les humains de la langue de la médecine, le système de gestion de base de données MySQL.

1.1.2. Utilisation

Il existe principalement trois manières d'utiliser le langage SQL, (*ibid*).

1. il peut être utilisé par un interpréteur standard avec un bloc de code exécuté comme la seule commande du SQL.
2. il peut être utilisé avec un fichier de commande PL/SQL.
3. il peut être utilisé aussi avec un programme de stockage de données.

Selon P. Rigaux (2005) « dans la plupart des cas les requêtes SQL exécutées dans les scripts PHP sont fixées par le programmeur et ce dernier connaît le type du résultat (nombre d'attributs et noms des attributs). Il peut arriver que les ordres SQL soient dynamiques, c'est-à-dire déterminés à l'exécution. C'est le cas par exemple quand on permet à l'utilisateur d'effectuer directement des requêtes SQL sur la base et d'afficher le résultat sous forme de table. On peut alors faire appel à des fonctions MySQL qui donnent des informations sur le type du résultat ».

Nous utilisons le langage SQL pour adresser des requêtes sur notre base de données au serveur MySQL, et recourons à PHP pour afficher les résultats sous forme de tableaux HTML.

1.1.3. Justification du choix

Le langage SQL est relativement simple à utiliser. C'est un langage déclaratif, c'est-à-dire qu'il consiste à décrire le résultat que l'on souhaite obtenir. Sa syntaxe et son vocabulaire se rapprochent du langage naturel. Nous avons utilisé le système de gestion de bases de données MySQL, qui fait partie des serveurs les plus utilisés.

1.2. Structure de la base de données

Nous avons constitué une seule et unique base qui contient quatre tables. Le nombre de tables correspond au nombre de dictionnaires servant de sources pour l'interface d'interrogation.

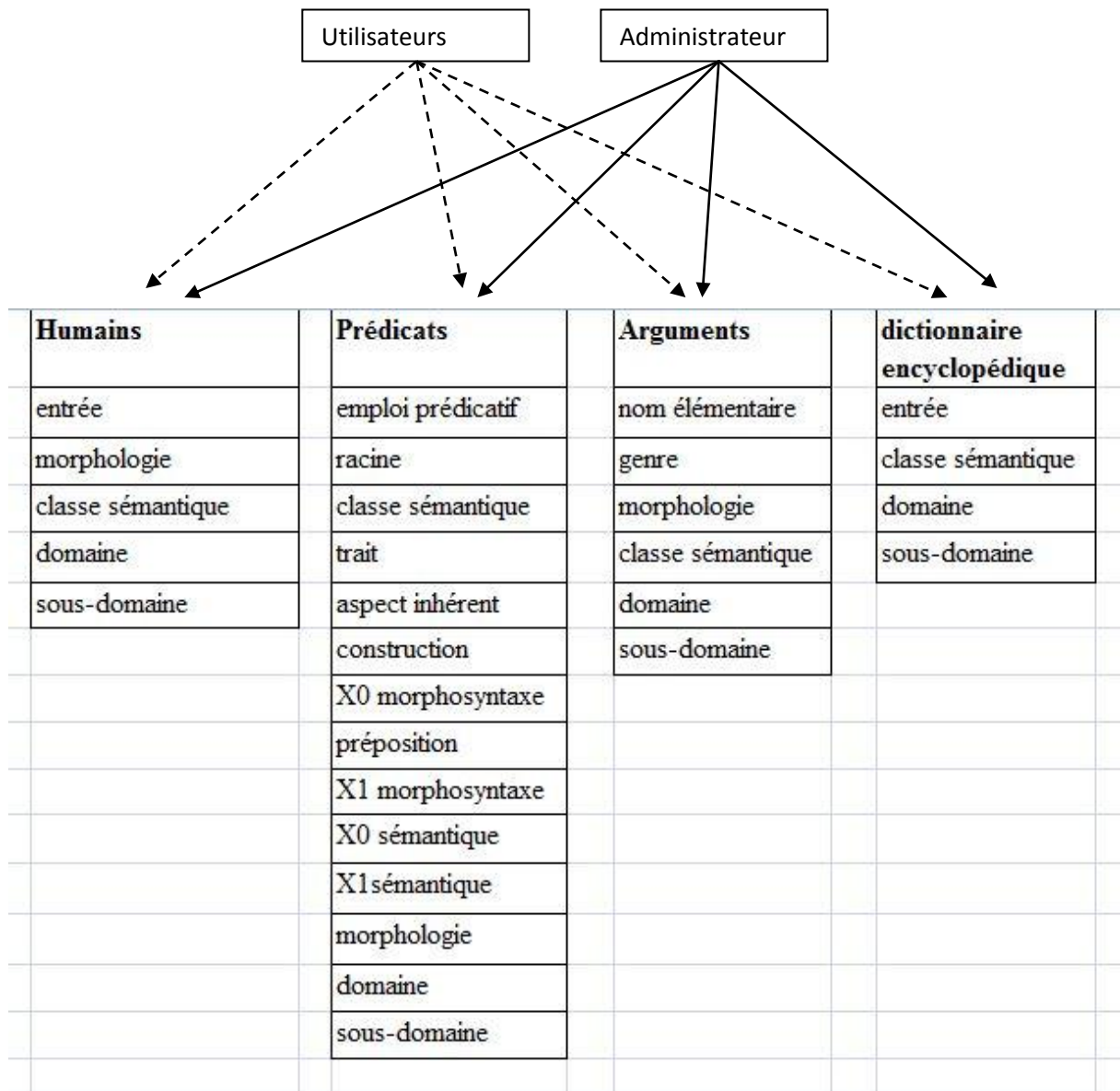
1.2.1. Définition de la base

Pour interroger nos dictionnaires, nous utilisons la base de données intitulée « dictionnaires ». La base de données contient quatre tables, en l'occurrence, la table des prédicats de la langue de la médecine, intitulée « prédicats », la table des arguments de la langue de la médecine, intitulée « arguments », la table des humains de la langue de la médecine, intitulée « humains » et la table du dictionnaire encyclopédique.

Notre base de données comporte l'ensemble des informations que contiennent nos dictionnaires. Les informations sont organisées de façon structurée afin de récupérer les informations facilement et de manière organisée.

1.2.2. Structure de la base

Pour expliquer la structuration de notre base de données, nous présentons un schéma relationnel décrivant les tables de la base.



—————> = interroger et mettre à jour les données

-----> = interroger la base

Schéma relationnel de la base

1.2.3. Fonctionnement de la base

Le rôle principal de la base de données se résume au stockage des données. Son utilisation consiste à interroger et exploiter les données stockées.

L'avènement des bases de données a révolutionné le mode d'accès à l'information ; cela veut dire que, contrairement au stockage des données dans les supports papiers ou électroniques, les bases de données facilitent la recherche documentaire.

Notre base de données permet non seulement d'interroger les données, mais aussi de les mettre à jour (modifier, ajouter et supprimer les données de la base).

La base de données contient deux tables supplémentaires. La première est réservée à l'administrateur : Elle lui donne accès à la base pour l'interrogation ainsi que les mises à jour. La seconde est réservée à l'utilisateur afin d'avoir accès à l'interrogation de la base.

1.3. Structure des tables

Les tables de la base de données correspondent aux tableaux enregistrés initialement dans le format Excel. Chaque table comporte des colonnes et des lignes, les colonnes correspondent aux attributs, les lignes aux données.

1.3.1. Table relative aux prédicats

La table relative aux prédicats est caractérisée par douze attributs (champs) qui sont du type chaîne de caractères. Les attributs sont les suivants :

- Emploi prédicatif : les emplois prédicatifs correspondent à toutes les entrées du dictionnaire des prédicats ayant une fonction prédicative.
- Racine : elle correspond aux racines des formes simples (les mots composés, les mots dérivés et les mots construits ne sont pas concernés par ce champ).

- Classe sémantique : elle informe sur la classe sémantique à laquelle l'entrée correspond.
- Trait : il correspond au type sémantique *action*, *état* ou *événement*.
- Aspect inhérent : il s'agit d'informations sur le temps interne au procès, comme le provisoire, le perfectif ou l'imperfectif.
- Construction : c'est la construction syntaxique dans laquelle apparaît l'entrée, autrement dit, sa combinatoire avec les autres unités de la phrase.
- X0 morphosyntaxe : correspond à la nature syntaxique du groupe nominal du premier argument.
- Préposition : c'est l'éventuelle préposition qui fonctionne avec l'entrée, si elle correspond à un verbe.
- X1 morphosyntaxe : correspond à la nature syntaxique du groupe nominal du deuxième argument.
- X0 sémantique : correspond à la nature sémantique du nom qui correspond au premier argument comme *humain*, *maladie*, *médicament*, etc.
- X1 sémantique : correspond à la nature sémantique du nom correspondant au deuxième argument comme *humain*, *partie du corps*, *maladie*, etc.
- Morphologie : c'est la morphologie de l'entrée. Elle peut être simple, composée, construite ou dérivée.
- Domaine : il correspond au secteur d'activité de l'entrée.
- Sous-domaine : il correspond à la spécialité médicale de l'entrée.

1.3.2. Table relative aux arguments

- Nom élémentaire : correspond aux entrées que comporte le dictionnaire des arguments.
- Genre : ce sont les genres qui correspondent à chaque entrée. Ils sont au nombre de quatre (féminin, masculin, féminin pluriel et masculin pluriel).
- Morphologie : c'est le champ qui fournit la catégorie morphologique de chaque entrée comme *N de N*, *NA*, etc.
- Classe sémantique : correspond aux classes sémantiques de chaque entrée.
- Domaine : il correspond au secteur d'activité de l'entrée.
- Sous-domaine : il correspond à la spécialité médicale de l'entrée.

1.3.3. Table relative aux humains

- Entrée : correspond aux entrées du dictionnaire des humains.
- Morphologie : donne la morphologie de chaque entrée.
- Classe sémantique : correspond aux classes sémantiques de chaque entrée du dictionnaire.
- Domaine : c'est l'information sur le domaine d'activité auquel correspond chaque entrée.
- Sous-domaine : il correspond à la spécialité médicale de l'entrée.

1.3.4. Table relative au dictionnaire encyclopédique

- Entrée : correspond aux entrées du dictionnaire des humains.

- Classe sémantique : correspond aux classes sémantiques de chaque entrée du dictionnaire.
- Domaine : c'est l'information sur le domaine d'activité auquel correspond chaque entrée.
- Sous-domaine : il correspond à la spécialité médicale de l'entrée.

2. Interface d'interrogation

Une interface d'interrogation sert à mieux explorer nos dictionnaires. Elle permet de faire des recherches ciblées. Etant donné que nos dictionnaires comportent des entrées définies selon leurs aspects syntaxiques, sémantiques et morphologiques, l'interrogation des dictionnaires se fonde sur la recherche des critères définitoires de chaque entrée.

2.1. Le langage PHP

Pour élaborer notre interface d'interrogation, nous avons utilisé le langage de programmation PHP. Nous proposerons ci-dessous une définition de ce langage, son utilité ainsi que les arguments justifiant notre choix.

2.1.1. Définition

Le PHP est un langage informatique dynamique (J-M. Defrance 2006). À l'origine, il a été créé particulièrement pour gérer des formulaires. Le PHP est à sa cinquième version. La première (PHP FI) est réservée à la gestion des formulaires. Son code s'est rapidement ouvert à la communauté internet, ce qui l'a aidée à évoluer rapidement. La deuxième version (PHP2) est diffusée en version bêta, mais a été éphémère. Elle a été remplacée par la version 3 (PHP3), mais cette dernière n'était pas performante. Cela a précipité son remplacement par la quatrième version (PHP4) qui affiche des performances plus élevées. En 2004, on a

développé la cinquième version (PHP5) avec l'objectif de rendre PHP plus performant, tout en restant un langage simple à utiliser (*ibid*).

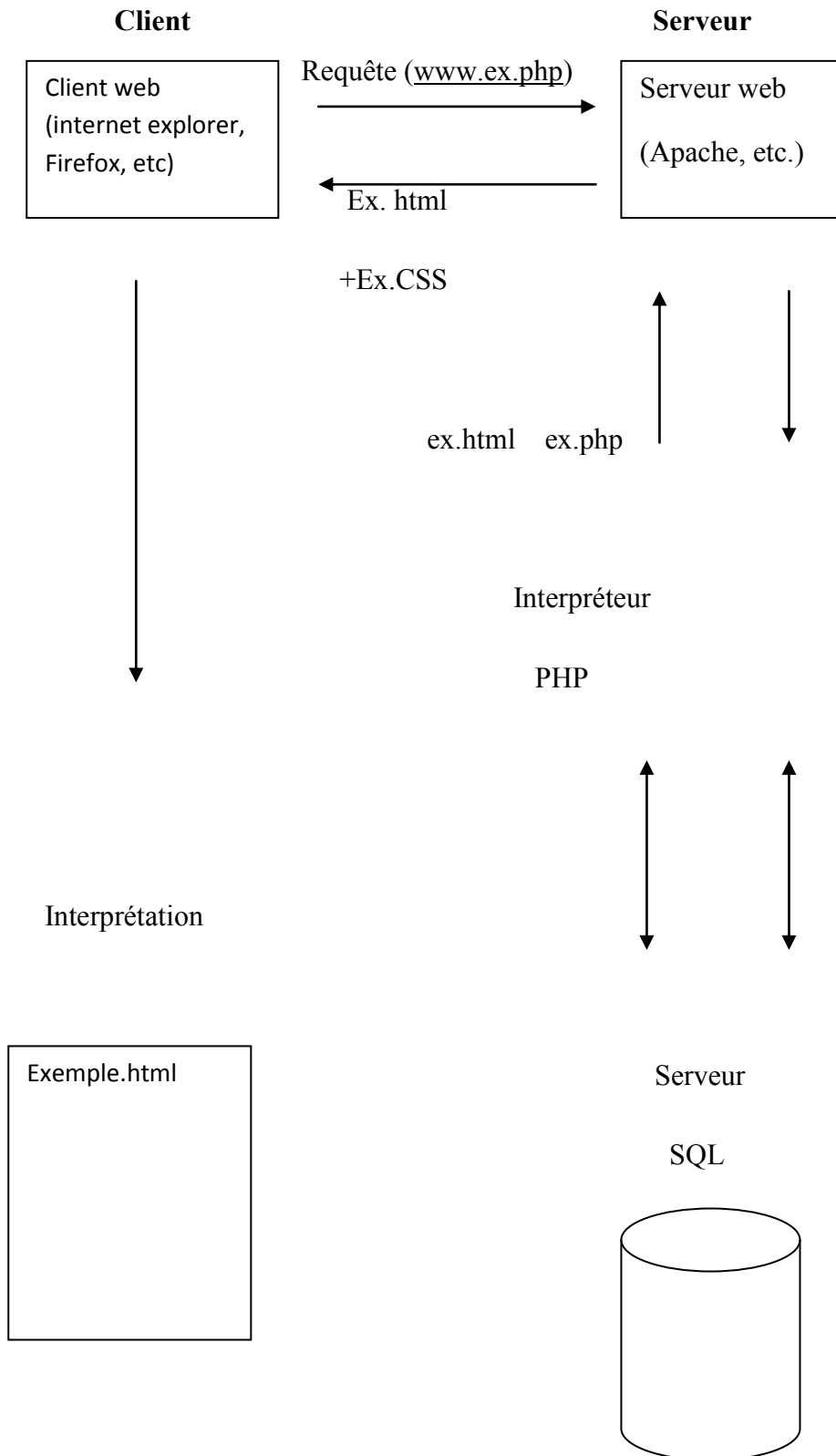
2.1.2. Fonctionnement

Le PHP est un langage de programmation interprété côté serveur. Il est utilisé pour générer des pages HTML de manière dynamique. Le langage PHP est proche syntaxiquement du langage C, (P. Rigaux 2005). Il peut être associé à un serveur de gestion des bases de données comme MySQL, ce que nous avons réalisé dans le cadre de notre application.

Lors de la phase d'interprétation, PHP construit une page HTML qui peut éventuellement contenir différents types de codes qui seront interprétés côté client (Javascript, Flash, Actionscript, etc.).

Par ailleurs, le langage PHP est conçu pour communiquer de manière simple et directe avec les serveurs SQL installés côté serveur. Ainsi, durant la construction de la page HTML, des informations peuvent être récupérées sur une base de données SQL (PHP se charge d'adresser les requêtes au serveur) et injectées au sein de la page.

Nous pouvons finalement illustrer le rôle de PHP à l'aide du schéma suivant :



2.1.3. Justification du choix

Selon P. Rigaux (2005), « PHP - le plus souvent associé à MySQL – est à l'heure actuelle le plus répandu des langages de programmations sur le Web ». Bien qu'il existe de nombreuses alternatives au couple PHP MySQL, ce dernier constitue indéniablement une solution robuste et répandue pour la construction d'applications web.

De plus, le PHP est un langage souple et facile à maîtriser par un débutant. Il a une syntaxe facile et il a tendance à proposer des solutions pratiques pour l'utilisateur. Il est libre, gratuit, et son installation est facile, sur Windows comme sur Linux.

2.2. Le langage HTML / CSS

Nous présentons ci-dessous le langage HTML et le langage CSS ainsi que leurs fonctionnements.

Les HTML est un formalisme qui sert à structurer le contenu des pages web. Le CSS est un formalisme qui sert à mettre en forme le contenu des pages web.

2.2.1. Définition du HTML

Le HTML (HyperText Mark-up Language), est un langage de balises qui fonctionne d'une manière similaire au XML. Ainsi, les contraintes syntaxiques (utilisation de chevrons pour les balises, non chevauchement de celles-ci, etc.) sont identiques. Néanmoins, dans la mesure où HTML est utilisé de manière massive pour publier les informations sur le web, son interprétation est plus souple que celle du XML : nombre de sites web contiennent des pages syntaxiquement mal formées sans pour autant que cela bloque leur fonctionnement.

L'interprétation des pages HTML est réalisée côté client par les navigateurs web. Aujourd'hui, avec la cinquième version du HTML, il est relativement facile de construire des pages contenant des types de données différents : images, vidéos, zones de dialogue, etc.

2.2.2. Définition du CSS

Le CSS est un langage informatique qui sert à décrire la représentation d'un document HTML (R. York 2007). Chacun des éléments définis par la structuration HTML peut ainsi faire objet d'une mise en forme spécifique : taille, police, couleur, etc.

S'il n'est pas indispensable à la consultation des pages web (d'ailleurs les navigateurs proposent tous des options autorisant la désactivation des propriétés CSS), le CSS est systématiquement utilisé pour faciliter leur lecture.

2.2.3. Utilisation

Quand on veut afficher un texte sur le Web, on utilise les balises HTML. Chaque partie du texte est insérée entre deux balises, une balise ouvrante et une balise fermante. Un fichier HTML peut être constitué d'un prologue. Il sert à définir la version HTML utilisée pour interpréter le code, par exemple `<! Doctype html>` pour la version 5. D'autres balises servent à définir les différentes parties du fichier comme `<head></head>` pour le titre et les feuilles de style et `<body></body>` pour le corps du fichier.

Nous illustrons cette brève définition avec la structure minimale suivante :

```
<html>  
<head>  
<title>  
    Titre de la page  
</title>  
</head>  
<body>  
    Contenu de la page  
</body>  
</html>
```

Le CSS peut être utilisé de trois façons différentes. Il peut tout d'abord être placé directement dans les balises de la page HTML, via l'attribut « style ». Il peut également être déclaré de manière globale dans l'en-tête du document HTML. Enfin, le code CSS peut être

enregistré dans une feuille de style externe à la page web. Il est alors nécessaire de déclarer dans cette dernière le chemin d'accès à la feuille de style. Cette troisième solution présente l'avantage de centraliser le code, de faciliter sa maintenance et de le rendre réutilisable (C. Schmitt 2005). Nous présentons ci-dessous quelques attributs CSS et leurs effets sur la mise en forme.

Center = centrer un élément

Height = hauteur de l'élément

Width = largeur de l'élément

Color = couleur des caractères contenus dans l'élément

Etc.

En guise de conclusion, nous dirons que le CSS sert uniquement à l'affichage des pages et des fichiers Web.

2.3. Les formulaires

Nous commençons par la définition des formulaires puis nous expliquons leur structuration et leur fonctionnement.

2.3.1. Définition

Un formulaire représente une boîte de dialogue sur internet. Il permet d'échanger les informations entre internautes.

Un formulaire est un espace où l'on saisit des données dans l'interface utilisateur. Il peut comporter plusieurs champs destinés à accueillir les données que l'on veut stocker dans la base. Un formulaire peut comporter plusieurs champs pour saisir des textes ou insérer des images, et des boutons pour la validation et pour l'envoi des données. Les boutons permettent d'envoyer des données au serveur.

La création d'un compte sur internet ou la diffusion et l'échange d'informations sur internet passe forcément par un formulaire.

Un formulaire reste toujours ouvert, cela veut dire qu'il peut être modifié ou complété à tout moment. Il permet d'apporter des éléments nouveaux à la base ou au contraire, de supprimer des éléments considérés inutiles.

2.3.2. Structuration des formulaires

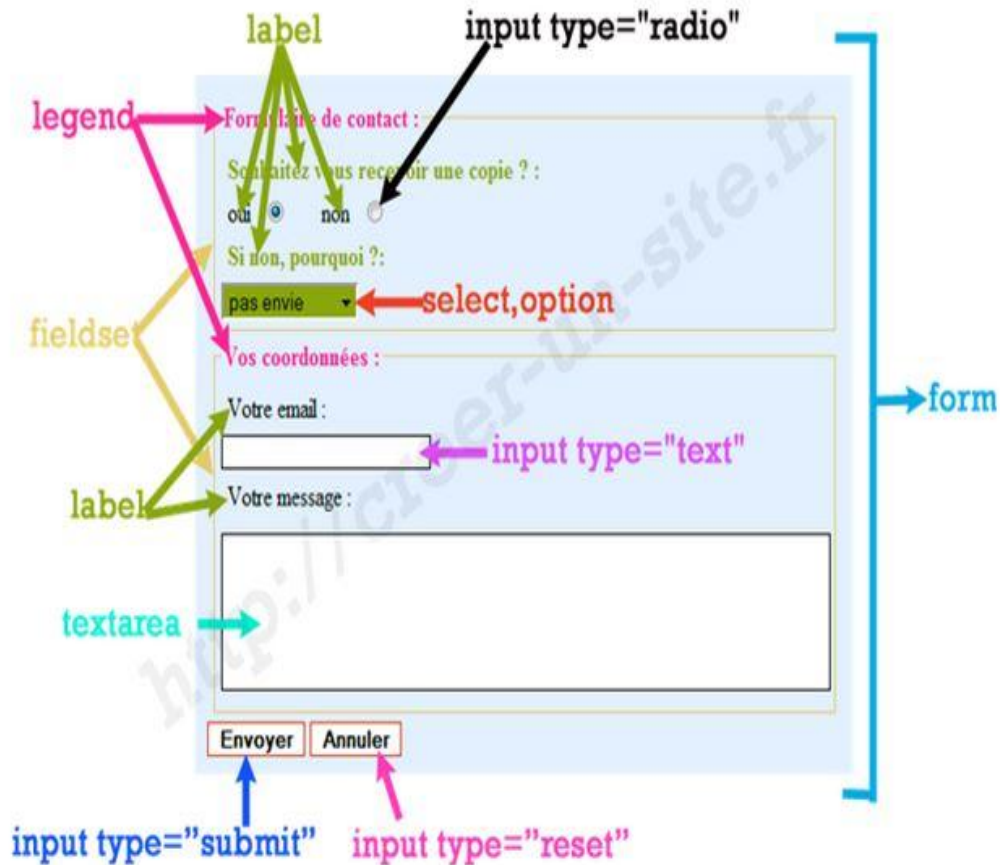
Les formulaires sont structurés et organisés selon le contenu et la nature des informations qu'ils contiennent ainsi que les besoins des utilisateurs.

Ils sont constitués des balises `<form></form>`. Les champs qui sont à l'intérieur sont des *input*. « action » et « method » sont deux attributs du formulaire qui sont insérés dans les balise `<form>`. L'attribut « action » est utilisé pour définir la destination du formulaire lors de l'envoi. L'attribut « method », dont la valeur est soit POST soit GET, est utilisé pour désigner la méthode utilisée pour la récupération des données renseignées par l'utilisateur. La méthode *post* est la plus utilisée.

Les deux méthodes sont utilisées aussi dans le partage des données et des variables globales.

La méthode *get* fait apparaître les variables globales transmises sur l'URL.

Le schéma ci-dessous illustre un exemple de formulaire.

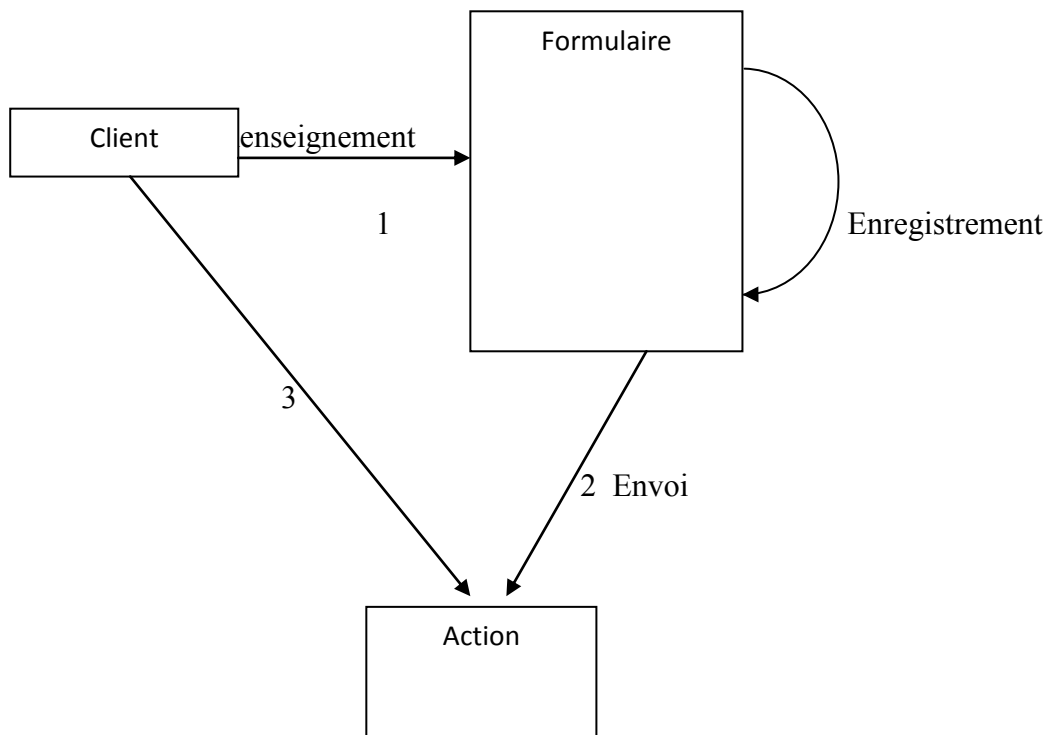


2.3.3. Fonctionnement des formulaires

Après avoir saisi toutes les informations que l'on veut stocker dans la base de données, les formulaires jouent le rôle d'intermédiaire entre les utilisateurs et le serveur.

La force des formulaires réside dans la rapidité de l'interaction homme-machine (la transmission des résultats est effectuée en quelques secondes). Les formulaires renvoient les résultats tels qu'ils sont construits par le programmeur : c'est de lui que dépendent la pertinence et la précision des critères de recherches.

Pour résumer, nous dirons que dans le fonctionnement des formulaires, l'utilisateur renseigne les données, le formulaire les récupère et les traite.



3. Mode d'emploi

Notre interface comporte quatre formulaires d'interrogation. Le premier est consacré au dictionnaire des prédicats (PRED_DIC), le deuxième est consacré au dictionnaire des arguments (ARGU_DIC), le troisième est consacré au dictionnaire des humains (HUM_DIC) et le quatrième est consacré au dictionnaire encyclopédique (ENCY_DIC).

Nous avons élaboré les formulaires afin de faciliter la recherche et filtrer les résultats. Ainsi, pour rechercher des informations sur un terme médical donné, il nous suffit de remplir le formulaire selon les caractéristiques et les besoins de notre requête. Les trois formulaires sont structurés différemment.

3.1. Requête sur les prédicats

Nous présentons dans cette partie le formulaire consacré au dictionnaire des prédicats, puis la requête associée à ce formulaire et enfin les scripts PHP.

3.1.1. Présentation du formulaire

Le formulaire relatif aux prédicats est composé de douze champs qui correspondent aux colonnes du dictionnaire. Dans le formulaire, chaque champ est présent sous forme d'une liste déroulante qui permet d'afficher les données. Nous avons utilisé des requêtes SQL pour afficher les données à partir de la base. Dans le formulaire, chacune des colonnes correspond à une des informations :

Emploi prédicatif

Racine

Classe

Type

Aspect inhérent

Construction

X0 morphosyntaxe

Préposition

X1 morphosyntaxe

X0 sémantique

X1 sémantique

Morphologie

Domaine

Sous-domaine

Dictionnaires de termes Médicaux

Dictionnaire des prédicats

Accueil Prédicats Arguments Humains

Recherche par **emploi prédicatif**

Recherche par **racine**

Recherche par **classe**

Recherche par **type**

Recherche par **aspect inhérent**

Recherche par **construction**

Recherche par **X0 morphosyntaxe**

Recherche par **préposition**

Recherche par **X1 morphosyntaxe**

Recherche par **X0 sémantique**

Recherche par **X1 sémantique**

Recherche par **morphologie**

Recherche par **domaine**

Recherche par **sous-domaine**

3.1.2. Présentation de la requête

Dans le dictionnaire des prédicats, la recherche d'une entrée peut être multicritère, c'est-à-dire que l'utilisateur peut exploiter tous les champs afin d'obtenir toutes les définitions que l'entrée du dictionnaire comporte. La recherche d'un terme médical peut s'effectuer dans la base en exploitant uniquement quelques champs. Autrement dit, la recherche se formule selon les besoins que nous avons en termes d'informations concernant telle ou telle entrée. De telle manière que nous pouvons afficher soit la morphologie, soit la syntaxe, soit la sémantique des mots. À cela, nous pouvons rajouter d'autres informations telles que l'emploi prédicatif, la classe sémantique, le type et l'aspect inhérent.

3.1.3. Script PHP

Le script PHP est un fichier exécutable codé en PHP. Pour illustrer la façon dont nous avons fait notre script, nous proposons l'exemple ci-dessous. Il s'agit d'une partie du script sur la sélection des emplois prédicatifs du dictionnaire des prédicats.

```
<p>
Recherche par <strong>emploi prédicatif :</strong><br/>
<select name="emploi Predicatif" >
<option value="" selected="selected" size="10"></option>
<?php
include("fonctions/connexion.php");
$query=mysql_query("SELECT * FROM predicat order by emploi Predicatif ASC");
$nb=mysql_num_rows($query);
while ($list=mysql_fetch_array($query))
{
```

```
echo "<option size =\"10\">".$list['emploi Predicatif']."</option>";  
  
}  
  
mysql_close();  
  
>  
  
</select>  
  
</p>  
  
<br/>
```

3.2. Requête sur les arguments

Nous présentons dans cette partie le formulaire consacré au dictionnaire des arguments, la requête associée à ce formulaire et enfin le script PHP.

3.2.1. Présentation du formulaire

Le formulaire relatif au dictionnaire des arguments est constitué de quatre champs correspondant aux colonnes qui comportent les informations sur les entrées du dictionnaire des arguments. Le premier champ correspond aux noms élémentaires, le deuxième correspond au genre, le troisième correspond à la morphologie et le dernier champ correspond à la classe sémantique.

3.2.2. Présentation de la requête

Le dictionnaire des arguments offre à l'utilisateur la possibilité de rechercher les termes du vocabulaire de la médecine ayant une fonction argumentale selon les informations fournies dans la base des données. La recherche peut s'effectuer soit par l'affichage de toutes les entrées du dictionnaire, soit par l'affichage des entrées du genre masculin, féminin, masculin pluriel ou féminin pluriel, d'une morphologie particulière, (par exemple N de N), ou d'une classe sémantique particulière (telle que les parties du corps).

La requête peut être multicritère. Elle se fait en exploitant plusieurs champs comme afficher les mots du genre masculin de la classe micro-organisme.

3.2.3. Script PHP

Pour présenter le script des arguments, nous reproduisons une partie du script consacré à la sélection des noms élémentaires du dictionnaire des arguments.

```
<p name = "N">
```

```
Recherchez par <strong>Nom(s) élémentaire(s) :</strong><br/>
```



```
<select name="nom" value="Choisissez">

<option value="" selected="selected"></option>

<?php

include("fonctions/connexion.php");

$query=mysql_query("SELECT * FROM argument order by nom asc");

$nb=mysql_num_rows($query);

while ($list=mysql_fetch_array($query))

{

echo "<option>".$list['nom']."</option>";

}

mysql_close();

?>

</select>

<br/>

</p>

<br/>
```

3.3. Requête sur les humains

Nous présentons dans cette partie le formulaire consacré au dictionnaire des humains, la requête associée à ce formulaire et enfin le script PHP.

3.3.1. Présentation du formulaire

Le formulaire relatif au dictionnaire des humains est constitué de quatre champs qui correspondent aux colonnes dans lesquelles sont inscrites les informations de chaque entrée du dictionnaire des humains du vocabulaire médical. Dans le premier champ, nous avons inscrit la liste des entrées, dans le deuxième champ, la morphologie qui correspond à chaque entrée, dans le troisième champ, la classe sémantique et dans le quatrième champ le domaine de chaque entrée du dictionnaire.

The screenshot shows a web interface for a medical dictionary. At the top, there is a header 'Dictionnaires de termes Médicaux' and a sub-header 'Dictionnaire des humains'. Below the sub-header, there are four tabs: 'Accueil', 'Préfixes', 'Argumens', and 'Humains'. The main content area contains five search criteria, each with a dropdown menu: 'Recherchez par entrée', 'Recherchez par morphologie', 'Recherchez par classe sémantique', 'Recherchez par domaine', and 'Recherchez par sous-domaine'. At the bottom of the search area, there are two buttons: 'Valider' and 'Annuler'.

3.3.2. Présentation de la requête

Le formulaire d'interrogation du dictionnaire des humains de la langue de la médecine comporte quatre champs. La recherche des termes s'effectue de la même manière que dans les deux premiers dictionnaires. L'utilisateur peut faire une recherche multicritère. Il peut afficher toutes les entrées ou une entrée avec ses caractéristiques. La recherche fonctionne également pour un seul champ, par exemple l'affichage de tous les termes d'une classe sémantique donnée ou les termes d'un domaine particulier.

3.3.3. Script PHP

Pour présenter le script du dictionnaire des humains, nous avons pris une partie du script consacré à la sélection des noms d'humains selon leur morphologie.

```

<p name="G">

Recherchez par <strong>morphologie(s) :</strong><br/>

<select name="morphologie">

<option value=" selected="selected"></option>

<?php

include("fonctions/connexion.php");

$query=mysql_query("SELECT DISTINCT morphologie FROM humain order by
morphologie ASC");

$nb=mysql_num_rows($query);

while ($list=mysql_fetch_array($query)){

echo "<option>".$list['morphologie']."</option>";

}

mysql_close();

?>

</select><br/>

</p>

```

3.4. Requête sur le dictionnaire encyclopédique

Nous présentons dans cette partie le formulaire consacré au dictionnaire des arguments, la requête associée à ce formulaire et enfin le script PHP.

3.4.1. Présentation du formulaire

Le formulaire relatif au dictionnaire encyclopédique est constitué de quatre champs correspondant aux colonnes qui comportent les informations sur les entrées du dictionnaire. Le premier champ correspond aux entrées, le deuxième correspond à la classe sémantique, le troisième correspond au domaine et le dernier champ correspond au sous-domaine.



The screenshot displays the 'Dictionnaires de termes Médicaux' application. At the top, there is a navigation bar with 'Accueil' and 'Dictionnaire encyclopédique' tabs. The main content area is titled 'Dictionnaire encyclopédique' and contains a search form with four dropdown menus labeled 'Recherchez par entrée', 'Recherchez par classe sémantique', 'Recherchez par domaine', and 'Recherchez par sous-domaine'. Below the dropdowns are two buttons: 'Valider' and 'Annuler'.

3.4.2. Présentation de la requête

Le dictionnaire encyclopédique offre à l'utilisateur la possibilité de rechercher les termes du vocabulaire de la médecine selon la classe sémantique, le domaine et le sous-domaine. La recherche peut s'effectuer soit par l'affichage de toutes les entrées du dictionnaire et obtenir les définitions encyclopédiques de toutes les entrées du dictionnaire, soit par l'affichage des informations selon les champs sélectionnés dans la requête.

3.4.3. Script PHP

Pour présenter le script du dictionnaire encyclopédique, nous reproduisons une partie du script consacré à la sélection des entrées du dictionnaire.

```
<p name = "N">  
  
Recherchez par <strong> entrée :</strong><br/>  
  
<select name="nom" value="Choisissez">  
  
<option value="" selected="selected"></option>  
  
<?php  
  
include("fonctions/connexion.php");  
  
$query=mysql_query("SELECT * FROM encyclopédie order by nom asc");  
  
$nb=mysql_num_rows($query);  
  
while ($list=mysql_fetch_array($query))  
  
{  
  
echo "<option>".$list['entree']."</option>";  
  
}  
  
mysql_close();  
  
?>  
  
</select>  
  
<br/>  
  
</p>  
  
<br/>
```

Conclusion

L'exploitation du couple PHP/MySQL nous a permis d'obtenir une interface d'interrogation exploitable en ligne. Elle permet à l'utilisateur d'interroger les dictionnaires électroniques de la médecine (PRED_DIC, AGRU_DIC, HUM_DIC) et d'obtenir des informations linguistiques et encyclopédiques sur les termes médicaux.

CHAPITRE 9

PRÉSENTATION DES RÉSULTATS

Introduction

L'arrivée de l'ordinateur a bouleversé le système du traitement de l'information, notamment l'information écrite. L'utilisation des bases de données est devenue un moyen incontournable pour le stockage et la consultation de l'information. Tout apprentissage ou exercice d'un domaine intellectuel nécessite une utilisation fréquente d'un dictionnaire. C'est le cas du domaine de la médecine qui est connu par sa complexité.

L'idée d'utilisation d'une base de données s'est imposée pour trouver un endroit où stocker toutes les données de nos dictionnaires et l'interface d'interrogation était pour nous le meilleur moyen de faciliter l'accès aux dictionnaires et ainsi faciliter la recherche. Dans ce cas, les recherches sont rapides et, surtout, ciblées.

1. Accueil

Dictionnaires de termes Médicaux

[Accueil](#) [Dictionnaires](#) [Contact](#)



Cette interface comporte deux types de dictionnaires. Il s'agit des dictionnaires de langue et d'un dictionnaire encyclopédique. Les dictionnaires de langue, en l'occurrence le dictionnaire des prédicats, le dictionnaire des arguments et le dictionnaire des humains, proposent une description spécifique des termes médicaux pour le traitement automatique. Ils comportent plus de 17500 entrées.

- Le dictionnaire des prédicats propose les définitions des termes selon des critères sémantiques, combinatoires et morphologiques.
- Le dictionnaire des arguments définit les entrées selon le genre, la morphologie, la classe sémantique, le domaine et le sous-domaine.
- Le dictionnaire des humains définit les entrées selon la morphologie, la classe sémantique, le domaine et le sous-domaine.
- Le dictionnaire encyclopédique définit les termes selon la classe sémantique, le domaine et le sous-domaine.



L'interface d'interrogation s'ouvre avec la page d'accueil. L'interface s'intitule dictionnaires des termes médicaux. Sa manipulation passe par trois icônes différentes sur lesquelles on peut cliquer. La première nommée *accueil* permet d'afficher le texte de présentation des dictionnaires. La deuxième icône nommée *dictionnaires* permet d'accéder à deux types de dictionnaire. Des dictionnaires linguistiques et un dictionnaire encyclopédique. La troisième icône nommée *contact* permet à l'utilisateur de contacter l'administrateur de l'interface pour s'inscrire ou demander des renseignements sur son fonctionnement.

2. Contact

Dictionnaires de termes Médicaux

Nous contacter

Nom
Votre nom

Mail
Votre mail

Message
Votre mail

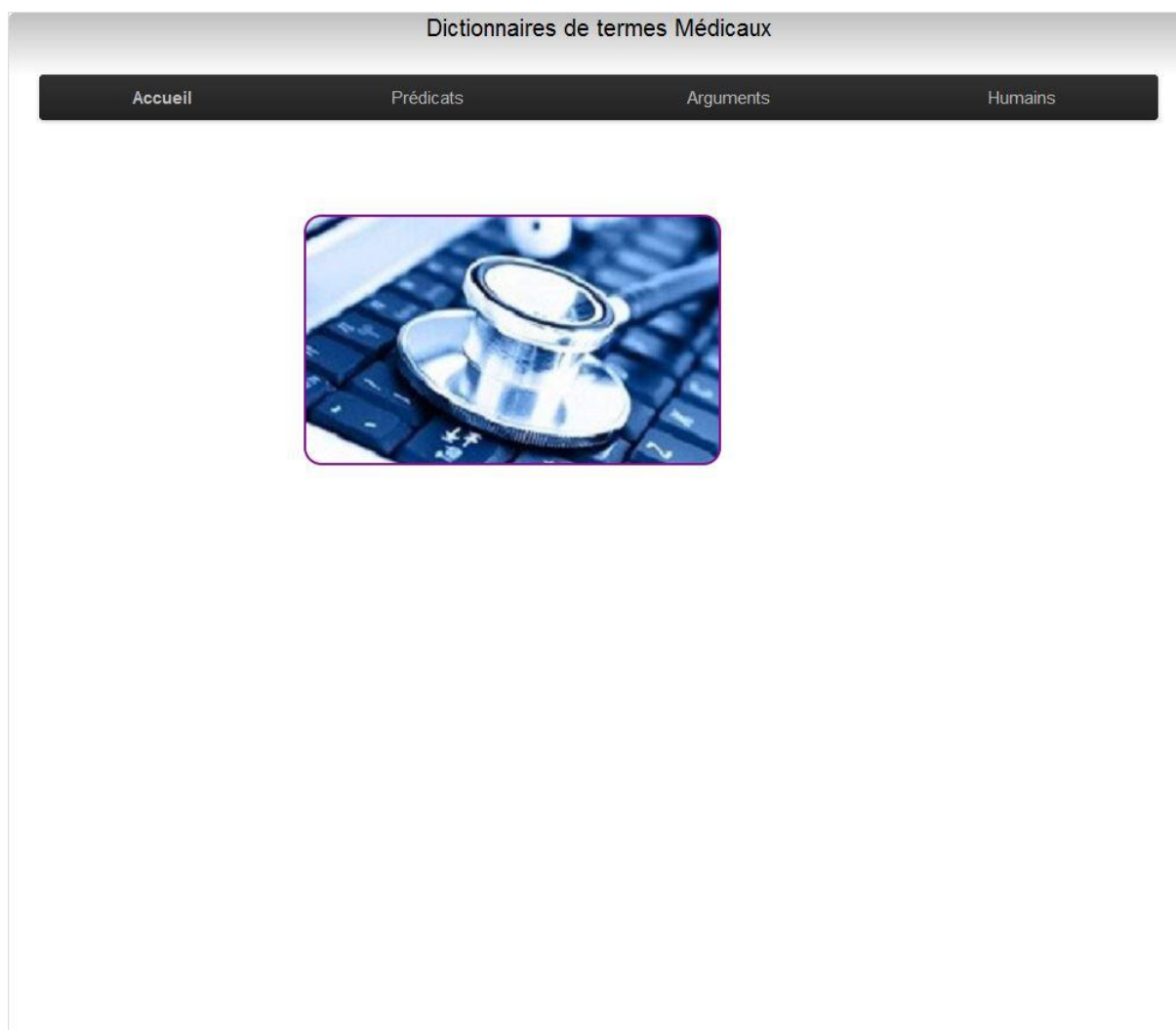
*2+2=? (Anti-spam)
2+2=?

[Envoyer](#) [Retour](#)

réservés. Copyright © 2013 Laboratoire LDI - Karim CHEBOUTI

Dans la page de contact, il est indispensable de remplir les champs indiqués pour pouvoir adresser un mail à l'administrateur. Il s'agit du nom, de l'adresse mail et du contenu du message.

3. Dictionnaires linguistiques



Cette page donne la possibilité de choisir entre trois dictionnaires linguistiques (le dictionnaire des prédicats, le dictionnaire des arguments et le dictionnaire des humains). Elle permet également de revenir sur la page d'accueil pour faire, éventuellement une nouvelle requête.

3.1. Dictionnaire des prédicats

3.1.1. Requête

Dictionnaires de termes Médicaux

Accueil
Prédicats
Arguments
Humains

Dictionnaire des prédicats

Recherche par **emploi prédicatif**

Recherche par **racine**

Recherche par **classe**

Recherche par **type**

Recherche par **aspect inhérent**

Recherche par **construction**

Recherche par **X0 morphosyntaxe**

Recherche par **préposition**

Recherche par **X1 morphosyntaxe**

Recherche par **X0 sémantique**

Recherche par **X1 sémantique**

Recherche par **morphologie**

Recherche par **domaine**

Recherche par **sous-domaine**

Le dictionnaire des prédicats est formé de quatorze colonnes, ainsi chaque colonne peut faire l'objet d'une requête. C'est-à-dire l'utilisateur peut afficher les résultats de ce dictionnaire d'une façon séparés (les entrées avec toutes leurs caractéristiques, les racines prédictives, les classes sémantiques, les types sémantiques, etc.) L'utilisateur peut choisir à sa guise s'il veut afficher les résultats d'une colonne entière ou bien d'une façon séparée. Dans ce cas de figure, il peut afficher, par exemple, toutes les classes qui figurent dans le dictionnaire des prédicats ou afficher seulement une classe particulière comme (maladie

contagieuse) tout en sachant qu'on peut combiner plusieurs colonnes afin d'afficher une recherche multicritère. Par exemple l'affichage d'une classe sémantique (maladie) du type état et d'une forme simple. La recherche multicritère peut se combiner selon le besoin de l'utilisateur.

3.1.2. Résultat

Dictionnaires de termes Médicaux

Dictionnaire Prédicat

Accueil | Prédicats | Arguments | Humains

Emploi prédecatif	Racine	Classe sémantique	Type	Aspet inherent	Construction
dépendance à la morphine		ADDICTION	eta	provisoire	X0 avoir DET N
dépendance à l'héroïne		ADDICTION	eta	provisoire	X0 avoir DET N
abus d'alcool		ADDICTION	eve	duratif imperfectif	X0 être A prep X1
abus d'amphétamine		ADDICTION	eve	duratif imperfectif	X0 être A prep X1
abus d'antidépresseur		ADDICTION	eve	duratif imperfectif	X0 être A prep X1
abus de barbituriques		ADDICTION	eve	duratif imperfectif	X0 être A prep X1
abus de cannabis		ADDICTION	eve	duratif imperfectif	X0 être A prep X1
abus de cocaïne		ADDICTION	eve	duratif imperfectif	X0 être A prep X1
abus de phencyclidine		EXAMEN_MEDICAL	act	duratif perfectif	X0 V sup DET N prep X1
				duratif	X0 V sup

Nous remarquons ici que le résultat concerne toutes les colonnes du dictionnaire, c'est-à-dire l'affichage des informations que comporte chaque terme. Cette première partie contient l'emploi prédecatif, la racine, la classe sémantique, le type, l'aspect inhérent et la construction.

Dictionnaires de termes Médicaux

Dictionnaire Prédicat

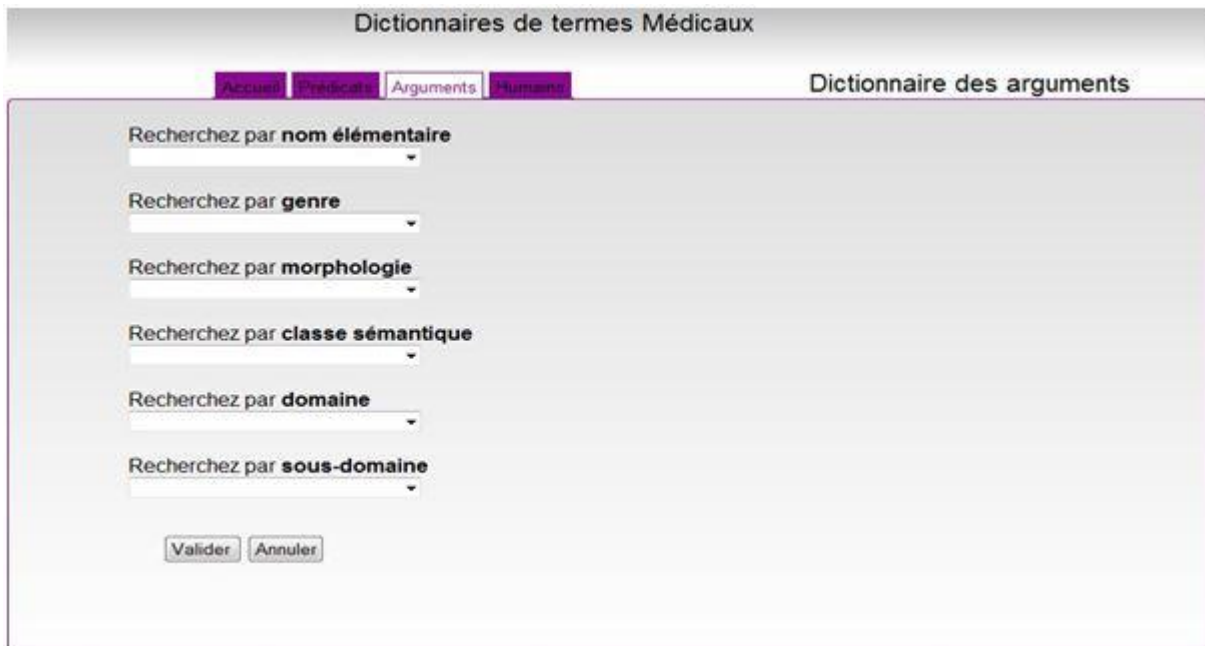
Accueil Prédicats Arguments Humains

X0 Morphosyntaxe	Préposition	X1 Morphosyntaxe	X0 Semantique	X1 Semantique	Morphologie	domaine
X0 = GN			X0 = humain		Mot composé[type:N à N]	Médecine
X0 = GN			X0 = humain		Mot composé[type:N à N]	Médecine
X0 = GN	pour	X1 = GN	X0 = humain	X1 = santé	Mot composé[type:N de N]	Médecine
X0 = GN	pour	X1 = GN	X0 = humain	X1 = santé	Mot composé[type:N de N]	Médecine
X0 = GN	pour	X1 = GN	X0 = humain	X1 = santé	Mot composé[type:N de N]	Médecine
X0 = GN	pour	X1 = GN	X0 = humain	X1 = santé	Mot composé[type:N de N]	Médecine
X0 = GN	pour	X1 = GN	X0 = humain	X1 = santé	Mot composé[type:N de N]	Médecine
X0 = GN	pour	X1 = GN	X0 = humain	X1 = santé	Mot composé[type:N de N]	Médecine
X0 = GN	à	X1 = GN	X0 = humain	X1 = humain	Mot composé[type:N de N]	Médecine
					Mot composé[type:N	

La deuxième partie de l’affichage des résultats comporte la distribution morphosyntaxique, la distribution sémantique, la morphologie, le domaine et le sous-domaine.

3.2. Dictionnaire des arguments

3.2.1. Requête



The screenshot shows a web interface for a medical term dictionary. At the top, there is a header 'Dictionnaires de termes Médicaux' and a navigation bar with tabs for 'Accueil', 'Prédicat', 'Arguments', and 'Humain'. The current page is titled 'Dictionnaire des arguments'. The main content area contains six search criteria, each with a dropdown menu: 'Recherchez par nom élémentaire', 'Recherchez par genre', 'Recherchez par morphologie', 'Recherchez par classe sémantique', 'Recherchez par domaine', and 'Recherchez par sous-domaine'. At the bottom of the search area, there are two buttons: 'Valider' and 'Annuler'.

Le dictionnaire des arguments comporte six champs, il s'agit de nom élémentaire, genre, morphologie, classe sémantique, domaine et sous-domaine. La recherche dans ce dictionnaire se fait selon le remplissage de ces champs.

3.2.2. Résultat

Dictionnaires de termes Médicaux

Dictionnaire Argument

Accueil Prédicats Arguments Humains

Nom élémentaire	Genre	Morphologie	Classe sémantique
<i>acupuncture</i>	<i>féminin</i>	<i>Mot dérivé</i>	<i>PROFESSION_SANTE</i>
<i>allergie et immunologie</i>	<i>féminin</i>	<i>Mot composé[type:N et NJ]</i>	<i>PROFESSION_SANTE</i>
<i>laboratoire d'analyses médicales</i>	<i>masculin</i>	<i>Mot composé[type:N de NA]</i>	<i>LIEU_SANTE</i>
<i>dents de Hutchinson</i>	<i>féminin pluriel</i>	<i>Mot composé[type:N de NJ]</i>	<i>PARTIE_CORPS_DENT</i>
<i>dents de lait</i>	<i>féminin pluriel</i>	<i>Mot composé[type:N de NJ]</i>	<i>PARTIE_CORPS_DENT</i>
<i>dents surnuméraires</i>	<i>féminin pluriel</i>	<i>Mot composé[type:NA]</i>	<i>PARTIE_CORPS_DENT</i>
<i>dents tachetées</i>	<i>féminin pluriel</i>	<i>Mot composé[type:NA]</i>	<i>PARTIE_CORPS_DENT</i>
<i>dent de Hutchinson</i>	<i>féminin</i>	<i>Mot composé[type:N de NJ]</i>	<i>PARTIE_CORPS_DENT</i>
<i>dent de sagesse</i>	<i>féminin</i>	<i>Mot composé[type:N de NJ]</i>	<i>PARTIE_CORPS_DENT</i>
<i>dent canine</i>	<i>féminin</i>	<i>Mot composé[type:NA]</i>	<i>PARTIE_CORPS_DENT</i>

Nous avons constaté que la recherche peut s'effectuer par le un seul ou plusieurs champs. Par exemple, l'utilisateur peut afficher les résultats par un seul champ comme l'affichage d'une seule entrée et une ou plusieurs de ses caractéristiques, comme il peut combiner la recherches selon les informations qu'il souhaite obtenir. Par exemple, afficher les termes d'une classe donnée, de formes simples, d'un genre particulier, etc. Dans ce cas, le résultat s'affiche selon la disponibilité des informations que comportent les entrées du dictionnaire.

Dictionnaires de termes Médicaux

Dictionnaire Argument

Accueil Predicats Arguments Humains

	Morphologie	Classe sémantique	Domaine	Sous-Domaine
	Mot dérivé	PROFESSION_SANTE	Médecine	Acupuncture
	Mot composé[type:N et N]	PROFESSION_SANTE	Médecine	Allergologie
	Mot composé[type:N de NA]	LIEU_SANTE	Médecine	Analyse médicale
riel	Mot composé[type:N de N]	PARTIE_CORPS_DENT	Chirurgie dentaire	Anatomie
riel	Mot composé[type:N de N]	PARTIE_CORPS_DENT	Chirurgie dentaire	Anatomie
riel	Mot composé[type:NA]	PARTIE_CORPS_DENT	Chirurgie dentaire	Anatomie
riel	Mot composé[type:NA]	PARTIE_CORPS_DENT	Chirurgie dentaire	Anatomie
	Mot composé[type:N de N]	PARTIE_CORPS_DENT	Chirurgie dentaire	Anatomie
	Mot composé[type:N de N]	PARTIE_CORPS_DENT	Chirurgie dentaire	Anatomie
	Mot composé[type:NA]	PARTIE_CORPS_DENT	Chirurgie dentaire	Anatomie

Cette partie comporte la suite des résultats en l'occurrence, les champs domaine et sous-domaine.

3.3. Dictionnaire des humains

3.3.1. Requête



The screenshot shows a web interface for a medical dictionary. At the top, there is a header 'Dictionnaires de termes Médicaux' and a navigation bar with tabs for 'Accueil', 'Prédicats', 'Arguments', and 'Humains'. The 'Humains' tab is selected, and the page title is 'Dictionnaire des humains'. Below the navigation, there are five search criteria, each with a dropdown menu: 'Recherchez par entrée', 'Recherchez par morphologie', 'Recherchez par classe sémantique', 'Recherchez par domaine', and 'Recherchez par sous-domaine'. At the bottom of the search area, there are two buttons: 'Valider' and 'Annuler'.

Le dictionnaire des noms d'humaines fonctionne exactement de la même façon que le dictionnaire des arguments. Ils ont le même mode de recherche. Le dictionnaire des noms d'humains comporte cinq champs, il s'agit d'entrée, morphologie, classe sémantique, domaine et sous-domaine. La recherche peut être combinée selon les besoins des utilisateurs.

3.3.2. Résultat

Dictionnaires de termes Médicaux

Dictionnaire Humain

Accueil Prédicats Arguments Humains

Entrée	Morphologie	Classe sémantique	Domaine	Sous-D
<i>personnel de laboratoire</i>	<i>Mot composé[type:N de N]</i>	<i>HUMAIN</i>	<i>Médecine</i>	<i>Analys médica</i>
<i>achondroplase</i>	<i>Mot dérivé</i>	<i>HUMAIN_MALADE</i>	<i>Médecine</i>	<i>Cancé</i>
<i>achondroplases</i>	<i>Mot dérivé</i>	<i>HUMAIN_MALADE</i>	<i>Médecine</i>	<i>Cancé</i>
<i>cancéreux</i>	<i>Mot dérivé</i>	<i>HUMAIN_MALADE</i>	<i>Médecine</i>	<i>Cancé</i>
<i>cancéreuse</i>	<i>Mot dérivé</i>	<i>HUMAIN_MALADE</i>	<i>Médecine</i>	<i>Cancé</i>
<i>cancéreuses</i>	<i>Mot dérivé</i>	<i>HUMAIN_MALADE</i>	<i>Médecine</i>	<i>Cancé</i>
<i>cancérologue</i>	<i>Mot dérivé</i>	<i>HUMAIN_MEDECIN</i>	<i>Médecine</i>	<i>Cancé</i>
<i>cancérologues</i>	<i>Mot dérivé</i>	<i>HUMAIN_MEDECIN</i>	<i>Médecine</i>	<i>Cancé</i>
<i>acataleptique</i>	<i>Mot dérivé</i>	<i>HUMAIN_MALADE</i>	<i>Médecine</i>	<i>Cardio</i>
<i>acataleptiques</i>	<i>Mot dérivé</i>	<i>HUMAIN_MALADE</i>	<i>Médecine</i>	<i>Cardio</i>
<i>cardiaque</i>	<i>Mot dérivé</i>	<i>HUMAIN_MALADE</i>	<i>Médecine</i>	<i>Cardio</i>
<i>cardiaques</i>	<i>Mot dérivé</i>	<i>HUMAIN_MALADE</i>	<i>Médecine</i>	<i>Cardio</i>
<i>cardiologue</i>	<i>Mot dérivé</i>	<i>HUMAIN_MEDECIN</i>	<i>Médecine</i>	<i>Cardio</i>
<i>cardiologues</i>	<i>Mot dérivé</i>	<i>HUMAIN_MEDECIN</i>	<i>Médecine</i>	<i>Cardio</i>

L'utilisateur de ce dictionnaire peut afficher chaque entrée du dictionnaire avec toutes ses caractéristiques, mais aussi une forme particulière, une classe sémantique donnée ou les noms d'humains d'un domaine particulier. Les champs du dictionnaire peuvent se combiner et afficher des résultats souhaités, c'est à dire selon les champs remplis.

4. Dictionnaire encyclopédique

4.1. Requête

The screenshot shows a web interface titled "Dictionnaires de termes Médicaux". At the top, there are two tabs: "Accueil" and "Dictionnaire encyclopédique", with the latter being selected. Below the tabs, the title "Dictionnaire encyclopédique" is displayed. The main content area contains four search criteria, each with a dropdown menu:

- Recherchez par **entrée**
- Recherchez par **classe sémantique**
- Recherchez par **domaine**
- Recherchez par **sous-domaine**

At the bottom of the search options, there are two buttons: "Valider" and "Annuler".

Le dictionnaire encyclopédique regroupe l'ensemble des termes médicaux des trois premiers dictionnaires, notamment le dictionnaire des prédicats, le dictionnaire des arguments et le dictionnaire des humains. Il comporte les champs suivants : *entrée*, *classe sémantique*, *domaine* et *sous-domaine*. Le dictionnaire encyclopédique est destiné aux personnes qui souhaitent obtenir uniquement des informations encyclopédiques (médicales) des termes médicaux.

4.2. Résultat

Dictionnaires de termes Médicaux

Accueil Dictionnaire encyclopédique Dictionnaire encyclopédique

Entrée	Classe sémantique	Domaine	Sous-domaine
<i>entrée</i>	<i>classe sémantique</i>	<i>domaine</i>	<i>sous-domaine</i>
<i>acupuncture</i>	<i>PROFESSION_SANTE</i>	<i>Médecine</i>	<i>Acupuncture</i>
<i>allergie et immunologie</i>	<i>PROFESSION_SANTE</i>	<i>Médecine</i>	<i>Allergologie</i>
<i>laboratoire d'analyses médicales</i>	<i>LIEU_SANTE</i>	<i>Médecine</i>	<i>Analyse médicale</i>
<i>dents de Hutchinson</i>	<i>PARTIE_CORPS_DENT</i>	<i>Chirurgie dentaire</i>	<i>Anatomie</i>
<i>dents de lait</i>	<i>PARTIE_CORPS_DENT</i>	<i>Chirurgie dentaire</i>	<i>Anatomie</i>
<i>dents surnuméraires</i>	<i>PARTIE_CORPS_DENT</i>	<i>Chirurgie dentaire</i>	<i>Anatomie</i>
<i>dents tachetées</i>	<i>PARTIE_CORPS_DENT</i>	<i>Chirurgie dentaire</i>	<i>Anatomie</i>
<i>dent de Hutchinson</i>	<i>PARTIE_CORPS_DENT</i>	<i>Chirurgie dentaire</i>	<i>Anatomie</i>
<i>dent de sagesse</i>	<i>PARTIE_CORPS_DENT</i>	<i>Chirurgie dentaire</i>	<i>Anatomie</i>
<i>dent canine</i>	<i>PARTIE_CORPS_DENT</i>	<i>Chirurgie dentaire</i>	<i>Anatomie</i>

Le mode de fonctionnement du dictionnaire encyclopédique est semblable à celui des dictionnaires de langue, c'est-à-dire que l'utilisateur affiche les résultats qu'il souhaite obtenir selon les champs du dictionnaire.

Conclusion

Ce chapitre est consacré à la présentation et aux commentaires des résultats obtenus de notre interface d'interrogation. Nous rappelons que nous avons exploité nos dictionnaires électroniques, notamment le dictionnaire des prédicats, le dictionnaire des arguments, le dictionnaire des humains et le dictionnaire encyclopédique du vocabulaire médical. Ces dictionnaires électroniques sont destinés au traitement automatique de la langue de la médecine. Pour des raisons pratiques, nous les avons exploités pour élaborer une interface d'interrogation afin de faciliter leur exploitation. Cette interface d'interrogation permet de faciliter la recherche des mots des dictionnaires. Il s'agit d'une recherche multicritère, notamment la recherche par des critères morphologiques, sémantiques, syntaxiques et encyclopédiques.

CONCLUSION GÉNÉRALE

A travers notre étude, nous avons tenté d'expliquer le fonctionnement linguistique du vocabulaire médical, du point de vue des trois fonctions primaires. Nous avons, ainsi, analysé les verbes, les noms et les adjectifs de la langue de la médecine.

Dans un premier temps, nous avons collecté plus de 17000 termes du vocabulaire médical. Nous les avons ensuite classés en trois grandes catégories différentes : les emplois prédicatifs (11300), les noms argumentaux (5580) et les noms d'humains (621). Nous les avons décrits pour élaborer des dictionnaires électroniques, notamment le dictionnaire des prédicats, le dictionnaire des arguments et le dictionnaire des humains. Nous avons affecté à chaque terme une classe sémantique. Pour ce faire, nous nous sommes basé sur les descriptions ontologiques de la langue de la médecine. Ainsi, le dictionnaire des prédicats contient 34 classes sémantiques. Elles regroupent les noms de maladies, les noms de maladies chroniques, les noms de maladies contagieuses, les noms de soins, les noms d'exams médicaux, etc. Le type sémantique *état*, *événement* et *action* est un autre descripteur attribué aux seuls emplois prédicatifs. L'aspect inhérent (duratif perfectif, duratif imperfectif, provisoire et ponctuel) exprime le temps interne de chaque emploi prédicatif. À chaque emploi prédicatif est associée la construction syntaxique la plus représentative du terme. Nous avons ainsi 25 constructions différentes. La distribution morphosyntaxique et la distribution sémantique sont d'autres informations syntactico-sémantiques associées à chaque emploi prédicatif. Nous avons également associé une catégorie morphologique à chaque terme. Nous avons dénombré 9679 mots composés, 1360 mots dérivés et 258 mots simples. Le classement par domaine (au nombre de 11) désigne le secteur d'activité du terme et le classement par sous-domaine (au nombre de 123) désigne la branche à l'intérieur du secteur d'activité, c'est à dire les spécialités médicales.

La terminologie du dictionnaire des prédicats est variée. Il contient environ 150 verbes, 277 adjectifs relatifs aux noms de maladies et 10873 substantifs qui sont généralement des noms de maladies.

Le dictionnaire des arguments contient uniquement des substantifs. Ils sont divisés en 47 classes sémantiques, 9 domaines et 68 sous-domaines. Les noms argumentaux diffèrent selon qu'ils appartiennent à la classe des parties du corps, des micro-organismes, ou du matériel médical, etc. La morphologie des arguments est très diversifiée, à ce niveau, nous avons recensé 189 formes différentes.

Le dictionnaire des noms d'humains médicaux contient 33 classes sémantiques différentes comme *humaine médecin* et *humain malade*. Du point de vue morphologique, ils sont variés. Nous avons dénombré 30 catégories morphologiques différentes dans le dictionnaire des humains. Chaque nom d'humain est associé à un domaine (au nombre de 5) ainsi qu'à un sous-domaine d'activité (au nombre de 36).

À partir des trois dictionnaires plurifonctionnels, notamment le dictionnaire des prédicats, le dictionnaire des arguments et le dictionnaire des humains, nous avons élaboré un dictionnaire encyclopédique pour que les utilisateurs obtiennent des informations encyclopédiques sur les termes médicaux.

Le dictionnaire encyclopédique regroupe tous les termes des trois premiers dictionnaires. Il comporte 17449 termes décrits par la classe sémantique (123), le domaine (17) et le sous-domaine (171).

Nous avons utilisé les informations que contiennent les entrées des dictionnaires afin de décrire et d'analyser le fonctionnement syntaxique et sémantique du vocabulaire médical, selon qu'ils correspondent à des emplois prédicatifs ou à des emplois argumentaux. Les prédicats et les arguments sont étudiés selon leurs propriétés morphologiques, syntaxiques et sémantiques. Le classement de la terminologie médicale a été effectué en termes d'emplois prédicatifs et d'emplois argumentaux.

Dans un second temps, nous avons élaboré une interface d'interrogation afin de rendre les dictionnaires plurifonctionnels pratiques. Cette application permet d'effectuer des recherches ciblées sur la terminologie médicale en un temps raccourci. Les dictionnaires donnent aux utilisateurs la possibilité d'obtenir des informations d'ordre sémantique, syntaxique, morphologique et encyclopédique. L'interface offre la possibilité d'effectuer des recherches selon des critères différents, c'est-à-dire qu'on peut choisir l'affichage des résultats des termes selon la classe sémantique, la morphologie ou la syntaxe. On peut également

afficher toutes les caractéristiques de toutes les entrées des dictionnaires à la fois. Pour résumer, on peut dire que l'interface offre les résultats selon le besoin de l'utilisateur.

A partir des dictionnaires électroniques que nous avons élaborés dans la cadre de la thèse, nous comptons développer une application dédiée à l'extraction d'information biomédicale. Cette application aboutira à un système de veille informatique à destination du grand public. L'application se basera sur la ressource existante composée des dictionnaires de prédicats et d'arguments et d'humains que nous devons compléter et enrichir. L'application doit récupérer sur le web des documents dans le domaine biomédical et en extraire des informations ciblées. Cette application affiche plusieurs objectifs :

- La mise en place d'une alerte épidémique
- Le relevé de nouvelles pathologies
- La découverte d'avancées thérapeutiques
- La surveillance sanitaire
- Le renvoi à des retours d'expérience de malades pertinents.

L'alerte épidémique doit permettre de l'apparition ou l'évolution d'une maladie. Elle sera basée sur la fréquence des mots dans les corpus. Elle permettra de prendre des mesures adéquates en cas d'apparition d'épidémie comme la grippe aviaire et la grippe porcine.

La découverte d'avancées thérapeutiques couvre aussi bien les nouvelles molécules, que les nouveaux médicaments, les nouveaux traitements et toute avancée aidant à la guérison des malades ou à la compréhension de leur maladie.

La surveillance sanitaire a un champ plus large, car elle peut aussi bien concerner l'industrie pharmaceutique avec l'établissement de nouveaux effets indésirables, que l'industrie de manière plus générale et les problèmes de santé publique comme lors de la découverte de la nocivité du Bisphénol A.

L'un des besoins des malades est de pouvoir recueillir des informations et des retours d'expérience de gens ayant vécu une épreuve similaire. La mise en relation des témoignages laissés sur des blogs par exemple permettrait d'apporter une information supplémentaire à celle fournie par le corps médical.

Conclusion générale

Cette application pourra, à la fois, extraire de nouveaux termes et les étiqueter afin d'enrichir la base de données. En faisant des requêtes sur des sites réservés à la santé comme les forums, les sites de vulgarisation, les sites qui proposent des terminologies médicales, etc.

BIBLIOGRAPHIE

- BAE, Hee Sook. Termes adjectivaux en corpus médical coréen : repérage et désambiguïsation. In *Terminology*, 2006, vol. 12, n° 1, p. 19-50
- BALLIU, Christian, 2001. Les traducteurs : ces médecins légistes du texte. In *Meta*, vol. 46, n° 1, p. 92-102
- BALLIU, Christian, 2005. La didactique de la traduction médicale, deux ou trois choses que je sais d'elle. In *Meta*, vol. 50, n° 1, p. 67-77
- BALLIU, Christian, 2010. Le traducteur, le médecin et le patient. In *Meta*, vol. 55, n° 1, p. 14-22
- BALLY, Charles, 1952. *Le langage et la vie*. 3^e éd. Genève : librairie Droz, 168 p. (Publications Romanes et Françaises ; 34)
- BLANCO, Xavier, 2000. Terminologie non nominale dans un dictionnaire électronique de médecine. In *Revue Informatique et Statistique dans les Sciences Humaines*, n° 26, p. 67-84
- BLANCO, Xavier, BUVET, Pierre-André, 2009. Présentation : les représentations des structures prédicat-arguments. In *Langages*, n° 176, p. 3-9
- BOISSEL, Jean-Pierre, DÜRRE, Françoise, 2005. Acception et acceptation : les professionnels de la santé devant les mots et syntagmes véhiculant de nouveaux concepts. In BELTRAN-VIDAL, Danièle, MANIEZ, François. *Les mots de la santé*. Lyon : Presses universitaires de Lyon, p. 11-21 (Travaux du CRTT)
- BOULANGER, Jean-Claude, 1978. Néologie et terminologie. In *Néologie en marche : Langue de spécialité, série b*, n° 4, 181 p.
- BUVET, Pierre André, LIM, Jung-Hae, 1996. Les déterminants nominaux aspectuels. In *Linguisticae Investigationes*, vol. 20, n° 2, p. 271-285
- BUVET, Pierre André, 1998. Déterminations et classe d'objets. In *Langages*, n° 131, p. 91-102
- BUVET, Pierre André, GIRARDIN, Chantal, GROSS, Gaston, GROUD, Claudette, 2005. Les prédicats d'affect. In *Lidl*, n° 32, p. 123-143

BUVET, Pierre-André, CARTIER, Emmanuel, ISSAC, Fabrice, MEJRI, Salah, 2007. Dictionnaires électroniques et étiquetage syntactico-sémantique. In *TALN-2007 & RECITAL-2007 : 14ème conférence sur le Traitement Automatique des Langues Naturelles, 5-8 juin 2007, Toulouse*. Toulouse : IRIT, p. 239-248

BUVET, Pierre André, 2008. Vers l'élaboration d'un dictionnaire unique des prédicats du français : DESSE. Dictionnaire Electronique Syntactico-Sémantique. In CONSTANT, Matthieu, DISTER, Anne, EMIRKANI, Louissette, PIRON, Sophie. *Description linguistique pour le traitement automatique du français*. Louvain-la-Neuve : Presses universitaires de Louvain, p. 23-42

BUVET, Pierre-André, 2009a. Quelles procédures d'étiquetage pour la gestion de l'information textuelle électronique ? In *L'information grammaticale*, n° 122, p. 40-48

BUVET, Pierre André, 2009 b. Des mots aux emplois : la représentation lexicographique des prédicats. In *Le Français Moderne*, vol. 77, n° 1, p. 83-96

BUVET, Pierre André, 2009. *Remarques sur la détermination du français*, pour l'habilitation à diriger des recherches, Université de Paris 13, 824 p.

BUVET, Pierre-André, 2012. Traitement Automatique du Discours Rapporté 2012. In *JADT 2012 : 11e Journée internationale d'analyse statistique des données textuelles, 13-15 juin 2012, Liège*. Liège : Université de Liège, p. 227-238

BUVET, Pierre André, 2011. Catégorisation sémantico-énonciative du lexique à partir d'un dictionnaire électronique. In CARDOSO, Suzana Alice Marcelino, MEJRI, Salah, MOTA, Jacyra Andrade. *Os di.ci.o.ná.rios : fontes, métodos e novas tecnologias*. [Brésil] : Instituto de Letras da Universidade Federal da Bahia, p. 75-96

BUVET, Pierre-André, NEVEU, Franck, 2013. *La dimension lexicale de la détermination en français*. Paris : Champion, 473 p. (Lexica. Mots et dictionnaires ; 26)

CAMPOS MARTIN, Natalia, 2009. Lexicographie et traduction (2) : macrostructure et microstructure des dictionnaires : les entrées, les articles et les définitions lexicographiques. In *Entreculturas*, n° 1, p. 351-364

CHARAUDEAU, Patrick, 1992. *Grammaire du sens et de l'expression*. Paris : Hachette, 927 p.

CHEVALLIER, Jacques, 1987. *Précis de terminologie médicale*. 5e éd. Paris : Maloine, 319 p.

COURSAGET-COLMERAUER, Colette, 1973. Les déterminants de la nominalisation. In *Cahier de linguistique*, n° 3, p 39-51

CUSIN-BERCHE, Fabienne. *Les mots et leurs contextes*. Paris : Presses de la Sorbonne Nouvelle, 2003. 202 p.

DEFRANCE, Jean-Marie, 2006. *PHP/MySQL avec Flash 8*. Paris : EYROLLES, 752 p. (Blanche)

DEGOULET, Patrick, FIESCHI, Marius, 1998. *Informatique médicale*. 3^e éd. Issy-les-Moulineaux : Masson, 246 p. (Abrégés)

DELBECQUE, Thierry, ZWEIGENBAUM, Pierre, 2005. Indexation UMLS en français : une expérience. In *11e Journées Francophones d'Informatique Médicale, 12-13 mai 2005, Lille*. Lille : Université de Lille 2 Droit et Santé, p. 2-8

DELISLE, Jean, 1997. Le métalangage de l'enseignement de la traduction d'après les manuels. In LEE-JAHNKE, Hannelore, DELISLE, Jean. *Enseignement de la traduction et traduction dans l'enseignement*. Ottawa : Presses de l'Université d'Ottawa, p. 185-246 (Regard sur la traduction)

D'HOUTAUD, Alphonse, 1998. *Sociologie de la santé : langage et savoirs, environnement et éthique*. Paris : L'Harmattan, 397 p. (Santé, sociétés et cultures)

DOMART, André, BOURNEUF, Jacques, 1992. *Nouveau Larousse Médical*. Paris : Larousse, 1142 p.

DOMENGEOT, Jacques, LE BEUX, Pierre, WEIL, Georges, 1994. *Informatisation de l'unité de soins du futur*. Paris : Springer-Verlag, p. 89-110. (Informatique et santé ; 7)

DUCOS, Joëlle, SALVADOR, Xavier-Laurent, 2011. Pour un dictionnaire de français scientifique médiéval : le projet Créalscience. In *Langages*, n° 183, p. 63-74

DUNHAM, George, GRAITSON, M., 1977. Traitement automatique du français médical. In *Cahiers de lexicologie*, vol. 30, n° 1, p. 47-74

EMIRKANIEN, Louise, FOUQUERE, Christophe, 2003. Présentation : TALN, Web et corpus. In *Revue Québécoise de Linguistique*, vol. 32, n° 1, p. 7-9

EVEILLARD, Philippe, 2002. Recherche d'information sur la toile (II) : conduites pratiques. In *Revue du Praticien : Médecine générale*, n° 596 et 597, p. 1873-1876

FAURE, Pascaline, 2010. Des discours de la médecine spécialisés multiples et variés à la langue médicale et unique et universelle. In *ASp*, n° 58, p. 73-86

FRANCŒUR, Aline, BRISEBOIS, Marie, 2001. Ressources documentaires médicales sur Internet : quantité, diversité et qualité. In *Meta*, vol. 46, n° 1, p. 128-144

FRYE, Curtis D., BURR, Emmanuelle, trad, CAMPILLO, Véronique, trad., 2007. *Excel 2007 : Microsoft*. Paris : Dunod, 378 p. (Étape par étape)

GALLIOT, Marcel, LAUBREAUX, Raymond, 1966. Le français langue vivante : grammaire complète. Toulouse : Privat-Didier, 374 p.

GAUTIER, Laurent, 2000. Du terme au phrasème en droit constitutionnel : un double aller-retour ? In GRECIANO, Gertrud (dir.). *Micro et macrolexèmes et leur figement discursif : études de linguistique comparée français/allemand*. Louvain-la-Neuve : Peteers, p. 284-296. (Bibliothèque de l'Information Grammaticale ; 43)

GEHENOT, Daniel, 1976. Le sigle : aperçu linguistique. In *Babel*, vol. 22, n° 3, p. 125-131

GEHENOT, Daniel, 1976. Le sigle : aperçu linguistique II. In *Babel*, vol. 22, n° 4, p. 173-177

GEHENOT, Daniel, 1977. Le sigle : aperçu linguistique III. In *Babel*, vol. 23, n° 1, p. 32-38

GEORGE, Kenneth E.M., 1977. La siglaison et les dérivés de sigles en français contemporain. In *Le Français moderne*, vol. 45, n° 1, p. 33-42

GHAZI, Joseph, 1985. *Vocabulaire du discours médical : structure, fonction et apprentissage*. Paris : Didier Erudition, 464 p. (Linguistique ; 16)

GORCY, Gérard, 1990. Le Trésor de la Langue Française (TLF) : son originalité et les voies ouvertes par son informatisation. In Institut National de la langue française. *Autour d'un dictionnaire : le « Trésor de la Langue Française »*. Paris : Didier Erudition, p. 187-207 (Dictionnaire et Lexicographie; 1)

GREZKA, Aude, BUVET, Pierre-André, 2007. Elaboration d'outils méthodologiques pour décrire les prédicats du français. In *Lingvisticae Investigationes*, vol. 30, n° 2, p. 217-245

GROSS, Gaston, VIVES, Robert, 1986. Les constructions nominales et l'élaboration d'un lexique grammairal. In *Langue française*, n° 69, p. 5-27

GROSS, Gaston, 1995. Pour une typologie de la détermination nominale. In *Linguisticae Investigationes supplementa*, n° 20, p. 161-171

GROSS, Gaston, 1996. *Les expressions figées en français, les noms composés et autres locutions*. Essentiel. Paris : Ophrys, 161 p. (L'Essentiel français)

GROSS, Gaston, MATHIEU-COLAS, Michel, 2001. Description de la langue de la médecine. In *Meta*, vol. 46, n° 1, p. 68-81

GROSS, Gaston, GUENTHNER, Franz, 2002. Comment décrire une langue de spécialité ? In *Cahiers de lexicologie*, vol. 80, n° 1, p. 179-199

GROSS, Gaston, 2008. Sur le statut syntaxique des substantifs humains. In LEEMAN, Danielle. *Des topoï à la théorie des stéréotypes en passant par la polyphonie et l'argumentation dans la langue : hommages à Jean-Claude Anscombe*. Chambéry : Université de Savoie, p. 27-41

GROSS, Gaston, 2012. *Manuel d'analyse linguistique : approche sémantico-syntaxique du lexique*. Villeneuve d'Ascq : Presses Universitaires de Septentrion, 370 p.

GROSS, Maurice, 1981. Les bases empiriques de la notion de prédicat sémantique. In *Langages*, n° 63, p. 7-52

GROUIN, Cyril, ROSIER, Arnaud, DAMERON, Olivier, ZWEIGENBAUM, Pierre, 2009. Une procédure d'anonymisation à deux niveaux pour créer un corpus de comptes rendus hospitaliers. In *13^e Journées Francophones d'Informatique Médicale, 28-30 avril 2009, Nice*. Paris : Springer-Verlag, p. 23-34

HAMBURGER, Jean., 1982. *Introduction au langage de la médecine*. Paris : Flammarion, 202 p.

HAMON, Albert, 1964. *Grammaire française : cycle d'observation 6^e-5^e*. Paris : Hachette, 191 p.

HARRIS, Zellig Sabbetai, GROSS Maurice, trad., 1976 *Notes du cours de syntaxe*. Paris : Le Seuil, 237 p.

HARRIS, Zellig Sabbetai, 1988. *Language and Information*, New-York : Columbia University Press, 120 p.

HATHOUT, Nabi, 2001. Analogies morpho-synonymiques : une méthode d'acquisition automatique de liens morphologiques à partir d'un dictionnaire de synonymes. In *TALN 2001-RECITAL 2001 : 8e conférence sur le traitement automatique des langues naturelles et 5e*

rencontre des étudiants chercheurs en informatique pour le traitement automatique des langues, 2-5 juillet 2001, Tours. Tours : Université de Tours, p. 223-232

JOUSSE, Anne-Laure, TREMBLAY, Ophélie, 2006. Le dictionnaire électronique : un outil précieux pour enseigner le lexique. In *Québec français*, n° 141, p. 115-116

KLEIBER, Georges, 1991. *Points de vue sur l'anaphore.* In *Journée SCICOGNEST, 8 Novembre 1990, Strasbourg.* Strasbourg : Groupe Anaphore et Deixis, (65 p).

KLEIBER, Georges, 1994. Anaphore associative, antécédent et définitude. In SCHNEDECKER, Catherine, CHAROLLES, Michel, KLEIBER, Georges, DAVID, Jean. *L'anaphore associative (aspects linguistiques, psycholinguistiques et automatiques).* Metz : Université de Metz, p. 153-173 (Recherches linguistiques ; 19)

KLEIBER, Georges, 2001. *L'anaphore associative.* Paris : PUF, 386 p. (Linguistique nouvelle)

KORTH, Henry F., SILBERSCHATZ, Abraham, 1988. *Systèmes de gestion des bases de données.* Bogota : McGraw-Hill, 546 p.

LEE-JAHNKE, Hannelore, 2001. L'enseignement de la traduction médicale : un double défi. In *Meta*, vol. 46, n° 1, p. 145-153

LE PESANT, Denis, 2000. *Six études de sémantique lexicale sur les noms locatifs*, décembre 2000, pour l'habilitation à diriger les recherches, Université de Paris 13, 231 p.

LE PESANT, Denis, 2002. La détermination dans les anaphores fidèles et infidèles. In *Langages*, n° 145, p. 39-59

LERAT, Pierre, 1995. *Les langues spécialisées.* Paris : Presses universitaires de France, 201 p.

LERAT, Pierre, 1997. Approches linguistiques des LSP. In *ASp*, n° 15-18, p. 1-10

MANIEZ, François, 2001. La traduction du nom adjectival en anglais médical. In *Meta*, vol. 46, n° 1, p. 56-67

MARCHELLO-NIZIA, Christiane, 1979. La notion de « phrase » dans la grammaire. In *Langue française*, n° 41, p. 35-48

- MARTIN, Robert, 1983. *Pour une logique du sens*. Paris : Presses Universitaires de France, 268 p.
- MARTINET, André, 1966. Le mot. In BENVENISTE, Emile. *Problèmes du langage*. Paris : Gallimard, p. 39-53 (Diogène)
- MARTINET, André, 1970. *Eléments de linguistique générale*. Paris : Armand Colin, 221 p.
- MARTINEZ, Camille, 2012. *L'orthographe des dictionnaires français : la construction de la norme graphique par les lexicographes*. Paris : Honoré champion, 648 p. (Lexica : Mots et Dictionnaires ; 22)
- MASSOUSSI, Taoufik, 2009. Transferts sémantiques et langues spécialisées. In *Cahiers de lexicologie*, n° 94, p. 107-135
- MATHIEU-COLAS, Michel, 1995. Syntaxe du trait d'union : structures complexes. In *Linguisticae Investigationes*, vol. 19, n° 1, p. 153-171
- MATHIEU-COLAS, Michel, 1996. Essai de typologie des noms composés français. In *Cahiers de lexicologie*, n° 69, p.71-125
- MATHIEU-COLAS, Michel, 1998. Illustration d'une classe d'objets : les voies de communication. In *Langages*, n° 131, p.77-90
- MATHIEU-COLAS, Michel, BUVET, Pierre-André, 1999. Les champs 'domaine' et 'sous-domaine' dans les dictionnaires électroniques. In *Cahiers de Lexicologie*, n° 75, p. 173-191
- MATHIEU-COLAS, Michel, 2009. MORFETIK : une ressource lexicale pour le TAL. In *Cahiers de lexicologie*, n° 94, p. 137-146
- MATORE, Georges, 1968. *Histoire des dictionnaires français*. Paris : Larousse, 278 p. (La langue vivante)
- MEJRI, Salah, 1995. *La néologie lexicale*. Tunis : Faculté des lettres de la Manouba, 379 p. (Linguistique, IX)
- MEJRI, Salah, 1997. *Le figement lexical : descriptions linguistiques et structuration sémantique*. Tunis : Faculté des lettres de la Manouba, 632 p.
- MEJRI, Salah, 2005. Figement, néologie et renouvellement du lexique. In *Linx*, n° 52, p. 163-174.

- MEJRI, Salah, 2008. La traduction des textes spécialisés : le cas des sciences du langage. In *Traduire : un métier d'avenir. Actes du colloque du 50e anniversaire de l'ISTI, 14-15 octobre 2005, Bruxelles*. Bruxelles : Editions du Hazard, p. 117-144
- MEJRI, Salah, 2009. Le mot, problématique théorique. In *Le Français Moderne*, vol. 77, n° 1, p. 68-82
- MEJRI, Salah, 2011. Figement, collocation et combinatoire libre. In MEJRI, Salah, ANSCOMBRE, Jean-Claude. *Le figement linguistique : la parole entravée*. Paris : Champion, p. 63-77 (Lexica. Mots et dictionnaires ; 18)
- MEJRI, Salah, 2011. Phraséologie et traduction des textes spécialisés. In GONZALEZ ROYO, Carmen, MOGORRON HUERTA, Pedro. *Fraseología contrastiva : lexicografía, traducción y análisis de corpus*. Alicante : Universidad de Alicante. Servicio de publicaciones, p. 125-138
- MEUNIER-CRESPO, Mariette, 1997. Les locutions nominales dans les dictionnaires de spécialités. In *Meta*, vol. 42, n° 1, p. 68-71
- MEYER, Eric A., 2007. *Conception de site Web avec les CSS : comprend des exercices et des vidéos de formation*. Londres : Pearson CampusPress, 480 p.
- MICHARD, Alain, 2000. *XML : langage et applications*. Paris : Eyrolles, 499 p.
- MOGORRON HUERTA, Pedro, MEJRI, Salah, 2008. *Las construcciones verbo-nominales libres y fijas. Aproximación contrastiva y traductológica*. Alicante : Universidad de Alicante, 221 p. (Encuentros Mediterráneos ; 1)
- MONIN, Sylvie, 1996. Termes éponymes en médecine et application pédagogique. In *ASp*, n° 11-14, p. 217-237
- MOULAY, Nadia, 1996. *Etude de la collocation lexicale dans une langue de spécialité en français : la sémiologie médicale*, pour l'obtention du diplôme d'études approfondies de sciences du langage, Université Paris 13, 197 p.
- MULLER, G., HERLICOVIEZ, M., 1982. Les grossesses prolongées. In CHAINE, Caroline, dir. *Encyclopédie médico-chirurgicale : Obstétrique. T. III*. Paris : éditions techniques, p. 77-106
- MURCIA, Colette, JOLY, Hubert, 2005. *Dictionnaire de mots nouveaux des sciences et des techniques 1982-2003*. Paris : CILF, 812 p.

MURRAY TURBAYNE, Colin, 1972. Metaphors for the mind. In RUDNER, Richard, ISRAEL Scheffler. *Logic and Art : essays in Honor of Nelson Goodman*. Indianapolis and New-York : The Boobs Merrill Compo, p. 61

NAKOS, Dorothy, 1994-1995. Les images en médecine: une perception vive et précise des réalités. In *ALFA*, n° 7-8, p. 269-282

NAMER, Fiammetta, 2005. Automatiser les définitions des termes médicaux: qu'est ce que le traitement automatique du langage apporte à la théorie morphologique? In *Journées Francophones d'Informatique Médicale, 12-13 mai 2005, Lille*. Lille : Université de Lille 2 Droit et Santé, p

NEVEU, Franck, 2004. *Dictionnaire des sciences du langage*. Paris : Armand Colin, 316 p.

PAJAUD, Daniel, 1990. *La taxinomie bionaturaliste 2 : Taxinomie et taxilogie, oppositions et congruences des concepts en ontologie*. Paris : Gaston Lachurie, 341 p.

PETIT, Gérard, 2008. Dénomination et figement. In BLUMENTHAL, Peter, MEJRI, Salah. *Les séquences figées entre langue et discours*. Stuttgart : Franz Steiner, p. 131-144 (Zeitschrift für französische Sprache und Literatur ; 36)

PRUVOST, Jean, 1997. *Les dictionnaires de la langue française et l'informatique : actes du colloque « La Journée des dictionnaires », 1995, Cergy-Pontoise*. Cergy-Pontoise : Centre de Recherche Texte-Histoire, 161 p.

PRUVOST, Jean, 1997. Les illustrations dictionnaires : d'une typologie et d'une terminologie pour le « dictionnaire papier » à une réflexion pour les « dictionnaires électroniques ». In *Les dictionnaires de langue française et l'informatique. Actes du colloque « La Journée des dictionnaires », 1995, Cergy-Pontoise*. Cergy-Pontoise : Centre de Recherche Texte-Histoire, p. 137-161

PRUVOST, Jean, 2000. *Dictionnaires et nouvelles technologies*. Paris : Presses Universitaires de France, 192 p. (Écritures électroniques)

PRUVOST, Jean, 2006. *Les dictionnaires français, outils d'une langue et d'une culture*, Paris : Ophrys, Paris, 200 p. (L'essentiel français)

QUEMADA, Bernard, 1968. *Les dictionnaires du français moderne 1539-1863*. Paris : Didier, 683 p.

- QUEMADA, Bernard, INSTITUT NATIONAL DE LA LANGUE FRANCAISE, 1990. *Autour d'un dictionnaire : Le "trésor de la langue française" : témoignages d'atelier et voies nouvelles*. Paris : Didier Erudition, 207 p. (Dictionnaire et lexicographie ; 1)
- REGENT, Odile, 1992. Pratique de communication en médecine : contextes anglais et français. In *Langages*, n° 105, p. 66-75
- REY-DEBOVE, Josette, 1971. *Etude linguistique et sémiotique des dictionnaires français contemporains*. The Hague : Mouton, 329 p.
- RIGAUX, Philippe, 2005. *Pratique de MySQL et PHP*. 3^e éd. Sebastopol : O'Reilly, 624 p.
- ROUBAKINE, A.N., 1974 *Dictionnaire médical français-russe : environ 45.000 termes*. Moscou : Encyclopédie soviétique, 616 p.
- ROULEAU, Maurice, 1994. La traduction médicale. Une approche méthodique. Brossard : Linguatex, 326 p.
- ROULEAU, Maurice, 2001. La facture des principaux dictionnaires médicaux français : point de vue d'un traducteur. In *Meta*, vol .46, n°1, p. 34-55
- SASTRE ALAIZ, Judith, 2007. *Dictionnaire électronique de catalan coordonné avec le français : le module NooJ*. Thèse de doctorat en langues et cultures romanes. Barcelone : Université autonome de Barcelone, 182 p.
- SAUSSURE, Ferdinand de, BALLY, Charles, SECHEHAYE, Albert, RIEDLINGER, Albert, 1965. *Cours de linguistique générale*. Paris : Payot, 331 p. (Bibliothèque scientifique)
- SAUSSURE, Ferdinand de, 1967. *Cours de linguistique générale*. Edition critique préparée par Tullio de Mauro. Paris : Payot, 520 p.
- SCHMITT, Christopher, GUERIN, James, trad., 2005. *CSS en action : des solutions pour implémenter les feuilles de style*. Paris : O'Reilly, 280 p.
- SFAR, Inès, 2007. *Les racines prédicatives entre équivalence et rupture d'emplois*. Thèse de doctorat en sciences du langage. Villetaneuse : Université Paris 13, 353 p.
- SKVORTSOV, Guennadi, 1994. Le français médical. In *Cahiers du CERTIL*, n° 4, p. 33-37
- SOURNIA, Jean-Charles, 1997. *Langage médical français*. Toulouse : Privat, 157 p. (Culture & professions de santé)

- SPILLNER, Bernd, 1992. Textes médicaux français et allemands. Contribution à une comparaison interlinguale et interculturelle. In *Langages*, n° 105, p. 42-65
- THOIRON, Philippe, 1991. Texte spécialisé et non-spécialiste : un problème de stratégie de décodage. In *Revue belge de philologie et d'histoire*, vol. 69, n° 3, p. 629-643
- TRÂN Duc Tuan, 1999. *La standardisation de la terminologie médicale vietnamienne : une approche socioterminologique*. Thèse de doctorat en sciences du langage. Rouen : Université de Rouen, 531 p.
- TROMEUR, Laurent, 2011. *Mise en place d'une interface en langue naturelle pour la plateforme Ontomantics*. Thèse de doctorat en sciences du langage. Villetaneuse : Université Paris 13, 253 p.
- ULLMANN, Stephen, 1957. *Style in the French novel*. Cambridge : the University press, 273 p.
- VANDAELE, Sylvie. Utilisation des bases de données bibliographiques spécialisées en traduction médicale. In *Meta*, 2001, vol. 46, n° 1, p. 103-116
- VAN HOOFF, Henri, 2001. La traduction des éponymes médicaux banalisés de langue anglaise. In *Meta*, vol. 46, n° 1, p. 82-91
- VAN LANCKER, Luc, 2008. *XHTML et CSS : les nouveaux standards du code source*. Saint-Herblain : ENI, 434 p. (Ressources informatiques)
- VENDRYES, Joseph, 1953. Le mot dans la phrase. In *Le français moderne*, 1953, vol. 21, n° 2, p. 81-90
- VERREAULT, Jean-Pierre, 1990. Les grossesses à risque. In *Journées médicales de l'hôpital du Saint-Sacrement [Québec], 1990, Laval*. Laval : Université de Laval, p.131-140
- YORK, Richard, 2007. *Débuter avec le CSS pour le webdesign*. Paris : First Interactive, 640 p.
- ZOLONDEK, Debbie, 1991. La siglaison : observations en télématique et en médecine. In *Terminogramme*, n° 62, p. 1-5

Dictionnaires

DUBOIS, Jean, LAGANE, René, NIOBEY, Georges, CASALIS, Didier, CASALIS, Jacqueline, MESCHONNIC, Henri, 1966. *Dictionnaire du français contemporain*. Paris : Larousse, 1224 p.

JEUGE-MAYNART, Isabelle, éd. FLORENT, Jacques, éd., 2012 *Le Petit Larousse illustré : en couleurs : 90000 articles, 5000 illustrations, 355 cartes, 125 planches, chronologie universelle*. [Paris] : le Grand livre du mois, XXXII-1934 p.

LAROUSSE, 1976. *Petit Larousse illustré : Dictionnaire encyclopédique pour tous*. Paris : Librairie Larousse, 1799 p.

ROBERT, Paul, REY, Alain, REY-DEBOVE, Josette, 1983. *Le Petit Robert : dictionnaire alphabétique et analogique de la langue française*. Nouvelle éd. revue, corrigée et mise à jour pour 1983. Paris : Le Robert, XXXI-2171 p.

REY, Alain, REY-DEBOVE, Josette, 1990. *Le Petit Robert : dictionnaire alphabétique et analogique de la langue française*. Nouv. éd. rev. corr. et mise à jour en 1990. Paris : Le Robert, XXXI-2171 p.

REY, Alain, REY-DEBOVE, Josette, 2010. *Le Nouveau Petit Robert : Dictionnaire alphabétique et analogique de la langue française*. Nouvelle édition. Paris : Le Robert, XLII-2837 p.

REY, Alain, TOMI, Marianne, HORDE, Tristan, TANET, Chantal, 1992. *Dictionnaire historique de la langue française : contenant les mots français en usage et quelques autres délaissés, avec leur origine proche et lointaine [...]*. Paris : Le Robert, 2 vol. XXI-2383 p.

REY, Alain, TOMI, Marianne, HORDE, Tristan, TANET, Chantal, 1995. *Dictionnaire historique de la langue française : contenant les mots français en usage et quelques autres délaissés, avec leur origine proche et lointaine [...]*. Nouvelle édition. Paris : Le Robert, 2 vol. XXI-2383 p.

REY, Alain, TOMI, Marianne, HORDE, Tristan, TANET, Chantal, 2012. *Dictionnaire historique de la langue française : contenant les mots français en usage et quelques autres délaissés, avec leur origine proche et lointaine [...]*. Nouvelle éd. augmentée par Alain Rey. Paris : Dictionnaires Le Robert, 3 vol. XVII-4168 p.

Webographie

Bibliographie

ATILF (pages consultées le 19 juin 2013). *Le Trésor de la Langue Française Informatisé*, [en ligne]. Adresse URL : <http://atilf.atilf.fr/>

CHAMBERLIN, Donald D., BOYCE, Raymond F (pages consultées le 6 juin 2013). *Structured Query Language*, [en ligne].

Adresse URL : http://fr.wikipedia.org/wiki/Structured_Query_Language

CISMeF (pages consultées le 28 janvier 2011). *Portail Terminologique de Santé*, [en ligne].

Adresse URL : www.chu-rouen.fr/ssf/arborescences.html

PLANTEVIT, Marc (page consultée le 25 juin 2013). *PL/SQL*, [en pdf].

Adresse URL : http://liris.cnrs.fr/~mplantev/documents/C5_PL:SQL_8:page.pdf

VIDONNE, Louis (pages consultées le 15 mai 2013). *Base Initiation à l'utilisation de SQL avec OOo Basic dans Base*, [en pdf].

Adresse URL : http://www.openoffice.org/fr/Documentation/How-to/Bdd/IntroSQL_Base.pdf

INDEX DES TERMES

A

Actualisateur : 22, 58, 75, 80, 81, 84, 85, 147, 181

Adjectif approprié : 171

Argument : 24, 25, 26, 28, 31, 38, 59, 60, 75, 81, 87, 88, 141, 145, 170, 178, 179, 180, 189, 207, 244, 260

B

Base de données : 17, 66, 216, 217, 218, 219, 236, 238, 239, 240, 253, 267, 286

C

Classe sémantique : 29, 34, 37, 86, 97, 99, 104, 106, 107, 127, 128, 174, 179, 214, 219, 244, 257, 259, 261, 262, 263, 264, 272, 273, 275, 278, 279, 280

Collocation : 16, 67, 144, 145, 146, 208, 213

Combinatoire : 177, 178

Composition : 27, 31, 40, 41, 42, 43, 49, 75, 102, 103, 109, 115, 116, 117, 118, 160, 161, 163, 183, 189, 190, 194, 208

D

Dictionnaire : 17, 84, 227, 228

E

Emploi argumental : 18, 38, 76, 149

Emploi prédicatif : 18, 24, 25, 31, 33, 34, 36, 38, 76, 87, 88, 121, 122, 142, 149, 198, 243, 255, 257, 273, 283

F

Figement : 16, 17, 27, 28, 77, 163, 208

H

Humain : 15, 16, 17, 18, 21, 33, 36, 37, 38, 57, 61, 71, 84, 87, 99, 104, 132, 166, 174, 176, 188, 189, 190, 194, 195, 197, 198, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 236, 240, 261

I

Inférence : 18, 23, 77, 209

M

Morphologie : 37, 40, 82, 97, 99, 100, 102, 104, 106, 115, 124, 146, 151, 152, 159, 161, 177, 191, 193, 195, 219, 236, 244, 245, 255, 257, 258, 259, 261, 262, 274, 275, 278, 284

N

Noms appropriés : 206, 209, 212

P

Polymorphie : 16, 18, 27, 28, 99, 124, 125

Polysémie : 16, 24, 25, 27, 29, 32, 56, 61, 75, 76, 77, 79, 138, 139

V

Verbes appropriés : 17, 83, 109, 131, 136, 155, 173, 177, 200, 207, 210, 212, 213

Verbes supports : 58, 59, 85, 89, 125, 126, 137, 143, 168, 210

RÉSUMÉ

LE VOCABULAIRE MÉDICAL DU POINT DE VUE DES TROIS FONCTIONS PRIMAIRES

Cette thèse s'inscrit à la fois dans le cadre de la théorie des trois fonctions primaires qui s'inspire de la théorie distributionnaliste de Z. Harris (la fonction prédicative, la fonction argumentale et la fonction actualisatrice) et dans la lignée des travaux qui ne distinguent pas la langue générale des langues spécialisées, du point de vue de leur fonctionnement syntaxique et sémantique. La terminologie médicale est décrite, ici, dans le cadre de la phrase. Nous avons analysé les termes médicaux selon leurs emplois prédicatifs ou argumentaux. Cette analyse a donné lieu à l'élaboration de bases de données. Une interface d'interrogation a été élaborée, par la suite, afin de permettre aux utilisateurs potentiels d'effectuer des recherches ciblées.

ABSTRACT

VOCABULARY MEDICAL PERSPECTIVE OF THREE PRIMARY FUNCTIONS

This thesis falls within the framework of the Three Primary Functions theory based on Z. Harris's distributionalist theory (predicative function, argumental function and actualizing function), as well as in the line of work that does not distinguish between general and specialized languages considering their syntactic and semantic functioning. The medical terminology is described here in the context of the sentence. We analyzed the medical terms according to their argumental or predicative employment. This analysis led to the development of databases. A query interface has been developed subsequently, to allow potential users to conduct targeted research.

DISCIPLINE

SCIENCES DU LANGAGE

MOTS CLÉS

prédicat, argument, terminologie médicale,

LABORATOIRE D'ACCUEIL : LEXIQUES, DICTIONNAIRES, INFORMATIQUE

LDI-CNRS-UMR 7187

Université Paris 13

99, Avenue de Jean Baptiste Clément

F-93430 Villetaneuse