

Article

« Description de la langue de la médecine »

Gaston Gross et Michel Mathieu-Colas

Meta : journal des traducteurs / Meta: Translators' Journal, vol. 46, n° 1, 2001, p. 68-81.

Pour citer cet article, utiliser l'information suivante :

URI: <http://id.erudit.org/iderudit/002976ar>

DOI: 10.7202/002976ar

Note : les règles d'écriture des références bibliographiques peuvent varier selon les différents domaines du savoir.

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter à l'URI <https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche. Érudit offre des services d'édition numérique de documents scientifiques depuis 1998.

Pour communiquer avec les responsables d'Érudit : info@erudit.org

Description de la langue de la médecine

GASTON GROSS

MICHEL MATHIEU-COLAS

Université de Paris XIII – Nord, Villetaneuse, France

RÉSUMÉ

Le traitement automatique des langues de spécialité exige une description qui rende compte de son fonctionnement. Le travail descriptif s'appuie sur la notion de phrase et l'établissement de la liste de tous les prédicats, ainsi que sur la détermination des arguments pour situer les emplois. L'article présente les diverses étapes de description illustrées par de nombreux exemples.

ABSTRACT

Computer processing of specialized languages requires a description of the process. Such a description must encompass the notion of "sentence," and identify predicates and appropriate arguments. Using several examples, this article outlines the various stages of the description.

MOTS-CLÉS/KEYWORDS

langue de la médecine, traitement automatique, prédicat, argument, codage catégoriel

1. Terminologie et traitement automatique

C'est dans les langues de spécialité que se manifeste actuellement le besoin le plus évident de traitement automatique. Les objectifs sont variés: recherche d'informations dans des sources documentaires et en particulier sur le Web, veille technologique, traduction assistée par ordinateur et, dans le meilleur des cas, traduction automatique. Les enjeux économiques sont plus clairs et plus immédiatement perceptibles que dans le cas de la langue générale. Or, il y a comme un hiatus entre cette demande de description des langues techniques et la réponse à cette demande que fournissent les travaux en terminologie. La plupart des lexiques ne sont pas en mesure de rendre compte du fonctionnement d'une langue de spécialité avec la précision nécessaire à des fins informatiques.

Ce constat est la conséquence du statut même de la terminologie telle que l'ont voulue ses créateurs (entre autres Wüster). Cette conception de base tourne le dos à la nature même des langues. Pour ces auteurs, le langage n'est que le reflet d'un découpage préalable existant dans les choses. Le lien entre les deux est le « terme » qui n'est qu'une dénomination, une étiquette collée à l'élément du monde qu'on a besoin d'isoler. Une conception du langage comme simple reflet du réel a des conséquences insoupçonnées. Comme dans la nature il existe des objets mais non des relations, on comprend qu'on n'ait pris en compte dans les langues que des substantifs et, dans le domaine technique, des suites plus ou moins figées créées pour dénommer des objets nouveaux. Voilà pourquoi, si on examine un dictionnaire technique, on est en présence d'une liste plus ou moins longue de noms composés.

Cette position théorique entraîne deux types de conséquences pour le traitement automatique.

D'une part, tout texte, même le plus spécialisé, comprend pour deux tiers de sa surface des mots de la langue générale. C'est la raison pour laquelle les traductions automatiques sont encore si déficientes : presque tous les termes de la langue générale sont polysémiques et, pour reconnaître et rendre compte de chaque signification, on doit être en mesure de décrire l'emploi auquel on a affaire. Il est donc impossible de traduire un texte spécialisé en n'ayant pas une description fine de la langue générale. Observons simplement qu'une langue aussi étudiée que le français n'a pas encore reçu une description capable de rendre compte de tous les emplois des verbes et à plus forte raison des autres prédicats, noms et adjectifs (nous reviendrons sur ce point). De plus, il n'existe pas de limite claire entre la langue générale et les langues spécialisées. Au demeurant, aucun système n'est en mesure de reconnaître si l'on est en présence de l'une ou de l'autre. Cette interférence n'est pas un des moindres obstacles au traitement automatique des textes. Il suffit de jeter un coup d'œil sur un texte technique donné pour se convaincre de ce qui vient d'être dit.

D'autre part, les auteurs d'inspiration wusterienne ont totalement négligé le fait que les langues techniques sont des « langues » et, comme telles, ne se réduisent pas à un vocabulaire. Elles forment un tissu constitué, comme tout texte, par des phrases, qui sont les unités constitutives de tout message linguistique. Or, les phrases sont formées de prédicats et d'arguments qui doivent être définis les uns par rapport aux autres, le rôle des prédicats étant fondamental. Voilà pourquoi les travaux d'inspiration wusterienne ne sont pas en mesure de rendre compte du fonctionnement des textes, lesquels nécessitent des descriptions linguistiques beaucoup plus précises et méthodiques que celles qui s'adressent à un cerveau humain, capable de pallier de façon indirecte les informations manquantes ou approximatives.

Les langues de spécialité comportent des prédicats qui leur sont propres et qui doivent figurer dans un dictionnaire du domaine. Il est clair qu'une description de la langue de l'informatique ne peut se passer de recenser des termes comme *saisir* dans *saisir (un texte, des données)*. Ce verbe, polysémique par ailleurs, est beaucoup plus difficile à décrire que *formater*, qui lui n'a qu'un seul sens. Dans toute langue de spécialité, on est amené à expliquer, raisonner, démontrer, illustrer. Les éléments lexicaux qui décrivent ces opérations ne sont pas des reflets de découpages qui existent dans la réalité. Leur rôle dans la phrase est le même que leurs équivalents de la langue générale. Aucun traitement automatique d'un texte technique donné n'est possible si on en cache la complexité et si on le considère comme fondamentalement différent de la langue générale.

2. Un concept commun : la notion de phrase

Ce qui manque dans la tradition terminologique, c'est l'idée que les mots ne sont pas des briques isolées mais entrent dans des phrases où ils prennent leur sens. C'est un fait connu qu'un mot comme « porte » ne peut pas être interprété hors contexte mais aura un statut différent s'il est précédé d'un groupe nominal ou de l'article : dans le premier cas, c'est un verbe et, dans l'autre, un nom. On fera remarquer que cette information est évidente, mais elle ne l'est pas pour le traitement automatique. Cette observation sur la catégorie grammaticale n'est qu'un exemple simple montrant que

les mots ne peuvent pas être décrits de façon autonome. Nous suggérons que le niveau d'analyse minimal ne peut pas être le mot : une entrée de dictionnaire doit être une phrase.

Il va de soi que si on veut savoir le sens d'un verbe comme *prendre*, il est impossible de le dire si on ne précise pas ses arguments : *prendre un steak* n'est pas *prendre de l'aspirine*, ni *prendre le bus*, ni *prendre une autoroute*, ni *prendre une gifle*, etc. Il convient donc, dans un premier temps, de donner une définition de la notion de phrase qui vaudra à la fois pour la langue générale et les langues de spécialité. Le mot *phrase* signifiera ici *phrase minimale*, ce que la tradition grammaticale a appelé *proposition*, et non une combinaison de propositions, comme c'est très souvent le cas dans les textes. Le passage de la phrase minimale à des textes requiert des descriptions différentes mettant en jeu des relateurs, des effacements, des relations anaphoriques, etc.

On définira la phrase simple comme un prédicat accompagné de ses arguments. C'est ce qu'on appelle un schéma d'arguments, dans lequel le prédicat est suivi par son sujet et son (ou ses) complément(s) : *rédiger (humain, texte)*. On voit que la nature sémantique des compléments dépend de celle du prédicat. Si on avait le prédicat *épeler*, l'objet ne serait pas *texte* mais *mot*.

Ce schéma doit être explicité en fonction de plusieurs paramètres. Tout d'abord, le prédicat ne doit pas être identifié au verbe. À la place de *rédiger*, nous aurions pu avoir *rédaction : rédaction (humain, texte)*. Une phrase doit ensuite être actualisée, c'est-à-dire inscrite dans le temps, ce qui revient à conjuguer le prédicat et à déterminer les arguments. La conjugaison du prédicat dépend de sa catégorie grammaticale. Un verbe est conjugué par des désinences propres et des verbes auxiliaires. Un prédicat nominal est « conjugué » par un type spécial de verbes appelés *verbes supports* : *faire (un voyage)*, *tirer (une conclusion)*, *prendre (une décision)*, *mener (un combat)*. Les adjectifs prédicats sont inscrits dans le temps à l'aide du verbe *être*. La description d'une langue a comme premier objectif de dresser la liste de tous les prédicats de cette langue et d'attribuer à chacun d'eux les substantifs qu'il peut avoir dans ses positions argumentales.

3. La notion d'emploi

Une constatation qui apparaît immédiatement, c'est que la plupart des prédicats sont polysémiques. Nous en avons donné un exemple plus haut avec le verbe *prendre*. Il se trouve que la notion de polysémie, qui semble pourtant claire, ne rend pas compte de la réalité du phénomène qui est ainsi évoqué. Revenons au verbe *prendre*. Si on remplace dans *prendre le bus* le complément *bus* par *train* (ou *autorail*, *métro*, *tramway*), alors le sens du verbe ne change pas. Si, au contraire, on utilise *autoroute* (ou *nationale*, *départementale*), alors nous aurons un sens différent. Quand donc on veut définir la signification d'un prédicat donné, il faut être en mesure d'établir des classes d'arguments. On appellera *emploi* une occurrence de prédicat avec des classes spécifiques de substantifs en position d'arguments. On observera qu'un prédicat, en cas de polysémie, ne peut pas avoir les mêmes classes d'arguments.

Cependant, le terme de *polysémie*, bien qu'il soit d'interprétation aisée et admise de tous, ne rend pas compte dans sa totalité du phénomène qu'il prétend décrire. Un prédicat polysémique n'a pas seulement plus d'un sens : ces « sens » sont aussi caractérisés par des propriétés différentes. Revenons à notre verbe *prendre*. On a déjà observé

que *prendre* n'a pas le même sens selon que le complément est un <aliment> ou un <médicament>. Mais ces deux « emplois » ont d'autres différences, comme la nominalisation. Dans l'emploi « médical », le verbe *prendre* a une forme nominale associée (*prendre un médicament, la prise d'un médicament*), ce qui n'est pas le cas dans l'emploi « alimentaire » (**la prise d'une tartine*). Cette notion nous semble une des plus importantes de la linguistique. En résumé, retenons que pour décrire un emploi, il faut être en mesure de reconnaître, automatiquement ou non, ses classes d'arguments.

4. Classer les arguments

Un des objectifs de la description linguistique en vue du traitement automatique est de réduire la polysémie pour être en mesure de reconnaître les emplois. La seule façon de le faire est de se servir de la nature sémantique des arguments. Il existe quatre types de prédicats de ce point de vue.

Il y a d'abord une première classe de prédicats qui n'imposent pas de restrictions à leurs arguments : ils les acceptent tous. C'est le cas, par exemple, de verbes comme *réfléchir* à l'égard de son complément : tout substantif peut en être le complément. Ces verbes, qui ne sont pas très fréquents, n'ont pas d'emplois différenciés et ne sont donc pas polysémiques.

Un deuxième groupe de verbes sélectionnent parmi tous les substantifs de grandes classes sémantiques, que nous appelons des *traits* et qui correspondent aux traits syntactico-sémantiques classiques. Ces classes sont au nombre de six, si l'on ne tient pas compte de la récursivité, c'est-à-dire des prédicats qui peuvent fonctionner en position d'arguments :

- Humain : *Paul a travaillé.*
- Animal : *Tous les animaux ont crevé.*
- Végétal : *Cette plante (germe, se fane).*
- Concret : *Nous avons déplacé ce caillou.*
- Locatif : *Cette route mène de Paris à Versailles.*
- Temps : *Cette cérémonie a duré deux heures.*

Ces classes constituent un premier outil pour lever les ambiguïtés qui caractérisent les prédicats. Ainsi, le verbe *réduire* peut-il être divisé en emplois différents grâce à ces traits : humain (*réduire un ennemi, réduire quelqu'un en esclavage*), concret (*réduire ce caillou en morceaux*), locatif (*réduire l'intervalle entre les deux repères*), temps (*réduire la durée du trajet*).

Un troisième groupe de prédicats est caractérisé par un spectre plus réduit. Un des emplois du verbe *nommer* peut être décrit grâce aux traits que nous avons énumérés plus haut comme ayant trois arguments humains, *nommer* (*hum, hum, hum*) correspondant à des phrases comme *On a nommé Paul président*. Mais cette description de *nommer*, si elle rend compte de phrases correctes comme celles que nous venons de donner, peut aussi générer des phrases fausses comme **On a nommé Paul voisin*, dont les trois arguments sont également humains. On voit donc que les traits que nous avons indiqués n'ont pas la précision voulue pour rendre compte de la construction de cet emploi. Ici, le deuxième complément ne représente pas n'importe quel humain : il désigne une sous-classe d'humains qu'on pourrait nommer

<fonctions> (*président, député, adjoint, délégué, etc.*). Nous appelons ces sous-classes des *classes d'objets*.

Si nous dressons la liste de tous les substantifs de <fonction> et si, d'autre part, nous avons une liste des substantifs humains, alors nous sommes en mesure de générer toutes les phrases possibles avec cet emploi de *nommer*. Il va de soi que le système ne confondra plus cet emploi du verbe avec un autre : *Ses parents l'ont nommé Paul*, car le substantif *Paul* n'est pas un nom de fonction. Voici un autre exemple. Le verbe *porter* peut avoir pour objet un concret (*Paul porte un seau de 10 kilos*). Cet emploi se traduit en anglais par le verbe *to carry*. Mais il se trouve que cette traduction ne vaut pas pour tous les concrets. Quand l'objet est un <vêtement>, qui est pourtant un concret aussi, la traduction de *porter* est *to wear*. Si donc on veut avoir la bonne traduction, il ne suffit pas d'avoir le trait <concret>, il faut aussi disposer de la classe d'objets <vêtements>.

Nous ne rappelons ici que pour mémoire le quatrième type de prédicats qui sélectionne, dans une position argumentale donnée, un seul substantif, comme c'est le cas d'*abaisser* dans *abaisser une perpendiculaire*. Il va de soi que ce que nous venons de dire pour la langue générale s'applique aussi aux langues techniques.

Résumons les observations que nous venons de faire. Comme les langues techniques sont des langues et non pas des lexiques seulement, l'unité minimale d'analyse est la phrase. Décrire ce type de langue, c'est être capable d'énumérer toutes les phrases qu'on est en mesure de générer. Cela signifie qu'on recense tous les prédicats, qu'on dispose de toutes les classes sémantiques d'arguments pour lever les ambiguïtés, c'est-à-dire pour séparer les emplois. Si, de plus, on décrit ces classes sémantiques en extension, alors on peut générer toutes les phrases possibles dans cette langue.

5. Classer les prédicats

Dans la description d'une langue à des fins de traitement automatique, on peut faire un pas supplémentaire. Nous venons de voir que les classes d'arguments permettent de reconnaître automatiquement l'emploi d'un prédicat dans un texte donné parmi tous les emplois potentiels de ce prédicat. Cette reconnaissance ne signifie peut-être pas grand-chose pour un lecteur ou un traducteur humains, mais c'est une étape d'une importance capitale pour le traitement automatique, d'autant que cela permet d'envisager une génération automatique de toutes les phrases possibles relevant d'un emploi déterminé, du fait du recensement systématique des éléments lexicaux de chaque classe d'objets. On voudrait faire le même travail sur les prédicats.

Une première étape importante consiste à montrer que la notion de prédicat ne se réduit pas à la classe des verbes. Une même racine prédicative peut avoir trois réalisations morphologiques (*il désire s'en aller, il a le désir de s'en aller, il est désireux de s'en aller*). Si l'on veut décrire les langues de spécialité, il est nécessaire, dans un premier temps, de dresser la liste pour chacune d'elles de tous les prédicats verbaux, nominaux et adjectivaux. Cette tâche n'est pas triviale. Nous avons considéré que les trois formes de la racine prédicative *désir-* étaient équivalentes du point de vue de la prédicativité. Ce fait ne repose pas sur des considérations morphologiques seulement (les racines sont les mêmes) mais aussi sur des critères syntaxiques (ces trois formes ont les mêmes arguments et les mêmes propriétés aspectuelles). Mais cela ne vaut pas

pour des mots reliés morphologiquement comme *peur* et *peureux*. Le premier est un nom de sentiment, qui peut avoir un complément phrastique (*il a peur de sortir*) et son aspect peut être ponctuel et duratif, alors que l'adjectif *peureux* désigne un trait de caractère, qu'il n'a pas de complément et que son aspect est seulement duratif. On voit donc que les relations entre prédicats verbaux, nominaux et adjectivaux ne sont pas simples et doivent être examinées au regard de toutes les propriétés linguistiques.

Par ailleurs, indépendamment de leur forme, les prédicats se différencient par leur nature sémantique. Même si ces classes sont déjà un peu anciennes, on a l'habitude de diviser les prédicats en *actions*, *événements* et *états*. Cette classification a un intérêt évident du point de vue sémantique et pour la reconnaissance des constructions syntaxiques. En effet, un même prédicat a une syntaxe spécifique selon qu'il a une lecture « active » ou « événementielle » : *Paul a réagi vivement* ; *il y a eu une réaction vive de la part de Paul*. Nous voudrions justifier les raisons d'un tel codage dans le cas des prédicats nominaux, qui se comportent différemment des verbes. Alors que, pour ces derniers, il est pratiquement impossible de faire une corrélation entre le sens et le type de conjugaison, c'est la règle pour les substantifs prédictifs. Prenons les prédicats nominaux d'action. On admettra que leur actualisation se fait généralement avec le verbe support *faire* : *Paul a voyagé*, *Paul a fait un voyage*. Mais tous les prédicats d'action n'ont pas ce support. Il est donc nécessaire de faire des sous-classes d'action pour prédire leur actualisation : les prédicats de <crimes> prennent *commettre* ou *perpétrer*, les <cris> prennent *pousser*, les <ordres> *donner* ou *intimer*, les <combats> *mener* ou *livrer*. On voit donc que, si on veut « conjuguer » tous les prédicats nominaux, on est obligé de dresser la liste des classes sémantiques qui les caractérisent. Il se pourrait que pour les actions et les événements il y en ait plusieurs centaines.

6. Étapes de la description

Si l'on envisage de décrire une langue de spécialité, il faudra procéder par les étapes suivantes, si l'on part d'un corpus informatisé. La première tâche consistera à établir la liste des formes du corpus et à lemmatiser les mots. Puis, on codera les parties du discours, et on séparera les prédicats des arguments. On décrira les classes de prédicats et leurs constructions. On établira ensuite les substantifs arguments en traits et classes d'objets, de sorte que l'on sera en mesure de décrire les prédicats avec précision, c'est-à-dire de reconnaître les emplois. On obtiendra ainsi deux types de dictionnaires, celui des prédicats et celui des arguments, dont la structuration sera différente mais complémentaire. Nous prenons comme exemple la langue de la médecine.

6.1. Nomenclature et lemmatisation

Il convient en premier lieu d'établir la liste des unités lexicales du domaine à partir de l'ensemble des données disponibles : connaissances personnelles, dictionnaires, ouvrages spécialisés, corpus divers. Il est évident que les informations accessibles sur le Web occuperont, dans ce dispositif, une place de plus en plus importante. S'agissant de la sélection des unités, on notera que les textes spécialisés juxtaposent trois types de vocabulaires :

- a) des unités de la langue générale, non spécifiques du domaine (« cette *question* mérite d'être discutée... », « nous *reviendrons* sur ce *thème* quelques *pages* plus loin », « au *moment* où *paraît* cet *article*... »);
- b) des éléments spécifiques du domaine: *hôpital* (*hospitaliser*, *hospitalisation*) *anesthésie* (*anesthésier*, *anesthésiant*, *anesthésique*), *cautériser* (*cautérisation*), etc.;
- c) des éléments du vocabulaire général mais affectés d'un sens particulier dans le domaine considéré: des noms comme *conduit*, *sillon*, *voies* (en anatomie), *circulation* (en physiologie), des verbes comme *admettre* (un malade dans un hôpital), *prendre* (un médicament), *se réveiller* (en parlant d'une douleur), des adjectifs comme *aigu*, *fonctionnel*, *indiqué*, *moteur*.

Seuls les éléments des types b) et c) peuvent être considérés comme caractéristiques de la spécialité.

La liste doit être ensuite lemmatisée, ce qui implique la réduction de toutes les formes fléchies à leur forme canonique: verbes à l'infinitif, noms au singulier, adjectifs au masculin singulier. Par exemple:

admettre	aigu
anesthésie	cautériser
conduit	fonctionnel
hospitalisation	indiqué
moteur	prendre
réveiller (se)	sillon

Les mots composés, naturellement, reçoivent le même traitement:

anesthésie de base	angine aiguë
anesthésie de contact	appendicite aiguë
anesthésie en selle	arthrite aiguë
anesthésie épidurale	bronchite aiguë
anesthésie générale	chorée aiguë
anesthésie locale	décubitus aigu
anesthésie loco-régionale	diabète aigu
anesthésie oro-trachéale	eczéma aigu
anesthésie péridurale	épiglottite aiguë
anesthésie régionale	glaucome aigu
anesthésie segmentaire	glomérulonéphrite aiguë
anesthésie visuelle	goitre aigu

6.2. Codage des catégories

On procède ensuite à l'identification des parties du discours, principalement les noms, les verbes et les adjectifs (éventuellement les adverbes). Les autres catégories (articles, prépositions, conjonctions) constituent essentiellement des outils grammaticaux et sont rarement spécifiques des langues spécialisées (voir cependant le cas du droit: prépositions comme *par devant*, adverbiaux comme *en foi de quoi*). On obtient ainsi autant de fichiers que de catégories morphologiques:

Noms:

abaisse-langue/nm	abcès anorectal/nm
abaisse-paupière/nm	abcès vertébral/nm
abaisseur de la lèvre/nm	abdomen/nm
abaisseur du sourcil/nm	abducteur/nm

abaque de Meiners/nm	abduction/nf
abasie/nf	aberration/nf
abcès/nm	aberration autosomique/nf

Adjectifs :

abdominal	adipeux
abortif	adrénergique
acromial	adsorbant
acromien	aérique
actif	agglutinant
activant	agnosique
adhésif	aigu

Verbes :

accoucher	admettre
administrer	agir
agoniser	alimenter
aliter	amputer
analyser	anémier
anesthésier	ankyloser
apaiser	appareiller

6.3. Dégrouperment des catégories

L'étape suivante consiste à identifier, pour chaque catégorie, le fonctionnement syntaxique. Les noms doivent être traités différemment selon qu'il s'agit de prédicats ou d'arguments. On notera que les substantifs humains peuvent relever des deux sous-ensembles.

Les adjectifs donnent également lieu à des sous-catégorisations. On distinguera notamment les adjectifs prédicatifs (*malade*) et les adjectifs de relation pouvant jouer le rôle d'arguments (*pulmonaire*): *examiner le poumon, examen du poumon, examen pulmonaire*. Nous regroupons dans une classe particulière les adjectifs causatifs (*diurétique, anti-inflammatoire*).

Quant aux verbes, ils peuvent théoriquement remplir plusieurs fonctions (prédicats, auxiliaires, verbes supports, éléments d'expressions figées). Mais ce sont surtout les verbes prédicatifs qui apparaissent comme caractéristiques des langues de spécialité.

6.4. Dictionnaire des arguments

Les substantifs arguments, comme on l'a vu plus haut, gagnent à être catégorisés en traits et en classes d'objets, en fonction des prédicats qui les sélectionnent. Dans les exemples qui suivent, on adoptera le codage élémentaire suivant :

- G: catégorie grammaticale (nm, nf, nmf)
- T: trait syntaxique¹
- C: classe d'objets

Humains non prédicatifs

Noms propres

Ambroise Paré/G :nm/T :hum

Pasteur/G :nm/T :hum

Docteur X (le)/G :nm/T :hum

Humains collectifs

équipe chirurgicale/G :nf/T :hum coll

corps médical/G :nm/T :hum coll

Conseil de l'ordre des médecins/G :nm/T :hum coll

Animaux

Animaux d'expérimentation

cobaye/G :nm/T :ani/C :ani. d'expér

souris/G :nf/T :ani/C :ani. d'expér

Végétaux

Plantes médicinales

agnus castus/G :nm/T :vég/C :plante méd

aloès officinal/G :nm/T :vég/C :plante méd

anémone hépatique/G :nf/T :vég/C :plante méd

aspérule odorante/G :nf/T :vég/C :plante méd

Micro-organismes

Protozoaires

amibe dysentérique/G :nf/T :micro/C :protoz

trypanosome/G :nm/T :micro/C :protoz

Bactéries

brucella/G :nf/T :micro/C :bactér

pneumocoque/G :nm/T :micro/C :bactér

bacille de Koch/G :nm/T :micro/C :bactér

Clostridium perfringens/G :nm/T :micro/C :bactér

Virus

adénovirus/G :nm/T :micro/C :virus

cytomégalovirus/G :nm/T :micro/C :virus

virus de la grippe/G :nm/T :micro/C :virus

virus de l'hépatite B/G :nm/T :micro/C :virus

Champignons microscopiques

actinomycète/G :nm/T :micro/C :champi

blastomycète/G :nm/T :micro/C :champi

botryomycète/G :nm/T :micro/C :champi

levure/G :nf/T :micro/C :champi

moisissure/G :nf/T :micro/C :champi

Concrets

Noms de parties du corps (npc)

tête/G :nf/T :inc/C :npc

nez/G :nm/T :inc/C :npc

main/G :nf/T :inc/C :npc

ventre/G :nm/T :inc/C :npc

<os>

radius/G :nm/T :inc/C :npc<os>
 cubitus/G :nm/T :inc/C :npc<os>
 tibia/G :nm/T :inc/C :npc<os>
 péroné/G :nm/T :inc/C :npc<os>

<muscles>

abaisseur de la lèvre/G :nm/T :inc/C :npc<muscle>
 biceps/G :nm/T :inc/C :npc<muscle>
 moyen fessier/G :nm/T :inc/C :npc<muscle>
 muscle cardiaque/G :nm/T :inc/C :npc<muscle>

<vaisseaux>

artère fémorale/G :nf/T :inc/C :npc<vaisseau>
 cubito-palmaire/G :nf/T :inc/C :npc<vaisseau>
 sous-clavière/G :nf/T :inc/C :npc<vaisseau>
 veine cave/G :nf/T :inc/C :npc<vaisseau>

<nerfs>

nerf crânien/G :nm/T :inc/C :npc<nerf>
 nerf oculaire/G :nm/T :inc/C :npc<nerf>
 grand hypoglosse/G :nm/T :inc/C :npc<nerf>
 glosso-pharyngien/G :nm/T :inc/C :npc<nerf>

<glandes>

glande lacrymale/G :nf/T :inc/C :npc<glande>
 thyroïde/G :nf/T :inc/C :npc<glande>
 hypophyse/G :nf/T :inc/C :npc<glande>

Liquides organiques

sang/G :nm/T :inc/C :liquide org.
 lymphe/G :nf/T :inc/C :liquide org.
 salive/G :nf/T :inc/C :liquide org.
 sueur/G :nf/T :inc/C :liquide org.

Hormones

hormone antidiurétique/G :nf/T :Npc/C :hormone
 contra-insuline/G :nf/T :Npc/C :hormone
 vaso-pressine/G :nf/T :Npc/C :hormone
 progestérone/G :nf/T :Npc/C :hormone

Substances (divers)

acide lactique/G :nm/T :inc/C :substance
 calcium/G :nm/T :inc/C :substance
 glucose/G :nm/T :inc/C :substance
 urée/G :nf/T :inc/C :substance

Médicaments

suppositoire/G :nm/T :inc/C :médic
 antihistaminique/G :nm/T :inc/C :médic
 sirop contre la toux/G :nm/T :inc/C :médic
 gouttes pour le nez/G :nfp/T :inc/C :médic

Instruments d'observation

amnioscope/G :nm/T :inc/C :instrum obs
 stéthoscope/G :nm/T :inc/C :instrum obs

Instruments de mesure

thermomètre/G :nm/T :inc/C :instrum mes
asthmomètre/G :nm/T :inc/C :instrum mes
pèse-urine/G :nm/T :inc/C :instrum mes
toise/G :nf/T :inc/C :instrum mes

Instruments chirurgicaux

scalpel/G :nm/T :inc/C :instrum chir
bistouri/G :nm/T :inc/C :instrum chir
porte-nitrate/G :nm/T :inc/C :instrum chir
serre-nœud/G :nm/T :inc/C :instrum chir

Contenants

armoire à pharmacie/G :nf/T :inc/C :contenant
trousse médicale/G :nf/T :inc/C :contenant

Locatifs

hôpital/G :nm/T :loc
centre de soins/G :nm/T :loc
pharmacie/G :nf/T :loc
cabinet médical/G :nm/T :loc

Parties de lieux

bloc opératoire/G :nm/T :loc/C :partie loc
salle d'attente/Gnf/T :loc/C :partie loc
salle d'opération/Gnf/T :loc/C :partie loc
salle de réanimation/Gnf/T :loc/C :partie loc
salle de garde/Gnf/T :loc/C :partie loc

Noms de temps

heures de consultation/G :nfp/T :tps
période d'incubation/G :nf/T :tps
temps de saignement/G :nm/T :tps
convalescence/G :nf/T :tps
congé de longue maladie/G :nm/T :tps

Adjectifs de relation

Adjectifs de localisation anatomique

abdominal/G :adj/C :loc anat
artériel/G :adj/C :loc anat
dentaire/G :adj/C :loc anat
fémoral/G :adj/C :loc anat
nasal/G :adj/C :loc anat
sous-cutané/G :adj/C :loc anat

6.5. Dictionnaire des prédicats

Ce dictionnaire décrit les verbes, les adjectifs et les noms prédicatifs, en indiquant dans chaque cas le nombre et la construction des arguments, ainsi que leur nature sémantique. Dans la mesure où les prédicats forment le noyau de la phrase, leur description est fondamentale pour le traitement des langues de spécialité. Nous reprenons ici, en l'appliquant à la médecine, la distinction habituelle en *actions*, *événements* et *états* dont nous parlions plus haut, en y ajoutant les « humains prédicatifs » (noms de fonction, de relation, etc.).

Outre la catégorie grammaticale, les traits et les classes (champs G, T, C), nous précisons quand il y a lieu les types d'arguments (N0 désigne le sujet, N1 et N2 les compléments éventuels). Pour les prédicats nominaux, il est fait mention du verbe support (champ W).

Prédicats d'action

Consultations

consulter/G :v/C :consult/N0 :hum/N1 :médecin/N2 :pour maladie
 aller voir/G :v/C :consult/N0 :hum/N1 :médecin/N2 :pour maladie

Examens

radio/G :nf/T :act/C :exam/N0 :médecin/N1 :de Npc/N2 :à hum/W :faire
 échographie/G :nf/T :act/C :exam/N0 :médecin/N1 :de Npc/N2 :à hum/W :faire
 (Avec partie du corps déterminée :)
 bronchoscopie/G :nf/T :act/C :exam/N0 :médecin/N1 :à hum/W :faire
 coloscopie/G :nf/T :act/C :exam/N0 :médecin/N1 :à hum/W :faire
 prise de sang/G :nf/T :act/C :exam/N0 :hum/N1 :à hum/W :faire
 analyse d'urines/G :nf/T :act/C :exam/N0 :hum/N1 :à hum/W :faire

Soins

traiter/G :v/C :soin/N0 :médecin/N1 :hum
 suivre/G :v/C :soin/N0 :médecin/N1 :hum
 envoyer/G :v/C :soin/N0 :médecin/N1 :hum/N2 :chez médecin (spécialiste)

Prescriptions

prescrire/G :v/C :prescript/N0 :médecin/N1 :médicament/N2 :à hum
 donner/G :v/C :prescript/N0 :médecin/N1 :médicament/N2 :à hum
 interdire/G :v/C :prescript/N0 :médecin/N1 :aliment, activité/N2 :à hum

(Suivre un traitement)

suivre/G :v/C :/N0 :hum/N1 :traitement
 prendre/G :v/C :/N0 :hum/N1 :médicament

Opérations

opérer/G :v/C :opér/N0 :chir/N1 :hum/N2 :de Npathol
 greffer/G :v/C :opér/N0 :chir/N1 :Npc/N2 :à hum
 opération//G :nf/T :act/C :opér/N0 :chir/N1 :sur hum/W :pratiquer
 intervention chirurgicale/G :nf/T :act/C :opér/N0 :chir/N1 :sur hum/W :pratiquer
 ligature/G :nf/T :act/C :opér/N0 :chir/N1 :de Npc/N2 :à hum/W :faire
 greffe/G :nf/T :act/C :opér/N0 :chir/N1 :de Npc/N2 :à hum/W :faire
 (Avec partie du corps déterminée :)
 césarienne/G :nf/T :act/C :opér/N0 :chir/N1 :à hum/W :faire
 trachéotomie/G :nf/T :act/C :opér/N0 :chir/N1 :à hum/W :faire

Ablations

enlever/G :v/C :ablat/N0 :chir/N1 :Npc/N2 :à hum
 extraire/G :v/C :ablat/N0 :chir/N1 :Npc (dent)/N2 :à hum
 amputer/G :v/C :ablat/N0 :chir/N1 :hum/N2 :de Npc
 amputation/G :nf/T :act/C :ablat/N0 :chir/N1 :de Npc/N2 :sur hum/W :pratiquer
 ablation/G :nf/T :act/C :ablat/N0 :chir/N1 :de Npc/N2 :sur hum/W :pratiquer
 (Avec partie du corps déterminée :)
 adénectomie/G :nf/T :act/C :ablat/N0 :chir/N1 :sur hum/W :pratiquer
 hystérectomie/G :nf/T :act/C :ablat/N0 :chir/N1 :sur hum/W :pratiquer

Anesthésies

anesthésier/G:v/C:anesth/N0:médecin/N1:hum

anesthésier/G:v/C:anesth/N0:médecin/N1:Npc

anesthésie/G:nf/T:act/C:anesth/N0:médecin/N1:à hum/W:faire

anesthésie générale/G:nf/T:act/C:anesth/N0:médecin/N1:à hum/W:faire

anesthésie péridurale/G:nf/T:act/C:anesth/N0:médecin/N1:à hum/W:faire

Prédicats d'événements

contracter/G:v/C:évé/N0:hum/N1:maladie infect

se rétablir/G:v/C:évé/N0:hum

se remettre/G:v/C:évé/N0:hum/N1:de maladie, d'opération

se déclarer/G:v/C:évé/N0:maladie

se répandre/G:v/C:évé/N0:maladie infect/N1:Nloc

se réveiller/G:v/C:évé/N0:douleur

Maladies infectieuses

rougeole/G:nf/T:évé/C:maladie infect/N0:hum/W:avoir

grippe/G:nf/T:évé/C:maladie infect/N0:hum/W:avoir

Inflammations

bronchite/G:nf/T:évé/C:inflammation/N0:hum/W:avoir

dermatite/G:nf/T:évé/C:inflammation/N0:hum/W:avoir

encéphalite/G:nf/T:évé/C:inflammation/N0:hum/W:avoir

« Attaques »

accident cardiaque/G:nf/T:évé/C:attaque/N0:hum/W:avoir

rupture d'anévrisme/G:nf/T:évé/C:attaque/N0:hum/W:avoir

embolie pulmonaire/G:nf/T:évé/C:attaque/N0:hum/W:avoir

attaque d'apoplexie/G:nf/T:évé/C:attaque/N0:hum/W:avoir

crise d'épilepsie/G:nf/T:évé/C:attaque/N0:hum/W:avoir

Divers

blessure/G:nf/T:évé/N0:hum/N1:à Npc/W:avoir

courbature/G:nf/T:évé/N0:hum/N1:à Npc/W:avoir

crampe/G:nf/T:évé/N0:hum/N1:à Npc/W:avoir

fracture/G:nf/T:évé/N0:hum/N1:à Npc/W:avoir

Adjectifs causatifs

analgésique/G:adj/C:caus

anti-inflammatoire/G:adj/C:caus

antiseptique/G:adj/C:caus

diurétique/G:adj/C:caus

Prédicats d'état

États pathologiques

anémie/G:nf/T:état/C:état pathol/N0:hum/W:souffrir de

anorexie/G:nf/T:état/C:état pathol/N0:hum/W:souffrir de

dépression/G:nf/T:état/C:état pathol/N0:hum/W:avoir

scoliose/G:nf/T:état/C:état pathol/N0:hum/W:avoir

hypertension/G:nf/T:état/C:état pathol/N0:hum/W:avoir

Adjectifs d'état

anémique/G:adj/C:état/N0:hum/W:être

anorexique/G:adj/C:état/N0:hum/W:être

dépressif/G:adj/C:état/N0:hum/W:être

malade/G:adj/C:état/N0:hum/N1:de Npc/W:être

Humains prédicatifs**Fonctions**

chef de service/G :nm/T :hum/C :fonc

interne/G :nmf/T :hum/C :fonc

Professions

aide-soignant/G :nmf/T :hum/C :prof

chirurgien/G :nm/T :hum/C :prof

Malades

diabétique/G :nmf/T :hum/C :malade

anorexique/G :nmf/T :hum/C :malade

tuberculeux/G :nmf/T :hum/C :malade

Conclusion

On aura remarqué que nous avons adopté, pour décrire une langue spécialisée, les mêmes outils que ceux qui servent à la langue générale. Ce faisant, nous ne séparons pas comme relevant de niveaux différents le lexique, la syntaxe et la sémantique. Ces trois niveaux sont confondus dans ce que nous avons appelé un « emploi ». Nous avons fait remarquer que l'on ne peut pas changer l'un sans porter des modifications dans les deux autres niveaux.

Nous avons voulu, d'autre part, montrer que la description des unités lexicales doit être suffisamment précise pour qu'un système informatique puisse reconnaître ou générer des phrases dans une langue technique. C'est seulement à cette condition que l'on peut espérer mettre au point des logiciels capables de traiter un texte spécialisé de façon automatique. Les classes sémantiques qui précèdent ne sont qu'un premier essai de description que nous donnons à titre méthodologique. Leur nombre pour un domaine donné est une question empirique : il en faut autant qu'il est nécessaire pour rendre compte de la compatibilité des prédicats et des arguments. Les classes une fois établies, il faudra faire le recensement exhaustif des éléments lexicaux qui les composent. On sera alors en mesure de rendre compte du fonctionnement d'une langue de spécialité comme celle de la médecine.

NOTES

1. *hum*=humain, *ani*=animal, *vég*=végétal, *inc*=inanimé concret, *loc*=locatif, *tps*=temps. La spécificité du vocabulaire médical nous conduit à attribuer aux micro-organismes un statut particulier (T: *micro*).

RÉFÉRENCES

- GROSS, G. (1994) : « Classes d'objets et description des verbes », *Langages*, 115, p. 15-30.
 LE PESANT, D. et M. MATHIEU-COLAS, dir. (1998) : *Langages*, 131, numéro spécial « Les classes d'objets ».
 LERAT, P. (1995) : *Les langues spécialisées*, Paris, Presses universitaires de France.