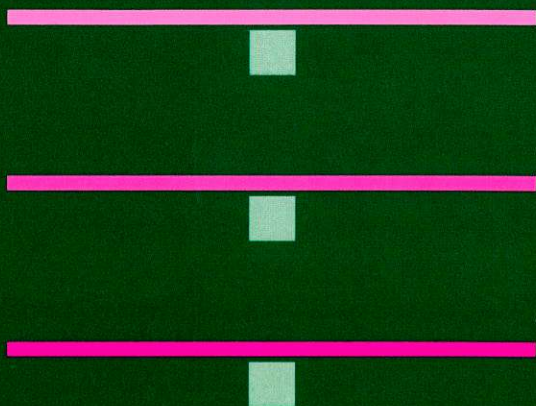


UNIVERSITÉS FRANCOPHONES



# INTRODUCTION À LA LEXICOLOGIE EXPLICATIVE ET COMBINATOIRE

Igor A. Mel'čuk / André Clas / Alain Polguère



 Editions  
Duculot



**INTRODUCTION  
À LA LEXICOLOGIE  
EXPLICATIVE  
ET COMBINATOIRE**

# Champs Linguistiques

Collection dirigée par

Dominique Willems (Université de l'État à Gand) et Marc Wilmet (Université libre de Bruxelles)

## *Recherches*

Bal W., Germain J., Klein J., Swiggers P., *Bibliographie sélective de linguistique romane et française.*

Brès J., *La narrativité.*

Cervoni J., *La préposition. Étude sémantique et pragmatique.*

Hadermann P., *Étude morphosyntaxique du mot Où.*

Jonasson K., *Le nom propre.*

Kleiber G., *Anaphores et pronoms.*

Léard J.-M., *Les gallicismes.*

Mélis L., *La voie pronominale. La systématique des tours pronominaux en français moderne.*

## *Manuels*

Klinkenberg J.-M., *Des langues romanes.*

Igor A. Mel'čuk, André Clas, Alain Polguère, *Introduction à la lexicologie explicative et combinatoire.*

## *Recueils*

Francard M., Latin D., *Le régionalisme lexical.*

UNIVERSITÉS FRANCOPHONES



# INTRODUCTION À LA LEXICOLOGIE EXPLICATIVE ET COMBINATOIRE

Igor A. Mel'čuk / André Clas / Alain Polguère



 Editions  
Duculot

La collection *Universités francophones*, créée en 1988 à l'initiative de l'UREF, propose des ouvrages modernes répondant aux besoins des étudiants de deuxième et troisième cycle universitaire ainsi qu'aux chercheurs francophones, et se compose de titres originaux paraissant régulièrement.

Leurs auteurs appartiennent conjointement aux pays du Sud et du Nord et rendent compte des résultats de recherches et des études entreprises en français à travers le monde. Ils permettent à cette collection pluridisciplinaire de couvrir progressivement l'ensemble des enseignements universitaires en français.

Enfin, la vente des ouvrages à un prix préférentiel destinés aux pays du Sud tient compte des exigences économiques nationales et assure une diffusion adaptée aux pays francophones.

Ainsi, la collection *Universités francophones* constitue une bibliothèque de référence comprenant des ouvrages universitaires répondant aux besoins des étudiants et des chercheurs en langue française.

Autres ouvrages de linguistique parus dans la collection *Universités francophones*

- *Inventaire des particularités lexicales du français en Afrique noire* (Equipe IFA), 1988.
- *Visages du français, variétés lexicales de l'espace francophone* (André Clas, Benoît Ouoba), 1990.
- *L'environnement traductionnel. La station de travail du traducteur de l'an 2001* (André Clas et Hayssam Safar), 1992.
- *Inventaire des usages de la francophonie: nomenclatures et méthodologies* (Danièle Latin, Ambroise Queffelec, Jean Tabi-Manga), 1993.
- *La traductique* (Pierrette Bouillon et André Clas), 1993.
- *Contribution à un inventaire des particularités lexicales du français de l'île Maurice* (Didier de Robillard), 1993.

© Duculot s.a. 1995  
Fond Jean-Pâques, 4 B - 1348 Louvain-la-Neuve

Toute reproduction d'un extrait quelconque de ce livre, par quelque procédé que ce soit, et notamment par photocopie ou microfilm, est strictement interdite.

*Printed in Belgium*

# PRÉFACE

La diffusion de l'information scientifique et technique est un facteur essentiel du développement. Aussi dès 1988, l'Agence francophone pour l'enseignement supérieur et la recherche (AUPELF-UREF), mandatée par les Sommets francophones pour produire et diffuser revues et livres scientifiques, a créé la collection « Universités francophones ».

Lieu d'expression de la communauté scientifique de langue française, *Universités francophones* vise à instaurer une collaboration entre enseignants et chercheurs francophones en publiant les ouvrages, coédités avec des éditeurs francophones, et largement diffusés dans les pays du Sud grâce à une politique tarifaire préférentielle.

Quatre séries composent la collection :

- Les manuels : cette série didactique est le cœur de la collection. Elle s'adresse à un public de deuxième et troisième cycles universitaires et vise à constituer une bibliothèque de référence couvrant les principales disciplines enseignées à l'université.
- Sciences en marche : cette série se compose de monographies qui font la synthèse des travaux de recherche en cours.
- Actualité scientifique : dans cette série sont publiés les actes de colloques organisés par les réseaux thématiques de recherche de l'UREF.
- Perspectives francophones : s'inscrivent dans cette série des ouvrages de réflexion donnant l'éclairage de la Francophonie sur les grandes questions contemporaines.

Notre collection, en proposant une approche plurielle et singulière de la science, adaptée aux réalités multiples de la Francophonie, contribue efficacement à promouvoir la recherche dans l'espace francophone et le plurilinguisme dans la recherche internationale.

Professeur Michel GUILLOU  
Directeur général de l'AUPELF  
Recteur de l'UREF





## SYMBOLES ET NOTATIONS

(Cette liste ne tient pas compte des symboles et notations spécifiques à l'informatisation du DEC, introduits au Chapitre VI.)

A	:	actant
ASém	:	actant sémantique
ASyntP	:	actant syntaxique profond
ASyntS	:	actant syntaxique de surface
BDTI	:	base de données textuelles informatisée
C	:	colonne (du tableau de régime)
CO	:	complément d'objet
DEC	:	Dictionnaire explicatif et combinatoire
DiCo	:	Dictionnaire de cooccurrences
ÉDiCo	:	Éditeur du DiCo [programme d'ordinateur]
<b>fam</b>	:	familier
FL	:	fonction lexicale
ILEC	:	<i>Introduction à la lexicologie explicative et combinatoire</i>
<i>L</i>	:	langue dont il est question
L	:	lexie
N	:	nom
PR	:	<i>Petit Robert</i>
PROP	:	proposition
RSém	:	représentation sémantique
SG	:	sujet grammatical
Sém	:	sémantique
<b>spéc</b>	:	langue spécialisée
SyntP	:	syntaxique profond
TALN	:	traitement automatique du langage naturel
V	:	verbe
X   Y	:	Y sont les conditions de l'application ou de l'usage de X
⟨X⟩	:	sens de X
⟨X    Y⟩	:	X est le présupposé, et Y est le posé [dans une définition]

// Y	:	Y est un élément fusionné de la valeur de la FL en question
X ⟨Y⟩	:	Y est une variante de X
Λ	:	l'ensemble vide
⊃, ⊂	:	inclusion ensembliste (X ⊃ Y signifie 'l'ensemble X inclut l'ensemble Y – comme son sous-ensemble')
∩	:	intersection ensembliste (X ∩ Y signifie 'l'ensemble X et l'ensemble Y ont des éléments en commun')
~	:	représente la lexie vedette dans une description lexicographique
◇	:	indicateur de la zone phraséologique où les phrasèmes sont sémantiquement apparentés à la lexie vedette
◆	:	indicateur de la zone phraséologique où les phrasèmes ne sont pas sémantiquement apparentés à la lexie vedette
*X	:	expression jugée sémantiquement ou syntaxiquement inacceptable
?X	:	expression jugée sémantiquement ou syntaxiquement difficilement acceptable
[L <sub>1</sub> + L <sub>2</sub> + ... + L <sub>n</sub> ]	:	expression phraséologique constituée de lexies L <sub>1</sub> , L <sub>2</sub> , ..., L <sub>n</sub> et qui est une vedette

# TABLE DES MATIÈRES

SYMBOLES ET NOTATIONS	VII
AVANT-PROPOS	5
PRÉFACE	9
<b>Chapitre I – Théorie lexicologique</b>	15
1 Lexie, unité de base de la lexicologie	15
2 Dictionnaire, produit final de la lexicologie	18
2.1 Concepts de base	18
2.2 La confection des dictionnaires	21
2.2.1 Difficultés linguistiques de la confection de dictionnaires	22
2.2.2 Difficultés psychologiques de la confection de dictionnaires	24
3 Lexicologie, discipline carrefour	26
3.1 Lexicologie vs lexicographie	26
3.2 La spécificité de la lexicologie	28
3.3 Le dictionnaire idéalisé : le DEC	29
<b>Chapitre II – Caractérisation générale du Dictionnaire explicatif et combinatoire [= DEC]</b>	33
1 Principes de rédaction du DEC	34
1.1 Principe de formalité	34
1.2 Principe de cohérence interne	36
1.3 Principe de traitement uniforme	40
1.4 Principe d'exhaustivité	42
2 Nomenclature du DEC	43
2.1 Principe de disponibilité	43
2.2 Principe de lexicalité	45
3 Documentation du DEC	48
3.1 Principe de la primauté du locuteur	49
3.2 Principe de consultation obligatoire des bases de données textuelles	50

<b>Chapitre III – Microstructure du Dictionnaire explicatif et combinatoire</b>	<b>55</b>
1 Concept de lexie	55
1.1 Éléментарité des articles de dictionnaire du DEC	55
1.2 Lexie, unité de base du DEC	56
2 Délimitation d'une lexie	57
2.1 Généralités	57
2.2 Critères de délimitation des lexies	59
3 Description d'une lexie = structure d'un article de dictionnaire	69
3.1 Généralités	69
3.2 Zone phonologique	71
3.3 Zone sémantique	72
3.3.1 Définition lexicographique	72
3.3.1.1 Présentation de la définition lexicographique	73
3.3.1.2 Concept fondamental de la définition lexicographique : actant sémantique	75
3.3.1.3 Règles d'élaboration de la définition lexicographique	78
3.3.1.4 Critères de formulation de la définition lexicographique	95
3.3.1.5 Structure et contenu de la définition lexicographique	102
3.3.1.5.1 Hiérarchisation des composantes sémantiques dans la définition lexicographique	103
3.3.1.5.2 Ordonnancement des composantes sémantiques dans la définition lexicographique	106
3.3.1.6 Une liste de vérification lexicographique [= une check-list]	107
3.3.2 Connotation lexicographique	111
3.4 Zone de combinatoire syntaxique	117
3.5 Zone de combinatoire lexicale	125
3.5.1 Présentation des fonctions lexicales	125
3.5.2 Inventaire des fonctions lexicales standard simples	129
3.5.3 Les fonctions lexicales « irrégulières »	148
3.5.3.1 Fonctions lexicales complexes	148
3.5.3.2 Configurations de fonctions lexicales	149
3.5.3.3 Fonctions lexicales non standard	150
3.5.4 Éléments fusionnés des valeurs des fonctions lexicales	151

3.6 Zone d'exemples	152
3.7 Zone phraséologique	153
<b>Chapitre IV - Macrostructure du Dictionnaire explicatif et combinatoire</b>	<b>155</b>
1 Structure d'un superarticle de dictionnaire : le vocable	155
1.1 Concepts pertinents	157
1.2 Regroupement des lexies en vocables	160
1.3 Ordonnancement des lexies d'un même vocable	162
1.4 Distances sémantiques entre les lexies d'un même vocable	166
1.5 Tableau synoptique du vocable	171
2 Structure d'une famille d'articles de dictionnaire : champ lexical	172
2.1 Champ sémantique	173
2.2 Champ lexical	176
2.3 Principe de traitement par champs sémantiques et lexicaux	178
2.4 Champs sémantiques et lexicaux dans le dictionnaire	178
<b>Chapitre V – Illustrations lexicographiques</b>	<b>181</b>
1 Détermination d'un champ lexical	181
2 Détermination d'un vocable	184
2.1 Division des acceptions	184
2.2 Regroupement des acceptions	186
2.3 Organisation du regroupement obtenu	186
2.3.1 Ordonnancement des articles de dictionnaire au sein d'un superarticle	187
2.3.2 Indication des distances sémantiques entre les lexies	187
3 Description d'un lexème : VENDREI.1	188
3.1 Zone sémantique	188
3.1.1 Définition lexicographique	188
3.1.2 Connotations	190
3.2 Zone syntaxique	190
3.3 Zone de combinatoire lexicale	191
3.4 Zone d'exemples	194
3.5 Zone phraséologique	194
4 Articles de dictionnaire illustratifs : VENDRE, SE VENDRE, VENDEUR	195
<b>Chapitre VI - Informatisation du Dictionnaire explicatif et combinatoire</b>	<b>205</b>
1 Qu'est-ce qu'un DEC informatisé ?	205

2	Description détaillée du Dictionnaire de cooccurrences [= DiCo]	210
2.1	Formalismes du DiCo	211
2.1.1	Structure générale d'un fichier de DiCo descrivant un vocable = un superarticle de DiCo	211
2.1.2	Analyse détaillée des composantes d'un superarticle de DiCo	214
2.2	Exemples de deux superarticles de DiCo	223
2.2.1	Traduction en format DiCo d'un superarticle de DEC : VENDRE	223
2.2.2	Un superarticle original de DiCo : ADHÉSION	226
3	Les éditeurs lexicologiques spécialisés	227
3.1	Qu'est-ce qu'un éditeur spécialisé de dictionnaire ?	227
3.2	Possibilités offertes par l'éditeur du DiCo	228
3.2.1	Opérations d'édition proprement dites	229
3.2.2	Opérations de recherche sur la base de données dictionnairiques	232
3.2.3	Opérations de vérification automatique de la cohérence intra- et interarticles	233
3.3	Possibilités additionnelles qui seraient offertes par un éditeur de DEC véritable	235
4	DEC et DiCo en tant que bases de données lexicales pour les systèmes de traitement automatique du langage naturel	236
	RÉFÉRENCES	241
	INDEX	247
	INDEX DES DÉFINITIONS	247
	INDEX DES TERMES ET DES CONCEPTS	248
	INDEX DES VOCABLES TRAITÉS	254

## AVANT-PROPOS

*L'Introduction à la lexicologie explicative et combinatoire* [= ILEC] est destinée à tous ceux qui s'intéressent, comme l'indique le titre, à la lexicologie – mais aussi à la lexicographie puisque, dans cet ouvrage, nous considérons que la lexicographie constitue une plate-forme d'essai pour les découvertes lexicologiques. Ce texte s'adresse également à ceux qui se passionnent pour la métalexicologie et la métalexicographie, et même à ceux qui, d'une façon ou d'une autre, veulent mieux connaître la linguistique générale et appliquée, dans ses rapports au mot. En fait, l'ILEC vise tous ceux qui ont un « intérêt » dans le lexique et dans les dictionnaires. Il répond à un grand nombre de questions lexicologiques ou lexicographiques, et il examine certains problèmes posés par la linguistique théorique et descriptive. Nous y abordons des interrogations théoriques, mais nous y présentons aussi de très nombreuses solutions lexicologiques pratiques. Dans l'ILEC, la lexicologie s'imprègne de lexicographie, et inversement. Le dictionnaire – en tant qu'objet d'études et résultat final de la recherche – prend donc une allure toute nouvelle dans cette méthodologie. En effet, nous proposons des principes, des critères et des règles strictes – des cheminements à respecter qui garantissent une qualité de réflexion et des résultats descriptifs qu'il n'est plus possible de mettre de côté.

Qu'est-ce que l'ILEC sinon une série de préceptes qui s'appuient sur des exemples et sont démontrés par eux ? En effet, de nombreuses illustrations sont présentées et finement décortiquées pour décrire des faits lexicaux de langue de façon systématique et précise. Ainsi, toute solution et toute procédure proposées visent à devenir un patron, un moule, peut-être même un modèle. Nous montrons comment sont les faits linguistiques, comment les interpréter et comment les exploiter, et surtout nous justifions la validité de la description lexicologique qu'on peut en tirer.

Bien sûr, nous n'avons pas tout inventé, loin de là. Nous avons largement suivi l'outillage connu des approches lexicologiques / lexicographiques anciennes et modernes de nos prédécesseurs; notre originalité se résume, avant tout, à avoir systématisé les descriptions et les présentations et surtout à avoir minutieusement respecté, parfois de façon quelque peu rigide, les principes de logique. Cela est en quelque sorte notre doctrine et nos lignes directrices. Nous avons résolument bâti notre approche sur la théorie linguistique Sens-Texte, qui sert de charpente et de fil directeur et nous fait aboutir à un modèle de dictionnaire, qui est un dictionnaire théorique – non pas dans le sens qu'il n'a pas d'existence concrète, mais dans l'acception qu'il répond à une certaine vision théorique de la réalité linguistique, que son pouvoir descriptif est maximalisé, et donc que la justesse de

la présentation des faits de langue rend justice à la théorie sous-jacente. Un tel dictionnaire correspond, nous en sommes convaincus, à un dictionnaire idéalisé.

Bien évidemment, en procédant comme nous l'avons fait, nous avons dû privilégier le recours à l'intuition des locuteurs ainsi qu'aux bases de données textuelles informatisées plutôt qu'aux dictionnaires existants; nous avons ainsi brisé la pratique, toujours dominante en lexicographie, de puiser les données lexicales dans des dictionnaires précédents. Mais attention ! L'intuition « moyenne » à laquelle nous faisons appel n'est qu'une abstraction; elle ne concorde pas nécessairement avec l'intuition particulière, celle du spécialiste du domaine ou du linguiste normatif, par exemple. Le lecteur peut donc se trouver en désaccord avec telle ou telle évaluation des faits : le français est une langue bien vivante, et par conséquent des flottements, des hésitations et des désaccords entre locuteurs sont à prévoir. Cela n'est point dangereux pour notre exposé : notre prétention est de donner une description formelle et précise des faits tels qu'ils sont établis par nous; la correspondance entre les données linguistiques citées et notre description de ces données est plus importante dans le cadre de l'ILEC que l'exactitude des données elles-mêmes. Il est clair, cependant, que nous récusons les manipulations des données qui cherchent à justifier certaines décisions théoriques *a priori* : c'est la théorie lexicologique qui doit refléter les faits de langue, et non l'inverse !

Ajoutons à cela que la connaissance explicative des faits linguistiques ne peut avoir de vraie validité que dans un cadre théorique bien précis, où le crible analyseur reste stable et logiquement explicite. Un tel cadre théorique sous-tend tous les concepts descriptifs utilisés dans cet ouvrage. Rappelons, si besoin est, qu'il n'y a pas de fait linguistique établi sans théorie et que seule la théorie permet d'interpréter le fait constaté et ainsi de faire progresser la connaissance. Nous espérons que le théoricien trouvera dans notre ouvrage ample matière à réflexion, à analyse, à expérimentation, à vérification et à prise de position, et que le praticien saura suivre méthodiquement les préceptes exposés, amasser les exemples nécessaires, et produire une œuvre où la description des données linguistiques relevées est systématique et complète pour le plus grand bien des usagers. L'ILEC est très précisément une présentation des concepts clés de la lexicologie/lexicographie. Est-ce à dire que tout a été abordé de façon complète et immuable ? Bien évidemment pas. À notre avis, on trouvera dans ce manuel l'équipement MINIMAL mais indispensable et suffisant pour pouvoir cheminer allègrement. Y a-t-il plus ? Bien sûr, certains concepts mériteraient d'être plus détaillés, plus décortiqués, plus explicités par des retours à des questions de linguistique générale ou plus longuement étudiés dans un traité particulier, ou tout simplement étayés par des indications bibliographiques précises. Nous avons, répétons-le, emprunté à de nombreux devanciers, mais nous n'avons pas systématiquement marqué ce que nous leur devons, et nous nous en excusons. Il eût d'ailleurs été délicat de le faire sans de longues et difficiles vérifications, dont les résultats n'auraient de toute façon satisfait personne. Notre objectif est de produire une



synthèse qui incorpore les « acquis » et les « nouveautés » avec les mêmes exemplifications standardisées, en essayant de donner à l'ensemble la cohérence logique la plus grande possible – même au risque de donner à l'ILEC une allure un peu impersonnelle. Il n'était peut-être pas souhaitable, pour ne pas dérouter le lecteur, de présenter tous les points de vue théoriques : l'unicité de la démarche et de la méthodologie en eussent été les victimes.

Il nous faut encore ajouter que, la terminologie en usage étant parfois flottante ou même divergente, nous avons pris soin de donner nos interprétations et nos définitions aux termes utilisés. Il était nécessaire de procéder ainsi pour clarifier certains concepts, mais surtout pour éviter d'égarer le lecteur sur de fausses pistes.

Si, selon le dicton, une image vaut mille mots et si, en le paraphrasant, on peut dire qu'un exemple vaut un paragraphe d'explications, alors nous avons bien suivi cette recommandation. Nous avons en effet, dans tous les cas, fourni de très nombreuses illustrations de tout genre. Parfois même des exemples développés remplacent des explications qui eussent été indispensables : nous espérons que l'excès d'exemples peut suppléer à la disette d'explications.

La recherche linguistique sur laquelle se base l'ILEC a été en partie financée par deux projets de recherche du professeur I. Mešćuk (subvention n° 96-ER-0618 des Fonds FCAR et subvention n° 410-91-1844 du CRSH). L'aide à la rédaction et à la publication du manuscrit nous est venue de l'Agence Franco-phonie pour l'Enseignement Supérieur et la Recherche.

Nous ne pouvons pas terminer ce bref avant-propos sans exprimer explicitement notre reconnaissance à divers chercheurs qui ont lu avec compétence et patience le manuscrit de l'ILEC. Leurs interrogations, leurs critiques, leurs réflexions et leurs propositions nous ont permis, dans certains cas, de pousser plus loin notre recherche et, dans d'autres cas, de corriger, de redresser ou d'amplifier des présentations. Nous aimerions plus particulièrement remercier Henri Béjoint, Daniel Blampain, Jean-Luc Descamps, Lidija Iordanskaja et Suzanne Mantha pour leur aide précieuse. Nous sommes également reconnaissants à Danielle Collignon, Lidija Iordanskaja et Suzanne Mantha pour leur participation dans la mise au point du manuscrit et des épreuves.

Les auteurs



## PRÉFACE

La plupart des lecteurs évitent sciemment les préfaces. Ils considèrent en effet qu'une préface ne peut avoir trop d'importance puisque l'ouvrage lui-même doit apporter tous les renseignements indispensables à une bonne compréhension. Les préfaces, on les écrit par politesse ou par tradition ! Pourtant, une préface peut renfermer des éléments absolument nécessaires pour assurer une lecture et une interprétation fructueuses du livre en entier. Nous prétendons que c'est justement le cas ici : cette préface a pour but de préparer le lecteur à percevoir notre texte DANS LA PERSPECTIVE QUE NOUS AVONS VOULU DONNER À L'OUVRAGE.

*L'Introduction à la lexicologie explicative et combinatoire*, ou ILEC, peut à maintes occasions surprendre ou même choquer par son rigorisme extrême et sa sécheresse formaliste. En effet, ni en lexicologie ni en lexicographie, on n'a l'habitude de rencontrer des constructions quasi mathématiques, avec des systèmes de définitions logiques et tout leur appareillage de déductions. Même si plusieurs disciplines de la linguistique moderne, surtout la syntaxe et la sémantique, ont déjà franchi le pas depuis des années, il n'en est pas ainsi en lexicologie/lexicographie. Nous croyons que l'ILEC est un des premiers ouvrages qui tentent d'implanter une approche logique et formelle dans l'étude des mots. Nous nous sommes donné une tâche bien spécifique :

|| proposer des méthodes de description rigoureuse, formelle et exhaustive du lexique.

Cette orientation impose inévitablement des cheminements rigides, des formulations catégoriques pouvant même agacer ou des analyses excessivement tranchées des matériaux linguistiques. Dans le domaine du lexique, il est habituel de procéder avec précaution, de proposer des descriptions approximatives et nuancées, en les situant de façon continue et en laissant un rôle vital au contexte et à l'intuition des locuteurs. Tout en reconnaissant le caractère naturel de cette façon de faire, nous avons opté pour une autre approche, celle de la prise de décisions discrètes (dans le sens mathématique) et absolues. Cette tendance s'intègre d'ailleurs parfaitement dans l'ensemble des tendances engendrées par la révolution informatique, ce qui justifie notre tâche et confirme sa justesse.

En suivant cette voie, nous ne voulons en aucun cas scandaliser ni d'ailleurs prétendre à la vérité immuable de nos constructions. Nous sommes, d'une certaine façon, prisonniers de notre option : formalisme oblige !

Nous espérons que le lecteur saura accepter, avec compréhension et indulgence, la brusquerie de nos propositions et de nos démonstrations, en nous

accompagnant pas à pas dans les pistes lexicales tracées. Pour l'aider, nous pensons utile d'éclaircir ici même, ne serait-ce que brièvement, les trois points suivants :

- les limites de notre étude;
- les sources de nos idées;
- la structure de l'ouvrage.

### Les limites de notre étude

L'*Introduction à la lexicologie explicative et combinatoire* peut étonner par au moins deux particularités. D'une part, par le titre même : que faut-il entendre par « explicative et combinatoire » ? Il s'agit en fait d'une allusion au *Dictionnaire explicatif et combinatoire (du français contemporain)* [= DEC], sur lequel nous travaillons depuis environ quinze ans (Mel'čuk *et al.* 1984, 1988, 1992). C'est un dictionnaire particulier dont il sera constamment question dans les pages qui suivent. Disons ici que *explicatif* dans son titre signifie que tout élément lexical qui y est consigné est infailliblement accompagné d'une EXPLICATION SÉMANTIQUE formelle, de sorte que les représentations sémantiques sont sous-jacentes au reste de la description lexicographique dont est munie chaque unité lexicale. L'adjectif *combinatoire*, quant à lui, indique que le DEC accorde une valeur particulière à la présentation rigoureuse et exhaustive de la COMBINATOIRE LEXICALE. Ainsi le DEC prend en charge, de façon systématique, l'axe paradigmatique dans le lexique (oppositions sémantiques, sélection sémantique d'unités) et l'axe syntagmatique (enchaînement des unités lexicales dans le texte). L'ILEC est en fait une théorisation du DEC; nous avons essayé de généraliser et de systématiser notre expérience lexicologique/lexicographique et, en même temps, d'établir des objectifs idéalisés à atteindre. En un mot, nous voulons proposer un modèle dictionnaire à discuter, à développer et à imiter.

D'autre part, nous venons d'avouer que l'ILEC traite, de façon essentielle, d'un dictionnaire; comment alors justifier *lexicologie* dans le titre si l'attitude générale consiste à associer le dictionnaire à la *lexicographie*, qu'on oppose toujours à la lexicologie ? Nous reprenons cette opposition dans le corps du texte (Ch. I, 3.1, p. 26); disons, pour le moment, que nous sommes convaincus qu'une étude lexicologique ne peut mener qu'à un dictionnaire, que la lexicographie n'est qu'une application de la lexicologie et que les deux disciplines ont la même importance pour la confection de dictionnaires.

Somme toute, l'ILEC est un ouvrage théorique sur le lexique, et donc le titre est justifié. Mais puisque toute étude lexicale d'envergure aboutit à un (fragment d'un) dictionnaire – un dictionnaire théorique, mais un dictionnaire quand même – il est logique que l'ILEC s'appuie sur un dictionnaire et le propose comme modèle idéalisé.

## Les sources de nos idées

Nous avons déjà prévenu le lecteur que nous n'avons pas systématiquement mis en relief tout ce que nous devons à nos prédécesseurs. Nous croyons quand même qu'il est important de rappeler que notre approche s'inspire fondamentalement des écrits, des idées et des théories que nous pouvons regrouper sous les quatre rubriques suivantes :

1. Les travaux de pointe de lexicologie et de lexicographie françaises. Nous pensons plus particulièrement au *Petit Robert* et aux ouvrages devenus classiques : Wagner 1967-70, Quemada 1968, Dubois & Dubois 1971, Rey 1970, 1973-76, 1977, 1982, Rey-Debove 1971, 1978, Pottier 1973. Il nous faut ajouter encore Collignon & Glatigny 1978, Picoche 1986 et Descamps *et al.* 1976.
2. Les travaux et les manuels généraux de lexicologie / lexicographie, tels que Hallig & von Wartburg 1952, Zgusta 1971, Hartmann 1983, Landau 1984, Tournier 1985, 1988, Lipka 1990, Béjoint 1994. Nous tenons particulièrement à mentionner la contribution de F. J. Hausmann, par exemple, Hausmann & Wiegand 1989 et Hausmann *et al.* 1989. Rappelons de plus le rôle important des revues spécialisées, telles que *Cahiers de lexicologie*, *Dictionaries*, *International Journal of Lexicography* et *Lexicographica*.
3. Les importants travaux sémantiques de A. Zholkovsky, de A. Bogusławski et surtout de A. Wierzbicka (Zolkovskij 1964a-b, Bogusławski 1970, Wierzbicka 1972, pour ne citer que les premières publications de ces auteurs). De nombreuses idées venant de leurs ouvrages ont été incorporées, sous une forme ou sous une autre, dans le DEC du russe (Mel'čuk & Zholkovsky 1984) ainsi que dans le DEC du français et continuent aujourd'hui encore à soutenir nos recherches lexicologiques.
4. Les recherches fondamentales en sémantique lexicale et la lexicologie de Jurij Apresjan. Il a exercé une influence toute particulière, qu'il est difficile de cerner, tant elle est présente et essentielle : de l'article Apresjan 1969, où l'auteur a jeté les fondements d'une lexicographie formelle (par l'énonciation de principes rigoureux de rédaction de dictionnaires), à la série de ses travaux lexicologiques (1979, 1986, 1988a-b, 1990a-b, 1992 et 1995), en passant par son livre classique (1974).

Dans la même veine, nous aimerions reconnaître notre dette vis-à-vis de Lidija Iordanskaja, dont nous avons exploité plusieurs idées tirées des trois volumes publiés du DEC français.

## La structure de l'ouvrage

Six chapitres forment la structure de l'ILEC.

Le Chapitre I caractérise la discipline dont il est question, la lexicologie. C'est clairement une DÉTERMINATION DU DOMAINE D'ACTIVITÉ et la pose de jalons conceptuels. En effet, il faut, pour cheminer normalement, centrer la réflexion sur la lexie, donc sur la lexicologie, et par ricochet sur le dictionnaire, produit et réalité vérificatrice finale de la théorie mise en œuvre. C'est encore le lieu pour marquer les interprétations et trancher le champ des investigations : quelle est la relation entre lexicologie et lexicographie ? De là, la réflexion et le partage des faits amènent à ce qu'est un dictionnaire, un dictionnaire quelque peu particulier, posé dans l'absolu, un dictionnaire que nous qualifions d'idéalisé. Ce dictionnaire idéalisé est incarné dans le Dictionnaire explicatif et combinatoire, ou DEC.

Le Chapitre II présente les CARACTÉRISTIQUES SPÉCIFIQUES DU DEC. Il explicite la nécessité des principes rigoureux de rédaction, en formule huit et en propose la justification. On le sait : un dictionnaire est affaire de rédaction, de nomenclature et de documentation; les principes exposés déterminent donc le cheminement de la rédaction, le choix de la nomenclature et le recours à la documentation.

Dans le Chapitre III, on étudie ce qu'il est convenu d'appeler la MICROSTRUCTURE du dictionnaire, c'est-à-dire l'organisation de ses articles, chaque article traitant d'une lexie. Là encore, se posent des questions fondamentales. Quelle est la nature d'une lexie ? Comment la déterminer et surtout comment la décrire avec efficacité ? C'est bien entendu le point central de toute entreprise lexicologique/lexicographique. C'est là qu'on trouve les points d'ancrage de la démarche proposée, ainsi que les critères et les règles appelés à orienter la recherche lexicale : ils serviront de « garde-fou » au lexicologue / lexicographe. C'est bien sûr le noyau dur de toute l'entreprise.

Le Chapitre IV aborde, dans le même esprit que le chapitre précédent, la MACROSTRUCTURE du dictionnaire, c'est-à-dire l'organisation des familles d'articles. D'autres règles sont proposées et étayées; elles servent de guide pour l'exploration des champs sémantiques et lexicaux.

Le Chapitre V se veut la zone d'ILLUSTRATIONS; nous y citons au complet quelques articles de dictionnaire du DEC, en les munissant des explications indispensables. De plus, nous présentons, de façon relativement détaillée, la procédure qu'un lexicologue / lexicographe est censé suivre dans l'élaboration de tels articles.

Le Chapitre VI explore les possibilités d'INFORMATISATION des descriptions lexicologiques / lexicographiques. Un dictionnaire explicatif et combinatoire est en effet un instrument d'une extrême utilité, tant pour la vérification des possibilités langagières que pour la génération automatique de textes ou pour le passage d'une langue à l'autre, c'est-à-dire pour les systèmes de traduction automatique ou assistée par ordinateur. Nous y ébauchons un système automatique

qui est appelé à servir comme rédacteur informatique pour l'élaboration de dictionnaires du type que nous lançons dans l'ILEC.

Notons que l'ILEC contient des notes en bas de page qui sont plus chargées de contenu que celles normalement trouvées dans les ouvrages de ce type. Dans nombre de cas, nous y avons relégué des informations pertinentes ou même indispensables mais qui auraient brisé la suite logique de l'exposé dans le cadre que nous nous étions fixé.

## Quelques remarques techniques

Dans notre présentation, nous avons souvent besoin de citer des exemples d'expressions inacceptables (= agrammaticales) ou difficilement acceptables (= semi-grammaticales). Les premiers, selon la tradition en linguistique, sont marqués par l'astérisque, qui signifie donc 'impossible', les seconds sont précédés par un point d'interrogation relevé, qui veut dire 'non souhaitable' :

- (1) a. *\*Les yeux de Jean sur la mer, il pensait à son avenir [= Ses yeux sur la mer, Jean pensait à son avenir].*  
 b. *?L'autorisation du contrôleur de décoller est requise [= Pour décoller, il faut obtenir l'autorisation du contrôleur].*

Comme on peut le voir dans ces exemples, après avoir cité une expression impossible ou non souhaitable, nous donnons au moins une « bonne » variante.

Les conventions typographiques adoptées dans l'ILEC sont les suivantes :

1. Les exemples linguistiques sont imprimés en caractères italiques.
2. Les noms de lexies sont imprimés en capitales.
3. Les premières mentions de termes importants introduits dans l'ILEC sont imprimées en caractères gras italiques; ces termes apparaissent dans l'index.
4. La mise en relief se fait par de petites capitales.
5. Les parenthèses entourant une expression sont utilisées pour marquer son caractère optionnel; les crochets renferment toutes sortes de commentaires et d'exemples informels; les chevrons signalent des variantes possibles.

Les exemples sont numérotés par des chiffres arabes entre parenthèses consécutivement dans chaque chapitre. Les principes de rédaction, les critères et les règles lexicologiques sont numérotés à travers l'ouvrage entier. Les définitions de concepts pertinents sont, au contraire, numérotées par des numéros doubles : le numéro du chapitre en chiffres romains, suivi du numéro de la définition dans ce chapitre; par exemple, **II.4** ou **IV.1**.

Le souci de désambiguïsation nous force à utiliser assez souvent les numéros lexicaux distinctifs, qui doivent identifier les acceptions voulues. Quand la lexie analysée a déjà été décrite dans les volumes publiés du DEC, nous empruntons

les numéros lexicaux à ce dernier – sans l'indiquer explicitement. Si la lexie qui nous intéresse n'a pas encore d'article dans le DEC, nous empruntons le numéro lexical au *Nouveau Petit Robert* (1993), en le signalant par l'abréviation PR.

Une dernière mise en garde : dans certains cas, nous reprenons les définitions et les autres informations lexicographiques des volumes publiés du DEC, en les changeant sans préavis – si nous pensons pouvoir présenter maintenant un meilleur traitement de l'unité lexicale concernée. Cela signifie que dans un cas de décalage entre la description proposée dans l'ILEC et celle trouvée dans le DEC publié, le lecteur est prié de prendre en compte celle de l'ILEC.



# CHAPITRE I

## THÉORIE LEXICOLOGIQUE

L'*Introduction à la lexicologie explicative et combinatoire* [dorénavant, ILEC] se veut une introduction à l'ÉTUDE DES LEXIES des langues naturelles. Dans notre perspective, cette étude doit répondre à deux impératifs majeurs :

1. Elle doit être MULTILATÉRALE. Cela signifie que la lexie est considérée simultanément

- sous sa facette sémantique,
- sous sa facette syntaxique,
- et sous sa facette lexico-combinatoire,

de façon à permettre une description globale qui prenne en charge toutes les caractéristiques dégagées.

2. Elle doit obligatoirement aboutir à l'élaboration de DICTIONNAIRES.

La conclusion qu'on doit tirer de cette prise de position est l'affirmation suivante :

|| Une étude de lexies visant l'élaboration de dictionnaires relève de la lexicologie théorique et descriptive.

Cette formulation entraîne la nécessité d'explicitier, ne serait-ce que brièvement, les trois éléments fondamentaux qui la composent, à savoir la lexie, le dictionnaire et la lexicologie. Nous en traiterons dans les trois sections qui suivent.

### 1 Lexie, unité de base de la lexicologie

La lexie est l'unité de base de la lexicologie – en fait, son objet central et même, en schématisant un peu, son seul et unique objet. Par conséquent, le concept de lexie devrait être défini dès à présent de façon rigoureuse et muni de toutes les explications indispensables. Cependant, comme ce concept est fort complexe, nous ne pouvons pas le traiter en profondeur ici et nous nous contenterons pour l'instant de quelques remarques préliminaires. Une discussion détaillée de la lexie sera proposée au début du chapitre III, 1.2, p. 56 *ssq.*

Le concept de lexie est une formalisation et, simultanément, une généralisation de la notion de MOT. Il n'est pas dans notre intention d'entreprendre ici une analyse poussée de cette notion, même si l'on sait bien que le mot est une unité centrale de la langue : on connaît tout aussi bien le caractère rétif du mot *mot*, qui, jusqu'à présent, a échappé aux tentatives de le circonscrire avec précision et a fait couler beaucoup d'encre pendant des décennies. Parmi les très nombreux ouvrages qui cherchent à expliciter la notion de mot, nous n'indiquerons que Juilland & Roceris

1972 (une bibliographie) et trois livres présentant des approches différentes : Krámský 1969, Pergnier 1986 et Di Sciullo & Williams 1987; on trouvera une étude détaillée de la question et des propositions de définitions dans Mel'čuk 1993 : Partie I, p. 97 *ssq.*

Pour le moment, il nous suffit de dire qu'une *lexie*<sup>1</sup> ou *unité lexicale*, est soit un mot pris dans une acception bien spécifique (= *lexème*), soit encore une locution, elle aussi prise dans une acception bien spécifique (= *phrasème*). Ainsi, si nous considérons le mot PONT, nous pouvons en dégager, de façon intuitive et en nous fiant à des dictionnaires, à peu près une quinzaine de lexies (plus loin, nous analyserons et formaliserons le processus d'isolation de lexies) :

Lexie 1 : PONT au sens de 'construction reliant les deux rives d'une étendue d'eau...';

Lexie 2 : PONT au sens de 'jours chômés entre deux jours fériés';

Lexie 3 : PONT au sens de 'ensemble des organes transmettant le mouvement... [dans une automobile]';

Lexie 4 : PONT au sens de 'circuit électrique formé par 4 composants...';

Lexie 5 : PONT au sens de 'plancher fermant par en haut la coque d'un bateau...';

...

De plus, on connaît toute une série de locutions comprenant le mot PONT, qui constituent d'autres lexies :

Lexie 6 : PONTS ET CHAUSSÉES au sens de 'service administratif des voies publiques...';

Lexie 7 : PONT AUX ÂNES au sens de 'banalité connue de tous';

Lexie 8 : PONT AÉRIEN au sens de 'liaison régulière par avion entre deux points...';

Lexie 9 : COUPER LES PONTS [avec N] au sens de 'cesser délibérément les relations [avec N] ...';

...

Telle que présentée dans ces exemples, une lexie est une entité trilatérale : elle a

- un sens (le *signifié* saussurien),
- une forme phonique/graphique (le *signifiant* saussurien),
- et un ensemble de traits de combinatoire (le *syntactique* de la théorie Sens-Texte : Mel'čuk 1993 : 117).

Les sens sont présentés – de façon très approximative – par les expressions entre guillemets « sémantiques »; les formes phoniques sont exprimées par

<sup>1</sup> Nous utilisons le terme *lexie* tel que proposé et défini par B. Pottier (1991).

l'écriture orthographique habituelle<sup>2</sup>; et les traits de combinatoire sont spécifiés par l'indication du genre grammatical (PONT est masculin), du régime (*pont* ENTRE N<sup>1</sup> et N<sup>2</sup>; cf. aussi *couper les ponts* AVEC N), etc.

Les lexies ainsi conçues forment la partie primordiale de la langue. En exagérant quelque peu, on pourrait même dire que l'ensemble des lexies est la langue. En effet, une langue est constituée de lexies et de règles servant à la manipulation de ces dernières. Les règles qui réunissent les lexies en syntagmes, les syntagmes en phrases, et les phrases en discours sont donc nettement secondaires par rapport aux lexies – en ce sens que leur nature et leur forme sont déterminées par les lexies.

On peut résumer ces affirmations par la comparaison suivante. L'ensemble des règles d'une langue  $\mathcal{L}$ , c'est-à-dire sa grammaire (au sens large du terme), est aux lexies de  $\mathcal{L}$  ce que l'ensemble des instructions d'assemblage d'un meuble en pièces détachées est à ces pièces. Les instructions d'assemblage sont obligatoirement écrites en fonction des pièces à assembler. Il en est de même pour les règles d'une grammaire : étant donné qu'elles nous permettent d'« assembler » des lexies, elles doivent être formulées en fonction des lexies ! Il s'ensuit que LE LEXIQUE D'UNE LANGUE PRIME LOGIQUEMENT SUR SA GRAMMAIRE. Rappelons, cependant, que ce n'est nullement le point de vue dominant dans la linguistique théorique actuelle, qui met un accent tout particulier sur la syntaxe et relègue le lexique au rôle d'un accessoire négligé, qui doit assurer les « insertions lexicales » après la construction des arbres syntaxiques<sup>3</sup>.

Notre démarche, par contre, s'appuie sur le point de vue inverse. Nous nous situons dans le cadre de la THÉORIE LINGUISTIQUE SENS-TEXTE<sup>4</sup>, où l'on part d'une représentation sémantique pour construire des arbres syntaxiques À L'AIDE DU LEXIQUE.

Avec cette mise en relief de l'importance capitale du lexique dans la langue, on comprend que nous voulons placer l'étude du lexique, c'est-à-dire la lexicologie, au cœur de la linguistique théorique. Cette attitude n'est pas très habituelle non plus. À la fin du XX<sup>e</sup> siècle, la lexicologie a encore peu d'importance et dans les manuels et dans l'enseignement de la linguistique. Par l'ILEC, nous espérons redresser quelque peu la situation.

2 Même si, la plupart du temps, l'écriture orthographique est suffisante pour la prononciation du français, il reste néanmoins de nombreux cas où la transcription phonologique est indispensable. Ainsi, il faut spécifier qu'on dit /pōtozan/ – avec les deux liaisons, mais /pōaerjē/, sans liaison. Pour d'autres langues, la situation peut être différente; pour l'anglais, par exemple, la nécessité de la transcription est tout à fait évidente.

3 Il est vrai que plusieurs études récentes (par exemple, Levin 1993) cherchent à mieux intégrer le lexique dans le modèle global de la langue en lui reconnaissant une place plus fondamentale.

4 Pour la théorie Sens-Texte, voir, entre autres, Mel'čuk 1974a, b, 1978, 1981, 1988 : 43-101 et 1993 : 41-79.

## 2 Dictionnaire, produit final de la lexicologie

Dans la logique de la démarche proposée, nous aboutissons inévitablement à l'affirmation que le dictionnaire est la cible et le produit final des recherches lexicologiques. Pour expliciter cette affirmation, nous procédons en deux étapes : d'abord nous allons éclaircir certains concepts fondamentaux, souvent confondus, de la lexicologie qui ont trait au dictionnaire – dans le but de préciser le concept de dictionnaire lui-même (2.1); puis, nous traitons, en termes assez généraux, de la confection des dictionnaires (2.2). Cela nous amènera à un schéma idéalisé de dictionnaire de langue, qui sera présenté dans le chapitre suivant.

### 2.1 Concepts de base

Pour éviter toute confusion dans l'utilisation des termes centraux – *lexique, vocabulaire, dictionnaire, glossaire* – et pour asseoir notre exposé sur des bases logiques, nous commencerons par leur définition.

Aussi étrange qu'il puisse paraître, on ne trouve pas de définition formelle rigoureuse de ces termes dans les manuels et les livres de référence dans le domaine de la lexicologie et de la lexicographie. Nous nous devons donc d'élaborer les termes ci-dessus à partir de l'usage observé. Cependant, cet usage est extrêmement flou. Même s'il est possible de dégager quelques lignes magistrales dans leur application, ces termes sont néanmoins souvent employés avec plusieurs sens différents, sans aucune distinction nette entre les diverses acceptions, et des termes différents sont en plus utilisés pour la même notion. Nous nous voyons donc forcés de proposer des DÉFINITIONS PRESCRIPTIVES qui stabilisent les usages les plus courants et les plus cohérents, en rejetant tous les autres.

#### Définition I.1 : lexique

|| Nous appelons *lexique* d'une langue  $\mathcal{L}$  l'ensemble de toutes les lexies de  $\mathcal{L}$ .

On peut donc parler de l'étude du lexique du français, dire que le lexique du japonais inclut des verbes spéciaux qui..., etc.

Le lexique (d'une langue) est quelque chose d'objectif, qui fait partie de cette langue et qui existe tout à fait indépendamment du linguiste ou du lexicographe qui veut l'étudier.

#### Définition I.2 : vocabulaire

|| Nous appelons *vocabulaire* (de X) un sous-ensemble du lexique de  $\mathcal{L}$  (déterminé par X).

On distinguera deux sous-sens du terme *vocabulaire*, selon qu'il s'agit d'un individu/d'un (ensemble de) texte(s) X ou d'un domaine X qui détermine le vocabulaire en question :

- Le vocabulaire de l'individu X ou de l'ensemble de textes particuliers X est l'ensemble de TOUTES les lexies de  $\mathcal{L}$  que X utilise/qu'on trouve dans X. Par exemple, *le vocabulaire de Malraux* est un sous-ensemble du lexique français utilisé par Malraux (y compris les conjonctions, les articles, etc.); c'est aussi cela qu'on exprime en disant *Le vocabulaire d'Isobel est époustouflant !, On étudie le vocabulaire de la Chanson de Roland, le vocabulaire des textes en géologie moderne*, etc.
- Le vocabulaire du domaine X est, par contre, l'ensemble des lexies de  $\mathcal{L}$  qu'on utilise UNIQUEMENT pour parler de X; cela exclut nécessairement les lexies utilisées également ailleurs. Ainsi, *le vocabulaire de la cuisine en français* est un sous-ensemble du lexique français utilisé uniquement pour parler de la cuisine; ce vocabulaire ne comprend donc pas les conjonctions, les articles, etc., ainsi que toutes les lexies de la langue commune.

#### Remarques

1. Un domaine peut être identifié par la désignation des personnes impliquées dans ce domaine : *le vocabulaire des amateurs de football en anglais américain, le vocabulaire du (des) chimiste(s), le vocabulaire des Précieuses*, ... Une telle désignation est en fait une référence au domaine plutôt qu'aux individus.
2. La définition proposée est assez flexible pour convenir à des tâches bien différentes. Ainsi, *le vocabulaire de la géologie* est censé n'inventorier que les termes spécifiques pour le domaine de la géologie (l'appartenance d'un terme au domaine spécifié n'est pas, bien sûr, pertinent pour cette discussion). Par contre, *le vocabulaire des textes de la géologie* doit inclure tout ce qui est contenu dans un corpus de textes de référence.

### Définition 1.3 : dictionnaire

|| Nous appelons *dictionnaire de  $\mathcal{L}$*  une description du lexique de  $\mathcal{L}$  – sous la forme d'une énumération (d'une partie relativement importante) de lexies de  $\mathcal{L}$ , dans laquelle chaque lexie est munie d'informations pertinentes.

La définition ci-dessus exige des précisions concernant les deux éléments suivants : partie relativement importante et informations pertinentes.

#### Partie relativement importante

Soit une liste de lexies de  $\mathcal{L}$  auxquelles on a associé les équivalents dans une autre langue ainsi que toutes les informations jugées nécessaires. Si cette liste ne contient qu'une cinquantaine de lexies, personne ne la qualifiera de dictionnaire de  $\mathcal{L}$ ; si, par contre, elle en contient plusieurs milliers, personne n'hésitera à l'appeler ainsi. Cette considération explique l'incise « une partie relativement importante [du lexique de  $\mathcal{L}$ ] » dans la définition. Notons qu'il n'est pas possible

d'indiquer le nombre exact de lexies suffisant pour attribuer à une telle liste sans hésitation l'étiquette de dictionnaire; la qualification 'partie relativement importante (du lexique)' doit donc rester vague. Cependant, tout auteur de dictionnaire, en appelant son ouvrage ainsi, doit déterminer si la partie du lexique consignée est importante – et de quel point de vue. « Mon premier dictionnaire » des élèves français à la petite école ne renferme que quelques centaines de mots, mais c'est une partie du lexique fort importante pour ce niveau. Le titre *Dictionnaire d'anglicismes* est tout à fait légitime, même si leur nombre n'est pas trop élevé, puisque les anglicismes sont considérés comme relativement importants. Il en est de même pour les *Dictionnaires de rimes*, *Dictionnaires de mots croisés*, *Dictionnaires de scrabble*, *Dictionnaires de difficultés de la langue L*, etc.

### Informations pertinentes

Les informations pertinentes dans un dictionnaire peuvent être de nature fort variée : prononciation, données grammaticales, explications sémantiques, synonymes ou antonymes, traductions, étymologies, fréquences... Une des façons (d'ailleurs, la plus habituelle) de classer les dictionnaires est d'après le type d'informations que le dictionnaire associe aux lexies consignées : dictionnaire de prononciation, dictionnaire grammatical, dictionnaire explicatif, dictionnaire de synonymes, dictionnaire bilingue... Un dictionnaire qui fournit toutes les informations ou au moins la plupart des informations pertinentes pour le lexique d'une langue donnée est appelé *dictionnaire de langue*. Dans ce livre, nous ne traiterons que de ce type de dictionnaire.

**NB** : Un dictionnaire doit spécifier explicitement dans son titre le type d'informations pertinentes qu'il contient : *Dictionnaire étymologique*, *Dictionnaire de régimes verbaux*, *Dictionnaire de style*, ... En l'absence de telles spécifications, donc, par défaut, l'« information pertinente » est comprise comme nécessaire et suffisante pour assurer l'utilisation et/ou la compréhension de chaque lexie consignée.

Un dictionnaire doit donc obligatoirement fournir des informations au sujet des lexies; une liste de lexies sans aucune information n'est qu'une liste de mots.

### Remarque

Dans les appellations du type *Dictionnaire des pierres précieuses* ou *Dictionnaire des hommes célèbres*, on utilise le terme *dictionnaire* dans un sens totalement différent de celui que nous venons de définir : il s'agit ici de listes de noms de choses/de personnes plutôt que de listes de lexies comme telles. En nous référant à l'expression allemande *Wörter und Sachen* ('mots et choses'), fort connue dans le domaine lexicologique, ce sont des « dictionnaires » de *Sachen*, alors que, dans le présent ouvrage, nous ne considérons que les dictionnaires de *Wörter*.

### Définition I.4 : glossaire

Nous appelons *glossaire* de X une description du vocabulaire de X – sous la forme d'une énumération des lexies (d'une partie relativement importante) du vocabulaire du domaine X ou de l'œuvre / l'individu X, dans laquelle chaque lexie est munie d'informations pertinentes.

Tout comme le dictionnaire, le glossaire liste des lexies avec des informations pertinentes. Cependant, il ne s'agit pas d'un type particulier de dictionnaire; un dictionnaire couvre la totalité ou une grande partie du lexique d'une langue, alors que le glossaire ne traite que de lexies SPÉCIFIQUES, peu importe leur nombre : par exemple, *Glossaire des termes géologiques*, *Glossaire des tragédies de Racine*, *Glossaire des jurons et des gros mots du français* ou *Glossaire de la culture rizière*. Par conséquent, un glossaire des termes utilisés dans un ouvrage peut se limiter à une vingtaine d'unités, ce qui ne serait aucunement suffisant pour un dictionnaire.

De ces quatre définitions, on peut tirer une conclusion importante : il y a une proportionnalité d'une part, entre les OBJETS DE DESCRIPTION (= lexique et ses sous-ensembles, vocabulaires spécifiques) et, d'autre part, entre les RÉSULTATS DE DESCRIPTION (= dictionnaires et glossaires). Schématiquement, cette proportionnalité s'exprime de la façon suivante :

$$\frac{\text{dictionnaire}}{\text{lexique}} = \frac{\text{glossaire}}{\text{vocabulaire}}$$

Notre intention dans cette étude de lexicologie n'est pas de présenter dans le détail tous les concepts introduits; en fait, les définitions de *vocabulaire* et de *glossaire* n'ont d'autre but que de mieux cerner les concepts centraux de *lexique* et de *dictionnaire*. Dans ce qui suit, nous nous concentrerons uniquement sur ces deux derniers termes.

## 2.2 La confection des dictionnaires

La description du lexique (= de l'ensemble des lexies<sup>5</sup>) d'une langue  $\mathcal{L}$  est, conformément à la définition I.3, un dictionnaire de  $\mathcal{L}$ . Par conséquent, dans notre optique, la lexicologie, qui traite du lexique, vise nécessairement la confection de dictionnaires.

Or, parler de dictionnaire implique au moins deux points fondamentaux qui nous paraissent être en conflit.

D'une part, les dictionnaires – nous nous limitons, comme nous venons de l'indiquer, aux dictionnaires de langue – sont universellement reconnus comme

<sup>5</sup> Nous ne voulons pas dire qu'il s'agit de l'inventaire EXHAUSTIF des lexies de  $\mathcal{L}$ ; tel que stipulé par la Définition I.3, il suffit que cet ensemble soit représentatif.

des sources indispensables d'informations sur le lexique de la langue et comme des instruments de travail irremplaçables pour de nombreuses professions langagières (traducteurs, réviseurs, journalistes, ...). On peut même prétendre que, dans une société technologique moderne, la langue n'existe que dans la mesure où elle est consignée dans un dictionnaire. La confection de dictionnaires s'avère donc comme une tâche de prime importance pour laquelle il existe une forte demande, et la communauté, en général, est prête à y investir temps et argent.

D'autre part, la réalisation d'un dictionnaire est une tâche longue et fort difficile, et elle n'est que très rarement menée à bonne fin. En effet, il est permis d'affirmer, sans malice, que tout dictionnaire est facilement critiquable et que même les meilleurs et les plus reconnus ne sont pas sans failles fort importantes. Le dictionnaire s'avère donc – de par sa nature – comme un instrument nécessairement incomplet et imparfait.

Il est évident qu'il y a contradiction entre ces deux constatations : l'offre ne correspond pas à la demande ! Et il est donc légitime de s'interroger sur les causes.

Nous pouvons signaler au moins deux types de facteurs qui rendent l'élaboration de dictionnaires hasardeuse et ingrate. Les uns sont liés à la substance même de la matière lexicale et forment les causes linguistiques, les autres relèvent de l'environnement social et idéologique et constituent les causes psychologiques.

### 2.2.1 Difficultés linguistiques de la confection de dictionnaires

On peut donner quatre grandes raisons découlant de la nature même de la langue qui rendent l'élaboration d'un dictionnaire extrêmement ardue.

#### 1. La QUANTITÉ D'UNITÉS à traiter

Si l'on admet que l'unité de base – la *lexie* – dans un dictionnaire est une acception d'un mot ou d'une expression figée, on constate immédiatement, en faisant un calcul très simple, que le nombre de ces unités est très élevé. Par exemple, un dictionnaire du français, contenant à peu près 60 000 entrées vedettes, dont chacune a en moyenne quatre à cinq acceptions, renferme par conséquent entre 200 000 et 300 000 acceptions. À cela, il faut encore ajouter un nombre important d'expressions figées et semi-figées, cachées à l'intérieur des articles, ce qui double ou même triple le nombre d'unités. Le lexicologue a donc à traiter à peu près un MILLION de lexies ! C'est évidemment une tâche énorme, de loin plus considérable que celle du linguiste phonologue, morphologue ou syntacticien.

#### 2. La QUANTITÉ D'INFORMATIONS à fournir pour une lexie

Si l'on part de l'hypothèse qu'un dictionnaire de langue doit fournir à l'utilisateur, pour chaque lexie L, toutes les informations dont celui-ci peut avoir besoin pour utiliser L correctement dans n'importe quel contexte, on se rend vite compte que la quantité des renseignements indispensables est extrêmement grande. Pour



chaque lexie à traiter, il s'agit tout d'abord de donner la définition lexicographique, c'est-à-dire écrire une représentation sémantique, qui est en elle-même un objet fort complexe. À la définition s'ajoutent les données du régime syntaxique avec tous les menus détails des conditions d'utilisation contextuelles. De plus, la cooccurrence lexicale restreinte de la lexie doit être spécifiée, et le lexicologue est ici confronté à un ensemble immense de données hétérogènes. On ne doit pas non plus oublier les indications précises sur la flexion, les marques d'usage (caractérisations géographiques, chronologiques, sociales, stylistiques), les données connotatives, etc. Et ce n'est pas tout : il faut encore les spécifications orthographiques, prosodiques, étymologiques, pragmatico-culturelles et même encyclopédiques ! Comme on peut le voir, déjà l'énumération des informations à fournir pour une seule lexie est assez longue. (On peut facilement se rendre compte de la complexité des informations lexicales à présenter en consultant les articles de dictionnaire donnés en exemple au Chapitre V, pp. 195-203.) Et toutes ces informations doivent être fournies pour chacune de ces quelques centaines de milliers de lexies !

### 3. La COMPLEXITÉ de l'information lexicographique

Non seulement les données descriptives à fournir dans un article de dictionnaire, c'est-à-dire les informations lexicographiques, ou *dictionnairiques*, sont extrêmement nombreuses et volumineuses, mais elles sont encore fort difficiles à dégager. Elles sont d'une part abstraites et non directement observables, surtout les données sémantiques. D'autre part, elles dépendent immédiatement des résultats de la linguistique théorique; en effet, elles mettent en jeu toutes les branches de la linguistique – de la sémantique, en passant par la syntaxe et la morphologie, jusqu'à la phonologie et la phonétique (recommandations orthoépiques). Le lexicologue est donc soumis à rude épreuve : il doit être compétent dans tous ces domaines, mais, par-dessus le marché, il doit être apte à opérer les liens nécessaires entre toutes ces spécialités.

### 4. La VARIATION des données lexicales

En plus des difficultés déjà signalées, le lexique à décrire ainsi que les informations à fournir sont instables. D'une part, des lexies apparaissent et disparaissent régulièrement; d'autre part, les informations lexicographiques inhérentes à une lexie particulière ne sont pas moins changeantes : le sémantisme, la syntaxe, la cooccurrence et même la morphologie d'une lexie peuvent varier. Il suffit de penser aux longues discussions à propos de RÉALISER au sens de 'se rendre compte' ou de CONTRÔLE au sens de 'maîtrise [d'un véhicule]', rejetés même de nos jours par certains puristes comme anglicismes inacceptables, alors que leur utilisation est quotidienne depuis des années. Nous avons découvert avec surprise que l'expression *X a coûté la vie à sa mère* au sens de 'la mère de X est morte en lui donnant la vie' produit un désaccord entre divers locuteurs : pour les uns, elle est tout à fait courante, alors que pour d'autres, elle est soit vieillie, soit uniquement littéraire, soit, encore, inusitée. Cependant, le conte de Blanche-

Neige comprend, dans le premier paragraphe, la phrase *Blanche-Neige a coûté la vie à sa mère* ! Même si pour de nombreux locuteurs le régime « après N » avec l'adjectif FÂCHÉ (– *Es-tu fâché après moi ?*) ne pose aucun problème, il ne semble pas acceptable pour d'autres; en tout cas, les dictionnaires courants ne l'indiquent pas. Il n'est pas nécessaire de multiplier les exemples; la preuve que la variation lexicale dépasse infiniment celle trouvée dans les autres domaines de la langue, par exemple, en morphologie ou en syntaxe, est facile à faire.

La variation lexicale s'effectue selon les trois axes majeurs suivants :

- variation SOCIALE ou *sociolectale* : éléments familiers, populaires, argotiques, officiels, ..., langues de spécialités, ...;
- variation GÉOGRAPHIQUE ou *topolectale* : éléments régionaux (pays francophones différents : belgicisms, canadianisms, africanisms, ...; régions différentes de la France : picardisms, parlars du Midi, parisianisms, ...);
- variation CHRONOLOGIQUE ou *chronolectale* : éléments archaïques, vieillis, vieillissants, néologiques, ...

De telles variations sont en plus accentuées par les variations INDIVIDUELLES – le flottement et le désaccord entre les divers locuteurs. Cette dernière variation « transperce » les variations sociolectales, topolectales et chronolectales. Ainsi, l'expression *C'est coulé dans le béton* ('C'est définitif) est perçue comme familière par certains, mais comme neutre par d'autres; un professeur de phonétique française prononce le nom GUILLEMET comme /gilme/ malgré l'indication des dictionnaires, qui est /gijme/; pour le lexème OIGNON, une partie des francophones préfère la prononciation /ɔ̃nɔ̃/, l'autre favorise /wɔ̃nɔ̃/; etc. À tout cela, s'ajoute encore le DÉCALAGE (bien connu) ENTRE L'USAGE LINGUISTIQUE d'un sujet N et le JUGEMENT SUBJECTIF de N sur son propre usage. Très souvent, le locuteur confronté à des expressions prétend les accepter ou les refuser, alors que son propre usage linguistique réel dément l'une ou l'autre de ses affirmations.

Pour les raisons énumérées – quantité d'unités, volume et complexité des informations, variabilité des données – la rédaction d'un dictionnaire ne peut être qu'une tâche hasardeuse et dont le résultat reste constamment à reprendre.

### 2.2.2 Difficultés psychologiques de la confection de dictionnaires

Elles relèvent presque de la quadrature du cercle : il s'agit en fait de réconcilier les exigences des usagers moyens et les nécessités d'une approche scientifique. Le public visé par les dictionnaires en exige une LISIBILITÉ PARFAITE. En effet, un dictionnaire doit satisfaire un large éventail d'usagers, tant le représentant d'une profession libérale que le travailleur manuel. Or cela se trouve en opposition directe avec les exigences scientifiques, qui présupposent une LOGIQUE SOPHISTIQUE (= la rigueur dans le traitement des lexies + une cohérence absolue) conjuguée à une recherche de l'EXHAUSTIVITÉ. Pour respecter toutes ces exigences, il faut recourir à des formalismes, qui engagent le dictionnaire infailliblement sur

la voie des « sciences dures » : on obtient un produit qui rappelle plutôt un traité de mathématiques qu'un dictionnaire traditionnel. Mais le public en général n'est pas enclin à suivre cette voie, et le lexicologue doit chercher un compromis difficile entre les nécessités scientifiques et la réalité sociale. Étant donné l'impossibilité pour la logique d'accepter des compromis, les solutions atteintes par le lexicologue ne sont, d'une part, jamais entièrement satisfaisantes, et d'autre part, la logique continue à pousser vers une formalisation toujours plus développée. On peut se rendre compte, en examinant certains dictionnaires récents, que les progrès dans le domaine sont avec évidence liés à une complexité formelle qui va croissante.

Le conflit entre les exigences scientifiques et celles du public est encore intensifié par les deux facteurs suivants.

Premièrement, les exigences scientifiques découlent nécessairement de la nature même du langage. La linguistique moderne a démontré de façon indubitable la véracité de la thèse saussurienne « La langue est un système où tout se tient ». Aucun dictionnaire ne peut plus en faire fi, et cela veut dire que tout article de dictionnaire doit présenter de façon homogène, cohérente, claire et exhaustive force données hétérogènes, incohérentes, compliquées et variables. Le résultat sera fatalement des articles de dictionnaire fort complexes.

Deuxièmement, il y a le contexte éditorial et surtout commercial. Avec les coûts extrêmement élevés de la préparation et de la fabrication d'un dictionnaire assez complet, aucun éditeur ne peut se payer le luxe d'ignorer les clients potentiels et encore moins leur nombre; mais en même temps, il ne peut non plus se permettre trop d'incohérences et de présentations impressionnistes.

Malgré toutes ces difficultés inhérentes à la tâche, nous croyons qu'on peut à l'heure actuelle produire un ouvrage lexicographique qui réponde – de la meilleure façon possible pour le moment – aux exigences scientifiques ainsi qu'à celles de l'usager moyen. En le disant, nous pensons tout particulièrement au *Dictionnaire explicatif et combinatoire du français contemporain* [= DEC] en cours d'élaboration au GRESLET (Groupe de Recherche en Sémantique, Lexicologie Et Terminologie) de l'Université de Montréal (voir Mel'čuk *et al.* 1984, 1988, 1992). En nous appuyant sur notre expérience de plus de quinze ans de recherches lexicographiques ainsi que sur les résultats d'autres chercheurs, nous pouvons proposer un schéma idéalisé du dictionnaire général de langue. Pour éviter tout malentendu, précisons qu'il s'agit :

- d'un dictionnaire DE LANGUE, qui est exemplifié à partir du français contemporain, mais que rien ne limite au français : la même méthodologie est applicable aux autres langues;
- d'un dictionnaire de la langue ACTUELLE (et non d'un trésor rassemblant les données linguistiques de toutes les époques, des origines de la langue à nos jours);

- d'un dictionnaire MONOLINGUE, mais tel que la description donnée s'applique à l'élaboration d'un dictionnaire bilingue.

Dans les pages qui suivent, nous traiterons du DEC en détail, tant d'un point de vue théorique que d'un point de vue descriptif. En fait, L'ENSEMBLE DU LIVRE EST CONSACRÉ AU DEC. Nous présenterons la théorie sous-jacente, nous donnerons les principes rédactionnels et exposerons les formalismes utilisés, et enfin nous analyserons en profondeur quelques articles de dictionnaire du DEC. Mais avant de le faire, nous croyons nécessaire de fournir le CADRE GÉNÉRAL dans lequel se situe la réflexion incarnée dans l'ILEC. En d'autres termes, nous devons expliquer pourquoi, dans notre approche, un manuel de lexicologie n'est en réalité que la description d'un dictionnaire idéalisé.

### 3 Lexicologie, discipline carrefour

Nous aborderons la discussion de la lexicologie en examinant les trois points suivants :

- les rapports entre la lexicologie et la lexicographie (3.1);
- le caractère spécifique de la lexicologie par rapport aux autres disciplines linguistiques (3.2);
- et le concept de dictionnaire idéalisé comme objet de la lexicologie (3.3).

#### 3.1 *Lexicologie vs lexicographie*

Traditionnellement, on oppose la lexicologie à la lexicographie comme suit : la lexicologie est une discipline théorique qui a pour objet l'étude générale du lexique, alors que la lexicographie est une discipline appliquée qui a pour objet l'élaboration de dictionnaires; la lexicographie doit faire siens et utiliser en pratique les résultats théoriques dégagés par la lexicologie.

Nous sommes pleinement d'accord avec cette position, mais nous nous voyons forcés d'y ajouter quelques précisions que nous jugeons indispensables. Nous le faisons en traçant un parallèle entre ces deux disciplines linguistiques et ce que l'on observe dans les sciences dites dures.

La différence entre la lexicologie et la lexicographie est la même que celle établie entre la physique et le génie, par exemple. La physique se préoccupe des lois générales de la matière, de l'énergie et du mouvement; le génie s'intéresse à l'application industrielle de ces lois : à la construction des ponts, à la fabrication d'avions, à l'extraction du pétrole, ... Comme on le sait, la physique est constituée de deux parties indissociables : la physique théorique et la physique expérimentale. La première se réalise sous forme de réflexions, d'équations et de calculs; c'est typiquement une science abstraite. La seconde, par contre, pénètre dans le

concret : sa réalité est basée sur tout un appareillage (on peut penser, par exemple, aux accélérateurs de particules, etc.). Aucune de ces deux physiques n'est pourtant autonome : la physique théorique ne peut pas exister sans mise à l'épreuve – par la physique expérimentale – de ses hypothèses et modèles, et la physique expérimentale a besoin d'hypothèses et de modèles à mettre à l'épreuve. Les deux physiques s'autofécondent, et ce fait est généralement reconnu. Ce que l'on tend à oublier, c'est que la même situation se présente en génie. La construction des ponts, la fabrication d'avions et l'extraction du pétrole ont bien leurs théories respectives, tout à fait abstraites, tout comme leurs pratiques industrielles, tout à fait concrètes. De la même façon que la physique se divise en physique théorique et physique pratique, c'est-à-dire expérimentale, le génie se divise, lui aussi, en génie théorique et en génie pratique, c'est-à-dire industriel. On observe donc l'intersection de deux oppositions : « science ~ application » et « théorique ~ pratique ». Les sciences, ainsi que leurs applications ont un côté théorique et un côté pratique, avec une rétroaction permanente de l'un sur l'autre.

Il convient de souligner qu'il en est parfaitement de même pour la paire « lexicologie ~ lexicographie ». La lexicologie (comme la physique) a sa théorie – l'étude générale du lexique des langues naturelles, au niveau formel et abstrait, tout en ayant, elle aussi, sa pratique – la description expérimentale du lexique. Mais celle-ci, conformément à la définition I.3 (p. 19), aboutit inévitablement à un dictionnaire ! Si la lexicologie théorique se doit de dégager les lois générales du lexique et d'en proposer une formalisation, la lexicologie expérimentale, par contre, a pour tâche de vérifier et de préciser ces formulations en les utilisant dans la description des ensembles des lexies. Son but final est donc la construction d'un DICTIONNAIRE IDÉALISÉ – le prototype des dictionnaires pratiques, un banc d'essai et une source capitale des données pour les théories lexicologiques.

Cependant, la lexicographie (comme le génie) a, elle aussi, sa théorie – l'étude spécifique du lexique d'une langue donnée, au niveau substantiel et concret. C'est une théorie de production des dictionnaires pratiques. Entre autres, elle doit se préoccuper des problèmes typographiques, commerciaux et pédagogiques sous-jacents à la publication du dictionnaire, aspects que la théorie lexicologique néglige. En même temps, la lexicographie a un côté pratique évident : la production effective de dictionnaires destinés à l'usage de divers publics. (C'est exactement cette activité qui correspond à ce que B. Quemada (1987 : 235) a justement nommé *dictionnairique*.)

Dans cet ouvrage, nous ne traiterons pas du tout de lexicographie pratique; notre seul objet est la lexicologie, mais sous les deux aspects mentionnés : la lexicologie théorique et la lexicologie expérimentale, c'est-à-dire descriptive.

### 3.2 *La spécificité de la lexicologie*

Même si la division introduite entre la lexicologie théorique et la lexicologie descriptive peut sembler novatrice, elle est observée de fait dans les autres disciplines linguistiques. Par exemple, en syntaxe, on a, d'une part, des études scientifiques « pures », qui concernent à la fois la théorie et sa vérification par une description expérimentale, et, d'autre part, des applications pratiques visant l'élaboration de manuels et de livres de référence sur la syntaxe d'une langue spécifique à l'usage du public (ces applications ayant, bien entendu, leur propre côté théorie et leur propre côté industriel), ainsi que la construction des systèmes d'analyse et de synthèse syntaxiques automatiques pour le traitement des langues par ordinateur. Il en est de même pour la sémantique, la morphologie ou la phonologie/phonétique.

En dépit des similitudes indiquées, la lexicologie n'est cependant pas une discipline comme les autres disciplines linguistiques. Ce n'est pas par hasard qu'elle est encore aujourd'hui une « parente pauvre » dans la famille linguistique : comme nous l'avons déjà dit, elle n'a que peu de place dans les cursus universitaires de linguistique. Cela tient encore à sa nature même : la lexicologie, contrairement aux autres disciplines, n'est pas une discipline linguistique autonome.

La sémantique, la syntaxe, la morphologie et la phonologie/phonétique – les disciplines de base – ont chacune pour objet une composante particulière de la langue. En effet, nous considérons la langue naturelle (conformément à la théorie Sens-Texte) comme un système de modules autonomes, un module étant un ensemble de règles chargées d'établir les correspondances entre les représentations linguistiques de deux niveaux adjacents dans le système adopté des quatre niveaux majeurs de représentation linguistique :

niveau sémantique ~ niveau syntaxique ~ niveau morphologique ~ niveau phonologique.

Ainsi, le module ou la composante sémantique assure la correspondance entre la représentation sémantique d'un énoncé ( $\approx$  son réseau sémantique) et sa représentation syntaxique ( $\approx$  son arbre de dépendances); le module syntaxique assure la correspondance entre la représentation syntaxique d'un énoncé et sa représentation morphologique ( $\approx$  sa chaîne de mots); enfin, le module morphologique assure la correspondance entre la représentation morphologique d'un énoncé et sa représentation phonologique ( $\approx$  chaîne de phonèmes). La sémantique est donc en charge du module sémantique, la syntaxe, du module syntaxique, et la morphologie avec la phonologie, des modules morphologique et phonologique. Ce sont ces disciplines linguistiques qui traitent des transitions, c'est-à-dire des correspondances mentionnées ci-dessus. Mais la lexicologie n'a pas de module linguistique correspondant et n'est pas liée à une transition particulière entre deux niveaux donnés de représentation. Par contre, la lexicologie prend en charge l'unité centrale de tous ces niveaux – la lexie. Elle analyse et décrit le comportement des lexies à travers tous leurs « avatars » d'un niveau à

l'autre; elle munit chaque lexie des informations nécessaires pour que cette lexie puisse être proprement traitée par la sémantique, la syntaxe, la morphologie et la phonologie. Pour ce faire, la lexicologie doit prendre en considération, et cela dans tous les détails, les traitements effectués par les quatre modules sous-jacents aux quatre disciplines linguistiques de base. Ainsi, s'élabore une boucle de rétroaction entre la sémantique, la syntaxe, la morphologie et la phonologie, d'une part, et la lexicologie, d'autre part. En se fondant sur cette rétroaction, la lexicologie théorique détermine le contenu et la forme des informations à assigner aux lexies. Une lexie munie de toutes ces données, organisées de façon systématique et cohérente, n'est pas autre chose qu'un ARTICLE DE DICTIONNAIRE. Le regroupement des lexies ainsi décrites constitue bien évidemment un dictionnaire. On voit donc que le dictionnaire est infailliblement l'objet d'étude, l'instrument de recherche et le produit final de la lexicologie. Comme nous l'avons déjà dit, il s'agit d'un DICTIONNAIRE IDÉALISÉ, qui ne tient nullement compte de considérations pratiques et commerciales et qu'on pourrait baptiser « dictionnaire du lexicologue ».

### 3.3 Le dictionnaire idéalisé : le DEC

Le « dictionnaire du lexicologue », contrairement au « dictionnaire du lexicographe », est un dictionnaire bien spécial; on peut même dire qu'un tel dictionnaire n'existe pas encore et que l'élaboration du DEC (mentionné plus haut) est la première tentative d'en créer un. En fait nous allons, dès à présent, assimiler le dictionnaire idéal de la lexicologie au DEC, et par conséquent, nous devons indiquer les deux caractéristiques du DEC qui autorisent cette assimilation.

D'une part, la création du DEC est une entreprise purement scientifique, un travail de recherche fondamentale. Autrement dit, il s'agit d'une étude très pointue de l'énorme domaine de la sémantique lexicale de la langue  $\mathcal{L}$ , ainsi que de la syntaxe de  $\mathcal{L}$ , surtout en ce qui concerne le comportement des lexies individuelles, et même de la morphologie et de la phonologie de  $\mathcal{L}$ . Pour illustrer la complexité linguistique de la tâche rédactionnelle du DEC, nous citerons trois exemples :

#### 1) Passivation

Soit la paire de phrases (1) :

- (1) *Ce tonneau peut contenir cent litres de vin.*  
 ~ *Cent litres de vin peuvent être contenus DANS ce tonneau.*

La relation entre les deux phrases est celle entre l'actif et le passif. Cependant, le complément d'agent est introduit ici par la préposition DANS, tournure non signalée dans les descriptions grammaticales usuelles du français. Le lexicologue qui s'occupe des lexies du type CONTENIR, etc. doit au préalable s'attaquer au problème du passif et même plus généralement de la voix en français.

## 2) Réflexivation

Soit les trois paires de verbes :

- (2) RASER ~ SE RASER,      REMPLIR ~ SE REMPLIR,  
COUCHER ~ SE COUCHER.

La grammaire traditionnelle se contente d'affirmer que le deuxième verbe de chaque paire est le réfléchi du premier. Ceci est bien entendu exact, mais nettement insuffisant pour le traitement des lexies de ce type dans un DEC. En fait, les rapports sémantiques entre les verbes de chaque paire sont tout à fait différents.

Dans la paire RASER ~ SE RASER, la description sémantique des deux membres est entièrement identique, parce que l'action de Jean qui rase Pierre et celle de Jean qui se rase sont les mêmes; la différence est au niveau de la joue où se place cette action. Autrement dit, SE RASER est le vrai réfléchi de RASER.

Dans la paire REMPLIR ~ SE REMPLIR, les choses se présentent de façon différente : 'X se remplit de Y' = 'X commence à être plein de Y'; par exemple, *Le caniveau se remplit d'eau* veut dire 'Le caniveau commence à être plein d'eau'. SE REMPLIR s'avère sémantiquement plus simple que REMPLIR, parce que 'Z remplit X de Y' = 'Z cause que X commence à être plein de Y' = 'Z cause que X se remplit de Y'; ainsi, *Jean remplit le caniveau d'eau* signifie 'Jean cause que le caniveau se remplit d'eau'. En clair, REMPLIR est le causatif de SE REMPLIR, ce qui n'est pas du tout le cas pour RASER et SE RASER.

La paire COUCHER ~ SE COUCHER manifeste une nouvelle différence : 'Z couche X' ≠ 'Z cause que X se couche' ! En effet, *Jean couche Pierre* ne veut pas dire 'Jean cause que Pierre se couche'. Si *Pierre se couche* signifie 'Pierre se cause qu'il est couché' (on a donc une autocausion ici), *Jean couche Pierre* veut dire 'Jean cause que Pierre est couché' (c'est de la causion directe). Dans cette paire, aucun des deux verbes ne peut être décrit par l'autre<sup>6</sup>.

En résumé, on peut affirmer que, sémantiquement, RASER et SE RASER sont identiques; REMPLIR inclut SE REMPLIR ('remplir' ⊃ 'se remplir'); et COUCHER et SE COUCHER montrent une intersection ('coucher' ∩ 'se coucher' = 'être couché').

On voit à l'évidence la difficulté à laquelle doit faire face le lexicologue du DEC, qui a l'obligation de décrire de façon cohérente et exhaustive les verbes réfléchis.

6 Lorsqu'un parent dit *J'ai couché les enfants* ou *J'ai été coucher les enfants*, il peut s'agir tout simplement d'une incitation aux enfants d'aller au lit, donc d'aller SE COUCHER. Cependant, la contradiction qu'on pourrait voir ici avec notre description n'est qu'apparente. De telles phrases s'appliquent habituellement à de jeunes enfants, qu'on présente ainsi comme incapables de se coucher par eux-mêmes. (Un père qui vient d'envoyer au lit des adolescents bruyants n'utiliserait pas ces phrases.) Autrement dit, on observe ici l'usage expressément vague du verbe COUCHER, usage typique de la langue en général.



### 3) Prosodisation

Les phrases (3) :

- (3) a. *Mon père est un VRAI médecin* 'Mon père exerce réellement la profession de médecin'.  
 ~ *Mon père est un vrai MÉDECIN* 'Mon père, qui n'est pas médecin, a des aptitudes pour soigner'.
- b. *C'est un VRAI Picasso* 'C'est un tableau de Picasso'.  
 ~ *C'est un vrai PICASSO* 'C'est un tableau qui n'est pas de Picasso, mais qui ressemble à un tableau de Picasso'.

permettent de dégager deux lexies VRAI, dont les sens différents entraînent la différence sémantique observée entre nos paires de phrases. Le fait important est que ces deux sens différents sont indissolublement liés à deux prosodies différentes. La description lexicographique de ces lexies doit par conséquent spécifier explicitement ces prosodies. Mais à l'heure actuelle, le lexicologue ne dispose pas encore d'un système de transcription suprasegmentale (= prosodique) développé au même degré que la transcription phonologique segmentale; pourtant le lexicologue du DEC ne peut contourner la question <sup>7</sup> ! (Voir aussi Ch. III, 3.2, p. 71.)

Malgré la brièveté de notre démonstration, on peut se rendre facilement compte dans quelle mesure l'élaboration d'un DEC est avant tout un travail de linguistique, et plus spécifiquement, une tâche de linguistique théorique. Nous croyons qu'on ne devrait aucunement qualifier ce travail, comme on le fait encore trop souvent, de simple « compilation » ou de « rédaction » ! Nous aimerions donc affirmer :

|| L'élaboration d'un DEC exige de la recherche linguistique de haut niveau.

En plus, le lexicologue du DEC doit négliger les considérations de temps et de coût : il a l'obligation impérieuse de fournir une description valide et complète du lexique de  $\mathcal{L}$  – coûte que coûte. Les restrictions typographiques, pédagogiques et financières, qui constituent un cadre de travail habituel pour les lexicographes (pratiques), doivent être mises de côté dans l'élaboration d'un DEC, de sorte que nous pouvons affirmer :

|| Le travail sur un DEC n'est pas subordonné à des considérations matérielles.

À présent, après avoir ainsi caractérisé la PHILOSOPHIE du DEC, ou bien notre cadre général de réflexion, nous pouvons aborder la STRATÉGIE de l'approche proposée. Le schéma idéalisé d'élaboration d'un DEC comporte les cinq éléments majeurs suivants :

- Principes de rédaction;
- Nomenclature;

<sup>7</sup> Le problème de la représentation des prosodies dans le dictionnaire a été formulé et traité dans Apresjan 1990a.

- Documentation;
- Détermination de l'entrée;
- Structure d'une entrée.

Nous analysons chacun de ces éléments dans une section spécifique du Chapitre II. Mais avant de le faire, nous devons préciser deux concepts indispensables pour les discussions à venir : la microstructure et la macrostructure d'un dictionnaire.

Il est évident que tout dictionnaire et plus spécifiquement encore un DEC ne peut être une collection disparate de données lexicographiques, mais doit être un système bien organisé, c'est-à-dire ayant une *STRUCTURE* complètement déterminée. Une telle structure se manifeste sur deux plans, d'une part au niveau de l'article de dictionnaire comme tel et d'autre part, au niveau du dictionnaire comme ensemble d'articles de dictionnaire.

Chaque article du DEC est élaboré suivant un schéma strict préalablement défini, ce qui assure à l'article une structuration rigide. En d'autres mots, un article de dictionnaire possède une structure dont les éléments sont les données lexicographiques; cette structure est appelée la *microstructure* du dictionnaire, en l'occurrence du DEC.

Le DEC dans son entièreté est, lui aussi, élaboré suivant un schéma préalable précis et manifeste donc également une structuration rigide : le DEC comme tel possède une structure dont les éléments sont des articles de dictionnaire; il s'agit, pour ainsi dire, d'une structuration du deuxième ordre. La structure d'un dictionnaire entier est appelée la *macrostructure* du dictionnaire, en l'occurrence du DEC<sup>8</sup>.

Dans les pages qui suivent, ce que nous disons à propos des principes généraux sous-jacents à l'élaboration du DEC réfère à la fois à la microstructure et à la macrostructure.

Deux remarques terminologiques importantes

1. Pour alléger notre exposé, nous nous contenterons d'utiliser dorénavant le terme *lexicologie* chaque fois qu'il s'agit en fait de lexicologie ET de lexicographie. Ainsi, par un abus de langage évident et innocent, la lexicologie englobera – sauf mention expresse du contraire – les deux domaines. La même remarque s'applique au terme *lexicologue* (qui tiendra lieu, là où c'est commode, du couple *lexicologue* et *lexicographe*).
2. L'adjectif *lexicographique* sera utilisé surtout avec le sens 'ayant trait au dictionnaire'. Par exemple, *description lexicographique de la lexie L* veut dire 'description de L par un <des> article<s> de dictionnaire'.

---

8 Les termes *microstructure* et *macrostructure* (d'un dictionnaire) ont été proposés par J. Rey-Debove (1971). Pour une discussion plus développée de ces notions, voir Hausmann & Wiegand 1989.

## CHAPITRE II

# CARACTÉRISATION GÉNÉRALE DU DICTIONNAIRE EXPLICATIF ET COMBINATOIRE

Avant d'aller plus loin, nous aimerions d'abord souligner ici que les caractéristiques majeures du DEC que nous présentons ne sont pas toutes des innovations fondamentales. Au contraire, certains principes sont connus et appliqués dans les dictionnaires depuis fort longtemps. Comme nous ne faisons pas une analyse systématique des dictionnaires existants, nous nous permettons de ne pas donner, là où elles s'imposeraient, de références nécessaires. Ainsi, quand nous introduisons une méthodologie comme particulière au DEC, cela n'exclut absolument pas son utilisation antérieure par d'autres lexicologues. Nous avons emprunté aux meilleurs ouvrages en faisant nôtres les bonnes idées et les descriptions intéressantes que nous avons trouvées chez nos prédécesseurs.

Cependant, nous croyons que l'originalité du DEC réside dans le fait que c'est la première tentative d'élaboration d'un dictionnaire de langue formel, ancré dans une théorie linguistique, elle aussi formelle. Sous l'angle d'une telle approche, il nous semble important et même indispensable de réunir, DANS UNE MÊME PERSPECTIVE LOGIQUE, les principes, les règles et les critères lexicographiques, connus et nouveaux, qui sont actualisés dans l'élaboration du DEC.

Comme nous nous lançons dans une entreprise lexicologique formelle, nous nous devons, conformément à notre propre attitude, d'être précis et rigoureux, donc, pour ainsi dire, de « prêcher par l'exemple ». Il nous faut donc dès maintenant expliquer la différence entre principe, règle et critère – trois termes importants qui reviennent souvent dans tout notre exposé.

Par *principe*, nous entendons une affirmation de type axiomatique que nous posons, sans justification spéciale, comme notre point d'ancrage. Les principes sont très généraux et, par conséquent, abstraits; ils limitent la marge de manœuvre du lexicologue.

Une *règle* est une prescription qu'un lexicologue se doit de suivre s'il veut obtenir certains résultats escomptés dans son traitement du matériel lexical. Les règles sont plutôt spécifiques et, par conséquent, concrètes; elles dictent une conduite au lexicologue.

Un *critère* formule un test qui permet au lexicologue de dégager certaines propriétés implicites des lexies ou de les systématiser. Les critères sont également relativement concrets; ils organisent l'analyse du lexique par le lexicologue.

De plus, comme nous parlerons d'articles de dictionnaire, précisons immédiatement qu'un article de dictionnaire est consacré à une seule lexie appelée *lexie vedette* (les définitions élaborées seront données plus loin).

Ainsi équipés, nous pouvons maintenant présenter les principes concernant les trois aspects fondamentaux de l'élaboration d'un DEC : la rédaction (1), la nomenclature (2) et la documentation (3).

## 1 Principes de rédaction du DEC

Même si tous les dictionnaires ont toujours cherché à respecter le plus possible une logique de rédaction et de présentation de leurs articles, le DEC est probablement le premier à ériger la rigueur logique en loi absolue de rédaction. Aucune entorse à cette loi ne doit être tolérée, même au risque d'augmenter la complexité de la description. En poussant les choses un peu loin, nous pouvons même dire que, dans le DEC, la logique prime même sur les constats linguistiques. Cela implique qu'il peut être préférable de commettre une erreur linguistique (= refléter une intuition fautive) que de pécher contre la logique. La rigueur logique et le système de description fondé sur celle-ci garantissent qu'une erreur de fait sera inévitablement découverte, alors qu'une erreur de logique rendrait la description non vérifiable.

L'exigence générale de rigueur logique est concrétisée par les quatre principes rédactionnels suivants, inhérents au DEC :

- principe de formalité (1.1);
- principe de cohérence (1.2);
- principe d'uniformité (1.3);
- principe d'exhaustivité (1.4).

Donc, pour nous

RIGoureux = FORMEL + COHÉRENT + UNIFORME + EXHAUSTIF.

### 1.1 Principe de formalité

|| Toute description dans un DEC est strictement formelle.

Dans le présent cadre, « formel » se manifeste par les deux caractéristiques suivantes.

Premièrement, toute description lexicographique est effectuée dans un MÉTALANGAGE FORMEL préétabli. Bien entendu, les dictionnaires existants emploient un certain métalangage lexicographique, mais généralement celui-ci se limite à la morphologie (indication du type de déclinaison et de conjugaison) et aux marques d'usage; dans certains cas, le dictionnaire formalise aussi la présentation de la prononciation (phonétique/phonologie). Sauf de rares exceptions<sup>1</sup>, le

1 Par exemple, pour la définition, le *Longman Dictionary of Contemporary English* (1987) utilise environ 2 000 mots, seuls admis dans les définissants de toutes les lexies consignées.

sémantisme, la syntaxe et surtout la cooccurrence lexicale de la lexie vedette ne sont pas traités à l'aide d'un métalangage précis et assez riche. Le DEC, au contraire, met l'accent sur un métalangage formalisé suffisamment expressif pour la description sémantique (la définition), pour la description syntaxique (le régime) et pour la description cooccurentielle (les fonctions lexicales). Le ou plus exactement les métalangages du DEC sont donnés par des règles de formation strictes et souples, qui permettent au lexicologue de décrire, de façon formelle, tous les phénomènes observés. On trouvera plus loin (Ch. III, 3.3 – 3.5) une caractérisation détaillée de ces métalangages.

Deuxièmement, toute description lexicographique est EXPLICITE : rien n'est laissé à l'intuition de l'utilisateur. Ainsi, le DEC ne peut se limiter à définir MAGAZINE comme 'publication périodique généralement illustrée', puisque cette définition ne distingue pas entre magazines, d'une part, et revues et journaux illustrés, d'autre part. Un magazine est différent d'un journal, même illustré, parce que les feuilles d'un magazine sont nécessairement attachées les unes aux autres (contrairement à un journal); un magazine est différent d'une revue parce qu'un magazine est destiné avant tout à divertir (cf. \**magazine mathématique* vs *revue mathématique*). Ces deux traits – format cahier et pour divertir – doivent être rendus explicites dans une définition du DEC. Par conséquent, la définition de MAGAZINE doit être comme suit :

MAGAZINE = 'publication périodique, sous forme d'un cahier, généralement illustrée<sup>2</sup>, destinée avant tout à divertir'.

Bien mieux, le DEC ne peut utiliser, dans l'entrée MAGAZINE, en tant que seul exemple illustratif, l'expression *magazine féminin*, parce qu'elle n'est pas explicite pour au moins trois raisons :

1. L'expression *magazine féminin* pourrait signifier soit 'magazine traitant de femmes' (cf. *magazine sportif* = '... traitant de sports'), soit 'magazine s'adres-

(suite note 1)

L'emploi rigoureux de ce métalangage définitoire a, comme le dit dans sa préface R. Quirk, «in many cases resulted in a fresh and revealing semantic analysis». Pour la combinatoire syntaxique, ce même dictionnaire a recours à un codage spécial, introduit par A.S. Hornby (*Oxford Advanced Learner's Dictionary of Current English*, 1948). Chaque code identifie une construction syntaxique parmi les quelques dizaines de constructions listées et exemplifiées au début du dictionnaire. Ainsi, le verbe PLUNGE 'plonger' est marqué «X9», ce qui veut dire que ce verbe s'utilise avec un CO<sup>dir</sup> et nécessite un circonstanciel de direction ('plonger qqch. quelque part').

Un dictionnaire plus récent de la même maison, *Longman Language Activator*, 1993, a encore davantage accentué la tendance à la formalisation du métalangage lexicographique. Non seulement il utilise un vocabulaire définitoire bien contrôlé de 1052 lexies, mais il donne en même temps une formalisation poussée de la description de la cooccurrence syntaxique et lexicale, ainsi que des champs sémantiques et lexicaux, etc. En fait, nous pensons que l'*Activator* est un dictionnaire modèle qui correspond bien à nos préoccupations, en combinant, de façon judicieuse, la manière de répondre aux exigences formelles et la présentation pédagogique. Il y a ici une preuve «vivante» de la possibilité d'élaborer un dictionnaire logique et formel qui soit accessible au grand public.

- 2 Cf. *magazine humoristique*, qui peut ne pas comporter d'illustration, *magazine de mots croisés*, etc.

sant aux femmes') (cf. *magazine pour enfants* = '...s'adressant aux enfants'); le DEC doit donc préciser qu'il s'agit seulement de 'magazine s'adressant aux femmes'.

2. Le DEC indiquera encore qu'on ne peut dire par analogie *\*magazine enfantin*, mais qu'on doit dire *magazine pour enfants*, alors que les deux expressions *magazine féminin/magazine pour femmes* sont possibles.
3. Il faut en plus signaler qu'on utilise [*magazine*] *politique/sportif* pour dire 'consacré à la politique/aux sports'.

L'analyse de ces exemples impose au lexicologue du DEC la définition finale suivante :

MAGAZINE = 'publication périodique, sous forme d'un cahier, généralement illustrée, (consacrée à un sujet déterminé X et) destinée avant tout à divertir le public Y'.

La composante entre parenthèses n'est pas obligatoire; en effet, un magazine peut traiter de sujets divers.

Pour les variables X et Y, qui apparaissent dans la définition, il faut spécifier – dans la sous-zone spéciale de l'article de dictionnaire appelée *Régime* (voir Ch. III, 3.4, p. 119) – la façon de les exprimer : X = *politique, sportif, de cinéma* (<*\*cinématographique*>), *de théâtre* (<*\*théâtral*>), *humoristique, ...*; Y = *féminin/pour femmes, pour enfants, pour les jeunes, ...* Bien entendu, tous ces renseignements seront présentés dans un métalangage spécial, du type mentionné ci-dessus.

Comme nous l'avons déclaré au début de cette sous-section, toute description dans le DEC doit être formelle. Par *être formel*, il faut entendre :

FORMEL = ÉCRIT DANS UN MÉTALANGAGE PRÉÉTABLI + EXPLICITE.

## 1.2 Principe de cohérence interne

La rigueur logique appliquée à un article de dictionnaire (donc au niveau de la microstructure) entraîne inévitablement la nécessité d'une cohérence interne :

|| Dans un DEC, les descriptions sémantique, syntaxique et cooccurentielle de la lexie vedette L montrent un accord complet.

Par « accord », nous entendons ici une correspondance réciproque – explicitement indiquée – entre les composantes sémantiques dans la définition de L, ses dépendants syntaxiques actanciels (= actants syntaxiques profonds)<sup>3</sup> et ses cooccurents lexicaux.

3 Un *actant sémantique* de L est un sens qui correspond à une des variables dans la définition de L ; cette variable représente un argument (au sens logico-sémantique) du prédicat 'L' qui est le sens de L. Un *actant syntaxique profond* de L est un syntagme qui exprime, au niveau syntaxique profond, un actant sémantique de L, voir p. 117. Pour les caractérisations plus détaillées, voir Mel'čuk 1993 : 47 *ssq.*

Ainsi, dans le cas du lexème MAGAZINE, les variables X et Y dans la définition assurent le lien entre le lexème lui-même et ses actants syntaxiques dans la phrase; de plus, pour chaque actant syntaxique, le régime spécifie l'expression correcte à la surface, c'est-à-dire dans la réalisation concrète.

Prenons un autre exemple. Supposons un dictionnaire qui définit le nom CÉLIBATAIRE comme suit :

CÉLIBATAIRE<sub>N</sub> = 'personne en âge d'être mariée qui n'a jamais été mariée'.

Cette définition serait en soi tout à fait acceptable. Cependant, elle ne tient nul compte des deux cooccurrents lexicaux usuels : VIEUX et ENDURCI. Le premier problème d'une telle définition est que ces adjectifs ne s'appliquent normalement qu'aux hommes :

(4) Pierre est un  $\left\{ \begin{array}{l} \text{vieux célibataire.} \\ \text{célibataire endurci.} \end{array} \right\} \sim * \text{Marie est une } \left\{ \begin{array}{l} \text{vieille célibataire.} \\ \text{célibataire endurcie.} \end{array} \right\}$

Cette particularité n'est pourtant pas prévue par la définition.

Le second problème de la définition considérée est que, l'adjectif ENDURCI étant perçu comme un intensificateur (*célibataire endurci* = 'très célibataire'), la définition citée ne contient pas de composante qui puisse accepter cette intensification.

Le principe de cohérence ne permet pas une définition qui ne prend pas en charge ces deux problèmes. Dans le DEC, l'article CÉLIBATAIRE devra obligatoirement refléter les deux faits mentionnés : *vieux célibataire* / *célibataire endurci* ne s'applique qu'aux hommes; *endurci* est un intensificateur. Une première solution consisterait à créer un autre lexème – CÉLIBATAIRE2, avec une définition adaptée aux deux adjectifs, soit 'HOMME en âge d'être marié qui n'a jamais été marié et qui VEUT rester tel' (les adjectifs VIEUX et ENDURCI ne seront compatibles qu'avec CÉLIBATAIRE2, et la composante 'veut' acceptera l'intensification). Mais cette solution se montre inacceptable – car le nom CÉLIBATAIRE utilisé seul, sans un des adjectifs en cause, n'a pas le sens 'homme en âge d'être marié qui n'a jamais été marié et qui veut rester tel'. La phrase (5a) en fournit la preuve :

(5) a. ?*Pierre ne l'épousera jamais, c'est un célibataire.*

Cette phrase est absurde, alors que la phrase (5b) est tout à fait normale :

b. *Pierre ne l'épousera jamais, c'est un célibataire endurci.*

Nous pouvons en conclure que le sens de CÉLIBATAIRE<sub>N</sub> ne contient pas la composante (... qui veut rester tel'. L'adjectif ENDURCI n'est donc pas un simple intensificateur de CÉLIBATAIRE, puisque ce dernier ne possède pas de composante à intensifier, mais a un sens indépendant. (Un intensificateur a un sens « dépendant » : intensificateur **de quoi** ?) Cela impose l'autre solution : garder un seul lexème CÉLIBATAIRE<sub>N</sub> et consigner les adjectifs en question dans l'article de dictionnaire de CÉLIBATAIRE<sub>N</sub>, en les munissant des restrictions correspondantes. Nous écrirons donc pour CÉLIBATAIRE<sub>N</sub> :

qui, étant un homme,

est C. depuis longtemps : vieux | prépos

qui, étant un homme,

veut rester C. :                   endurci

[La barre verticale | sépare l'élément lexical présenté à gauche des conditions de son emploi listées à droite. L'abréviation « prépos » indique que l'adjectif en cause est toujours préposé : pour garder le même sens, on ne peut pas dire \**célibataire vieux*. Comme la plupart des adjectifs français sont normalement postposés, on ne met aucune indication pour *endurci*; la postposition d'un adjectif est marquée par défaut.]

Le principe de cohérence, qui exige que, dans un article de dictionnaire, « tout colle », ne permet pas d'avoir dans un même article la définition et les cooccurrents cités sans démontrer l'existence d'un lien explicatif. Cela nous a conduits à une petite recherche, qui a abouti à la solution suivante : la définition de CÉLIBATAIRE<sub>N</sub> citée plus haut est parfaitement valable et les deux adjectifs « problématiques » doivent être décrits de la façon indiquée – par des *fonctions lexicales non standard*, voir Ch. III, 3.5.3, p. 150.

Pourtant, comme on peut s'y attendre, d'autres lexies exigent d'autres traitements. Ainsi, par exemple, pour l'expression *grand blessé*, le DEC favorise une solution légèrement différente. On ne peut décrire le sémantisme de l'adjectif GRAND de cette expression de façon indépendante, comme nous l'avons proposé pour VIEUX et ENDURCI avec CÉLIBATAIRE. L'expression *grand blessé* signifie 'blessé<sub>N</sub> – tel que sa blessure porte BEAUCOUP atteinte à sa santé', où l'apport de GRAND est juste ('beaucoup'), le reste – '[sa blessure] porte atteinte à sa santé' – faisant partie de la définition de BLESSURE, qui, à son tour, fait partie de la définition de BLESSÉ<sub>N</sub> :

BLESSÉ<sub>N</sub> = 'personne vivante qui a une ou plusieurs blessures'<sup>4</sup>.

L'adjectif GRAND est donc ici un vrai intensificateur, de sorte que la définition courante de BLESSURE – 'lésion visible faite aux tissus d'un être vivant par une cause extérieure' – doit être modifiée. Le principe de cohérence exige qu'on y ajoute la composante susceptible d'intensification : '... et qui porte ATTEINTE à la santé de l'être'. La définition finale de BLESSURE se lit donc :

BLESSURE = 'lésion visible faite aux tissus de l'être vivant X par une cause extérieure Y et qui porte atteinte à la santé de X'.

Ainsi la composante intensifiable ('porter atteinte') permettra de rendre compte, de façon naturelle et systématique, des cooccurrents comme *blessure grave* (*sérieuse, mortelle, légère, ...*). Par ricochet, la même composante justifie la

4 La composante 'vivant' est nécessaire car un cadavre couvert de blessures ne peut être appelé un *blessé*. Pour la condition de nécessité, voir Ch. III, 3.3.1.3, p. 92.



cooccurrence de BLESSÉ (adjectif et nom) : *blessé grave* (\**lourd*) vs *blessé léger, grièvement blessé* (\**malade*) vs *gravement malade*, etc.

Dans l'article de BLESSÉ<sub>N</sub>, l'adjectif GRAND apparaîtra comme une valeur de la *fonction lexicale standard simple Magn* (voir Ch. III, 3.5.2, p. 136) :

Magn : grand | prépos

Comme on le voit, le traitement simultané et comparatif des données sémantiques, syntaxiques et cooccurentielles, qui est imposé par la recherche de la cohérence absolue, aboutit à des résultats concernant les trois types de données. Parfois il faut changer la définition de départ, en l'adaptant à la syntaxe et à la cooccurrence observées; tel est le cas pour le nom BLESSURE dans notre exemple. D'autres fois, on décide de la description des cooccurents particuliers par des fonctions lexicales particulières; il en est ainsi pour VIEUX et ENDURCI (dans l'article CÉLIBATAIRE). Bien entendu, de nombreux autres cas, et souvent beaucoup plus complexes, se présentent : par exemple, le problème de cohérence entre la définition de la lexie vedette et l'inventaire de ses actants syntaxiques. Ainsi, étant donné qu'on *écrit* nécessairement dans une langue, naturelle ou artificielle, le sémantème 'langue' doit apparaître dans la définition de ÉCRIRE; il s'ensuit nécessairement qu'une dénomination de langue qui modifie le verbe ÉCRIRE dans une phrase doit être considérée comme un actant sémantique (et par conséquent syntaxique). Normalement les actants montrent un comportement idiomatique, restreint, irrégulier ; cependant, une expression comme EN FRANÇAIS auprès du verbe ÉCRIRE n'a rien d'irrégulier – elle semble être utilisée tout à fait librement (*Il a présenté sa thèse en français; Il donne des ordres en français*; etc.). On doit donc se demander si, dans *écrire en français*, l'expression EN FRANÇAIS est vraiment un actant de ÉCRIRE. Pour répondre complètement à cette question, il faut élaborer les articles de dictionnaire pour ÉCRIRE et ses verbes apparentés (EXPOSER, PRÉSENTER, DÉCRIRE, ...) ainsi que pour l'expression EN X (où X est un nom de langue) et ses parents sémantiques (par exemple, DANS LA LANGUE X : *Il a décrit tout cela dans une langue inconnue*). Dans l'état actuel des choses, nous devons rester fidèles à nos principes, c'est-à-dire accorder la primauté absolue à la logique et supposer que le nom de la langue est un actant du verbe ÉCRIRE. La recherche lexicologique approfondie peut invalider cette prise de position en démontrant ses conséquences indésirables. Mais alors il faudra changer les principes logiques, en les affinant, en les munissant de conditions plus restrictives, etc. En fait, il n'y a rien de particulier ici : c'est le cheminement normal de toute recherche scientifique.

Pour résumer succinctement le contenu du Principe de cohérence interne, nous pouvons retenir l'affirmation ci-dessous :

[un article de dictionnaire de L] *cohérent* = tel que le sémantisme, la syntaxe et la cooccurrence lexicale de L y sont bien mis en correspondance.

### 1.3 Principe de traitement uniforme

En généralisant l'exigence de cohérence à la macrostructure de dictionnaire, nous pouvons dégager le Principe de traitement uniforme.

|| Dans un DEC, les descriptions des lexies sémantiquement liées sont réalisées strictement de la même façon.

Exprimé différemment, ce principe affirme que les articles de dictionnaire des lexies d'un même *champ sémantique* (Ch. IV, 2.1, p. 173) doivent montrer un accord aussi complet que possible. Ainsi, l'article de dictionnaire de MAGAZINE doit correspondre à ceux de REVUE, de JOURNAL, de GAZETTE, etc.

Ce principe ne semble jamais observé dans les dictionnaires traditionnels. Citons, à titre d'exemple, le traitement des noms ethniques. Dans un dictionnaire courant, [un] FRANÇAIS est défini comme 'personne de nationalité française' (ce qui est inexact : *un Français* est quand même une 'personne de nationalité française DE SEXE MASCULIN'); [un] CHINOIS n'a pas de définition, mais est considéré comme un sens à part (2°, marqué Subst. : *Un Chinois*); [un] ALLEMAND n'a ni définition ni sens à part (pourtant, il y a la marque Subst. : *Les Allemands* [notez le pluriel !]; [un] RUSSE n'a pas de définition non plus, mais est considéré comme un sens à part (2°, marqué tout simplement « N » : *Un Russe*); [un] ESPAGNOL est traité comme [un] RUSSE, avec, en plus, l'indication du féminin possible; etc. Il semble, enfin, qu'on ne trouve pas deux noms ethniques ayant une structure de description identique. Or, dans le DEC une telle présentation n'est pas pensable : tous les noms ethniques doivent être décrits de la même façon. Plus concrètement, nous pensons qu'ils doivent tous être décrits au pluriel<sup>5</sup> et par rapport à un lieu géographique et une langue maternelle spécifiques, ce qui donnerait :

les FRANÇAIS = ethnie originaire de la France dont la langue maternelle est le français;  
 les CHINOIS = ethnie originaire de la Chine dont la langue maternelle est le chinois;

5 Les caractéristiques ethniques de par leur nature ne sont attribuables qu'à des collectivités. Les autres cas où la lexie définie doit être au pluriel incluent les noms de dispositifs et d'organes constitués de deux entités, tels GANTS, SKIS, YEUX, BRAS, ... (cf. p. 111, item 18). Les noms de ce type sont, dans un certain sens, des *pluralia tantum*. Le singulier morphologique correspondant signifie 'un représentant de l'ethnie en question'; ce singulier peut être pluralisé à son tour, en donnant le sens 'plusieurs représentants de l'ethnie en question'. La forme plurielle est donc ambiguë : elle signifie soit 'ethnie ...', soit 'plusieurs représentants de l'ethnie ...'.

NB : La corrélation sémantique observée entre les formes du pluriel et du singulier dans les lexies comme [les] FRANÇAIS ou [les] GANTS est très régulière et productive; le DEC devrait avoir une règle exprimant, de façon explicite, la généralisation couvrant ce fait. Ce type de règle constitue ce que Apresjan a nommé, il y a 30 ans, la *grammaire du dictionnaire*. Cependant, malgré l'importance de ce concept pour le DEC, il ne nous est pas possible de le développer davantage ici.

- les ALLEMANDS = ethnie originaire de l'Allemagne dont la langue maternelle est l'allemand;
- les RUSSES = ethnie originaire de la Russie dont la langue maternelle est le russe;
- les BRETONS = ethnie originaire de la Bretagne dont la langue maternelle est le breton; etc.

Cette présentation n'est qu'une ébauche rudimentaire : 1) il reste à définir les termes 'ethnie' et 'langue maternelle'<sup>6</sup>; 2) elle ne couvre pas le cas d'États multinationaux (*les Belges, les Suisses, les Canadiens, les Congolais, ...*); 3) elle ne couvre pas non plus le cas des ethnies pour lesquelles il est impossible de spécifier un seul pays d'origine (tels *les Peuls, les Berbères, ...*); 4) elle laisse de côté certaines populations pour lesquelles il est difficile de déterminer l'ethnie (*les Alsaciens, les Corses, ...*); 5) elle ne prévoit pas les cas non prototypiques : un Français né en Chine et qui ne parle pas français; etc. Nous voulions simplement mettre en relief l'UNIFORMITÉ des descriptions lexicographiques des lexies apparentées exigée dans le DEC.

**NB** : Notre exigence d'uniformité de description lexicographique ne doit pas être interprétée comme une obligation absolue de décrire TOUTES les lexies d'un même champ sémantique strictement de la même façon. C'est impossible, comme le montre déjà notre exemple ci-dessus : ainsi, le nom SUISSES ne peut pas être défini comme \*'ethnie originaire de la Suisse...'; la bonne définition dans ce cas est [*un*] SUISSE = 'originaire de la Suisse'. Dans le même ordre d'idées, le nom BERBÈRES doit être défini comme 'ethnie dont la langue maternelle est le berbère' (sans spécification du pays d'origine), etc. Le Principe d'uniformité n'exige l'uniformité que là où elle est possible et justifiée par la réalité linguistique.

Il est évident que seule l'approche par champ sémantique peut garantir la validité absolue du Principe de traitement uniforme. Alors que les dictionnaires sont habituellement rédigés par ordre alphabétique, c'est le concept de champ sémantique qui devient central dans l'élaboration du DEC. Par exemple, tous les noms ethniques sont décrits d'après le même canevas, ainsi que tous les noms de moyens de transport, tous les noms d'émotions, toutes les désignations pour les permissions / interdictions, pour les mesures (LARGE, HAUT, CHAUD, ...), pour les croyances, etc. L'importance du champ sémantique dans le DEC se dégagera mieux plus loin (Ch. IV, 2.1, p. 173), après l'explicitation des notions pertinentes.

6 Pour en montrer la faisabilité, esquissons ici une première définition d'ETHNIE :  
*ethnie* = ensemble d'individus possédant des caractéristiques communes linguistiques, culturelles, sociales et psychologiques dues, dans la plupart des cas, à une origine commune.

Même si dans le produit final – dans le DEC publié – on utilise la disposition alphabétique des entrées, il ne s'agit que d'une commodité de consultation; le DEC est un dictionnaire où le champ sémantique domine.

Nous pouvons donc dire :

TRAITEMENT UNIFORME = DESCRIPTION SIMILAIRE POUR TOUTES LES UNITÉS D'UN MÊME CHAMP SÉMANTIQUE.

L'application du Principe de traitement uniforme entraîne comme corollaire l'obligation de déterminer, pour chaque champ sémantique, le *schéma généralisé* ou format standard des descriptions lexicographiques des lexies du champ concerné. Dans le cas des ethnies, considéré ci-dessus, un tel schéma peut se présenter de la façon suivante :

*les*  $L_{(ethnique)}-s$  = ethnie originaire de ... dont la langue maternelle est ...

L'usage de tels schémas, qui ressemblent à ce que les écoliers connaissent depuis toujours comme « exercices à trous », ne se limite pas uniquement à la définition lexicographique. Bien au contraire, le même type de schémas généralisés est valable pour les autres parties d'un article de dictionnaire. En fait, le DEC a besoin de schémas généralisés d'un niveau plus élevé – pour le traitement uniforme des vocables apparentés d'un même champ lexical; mais pour le moment, nous ne sommes pas en mesure de développer davantage cette question importante.

#### 1.4 Principe d'exhaustivité

Comme tous les dictionnaires généraux de langue, le DEC a pour ambition de décrire le lexique d'une langue le plus complètement possible, c'est-à-dire de traiter TOUTES les lexies de la langue inventoriées à un moment donné. Le but ainsi visé pourrait être qualifié d'exhaustivité « externe ». C'est bien entendu le problème de la nomenclature du dictionnaire, qui sera repris ci-dessous, en 2. Cependant, le DEC se différencie des autres dictionnaires par son exigence d'une exhaustivité « interne », c'est-à-dire dans le traitement de CHAQUE lexie. (Ce principe concerne la microstructure du dictionnaire.)

|| Dans un DEC, l'article d'une lexie L inclut tous les renseignements lexicographiques nécessaires, d'une part, pour bien utiliser L elle-même, et d'autre part, pour pouvoir trouver les autres lexies  $L_1, L_2, \dots$  sémantiquement liées à L.

Ainsi, pour la lexie TRANSPORT EN COMMUN, il faut inclure les noms des divers moyens de transport et des lexies comme TITRE DE TRANSPORT, VOYAGEUR, CORRESPONDANCE, RÉSEAU, DESSERVIR, TRANSPORTER, VOYAGER, COMPOSTER, ... et même RESQUILLER ! Pour VACHE, il faut spécifier MEUGLER et MEUH; pour CANARD, CANCANER et COIN-COIN. De la même façon, on indique, sous la lexie correspondante, les interjec-

tions pertinentes : pour DÉGOÛT – POUAH ! ou BEURK !; pour PLAISIR – AH !; pour DOULEUR – AÏË !, AÏË-AÏË-AÏË !, OUILLE !; etc.

De plus, pour les interjections, il faut préciser les prosodies, puisque AH ! de plaisir et AH ! d'étonnement ne se prononcent pas de la même façon. On trouvera plus loin les détails de ce qui constitue les informations lexicographiques indispensables; il ne s'agit ici que d'insister sur la nécessité de les fournir complètement. Même au prix d'articles de dictionnaire très longs et complexes, le DEC doit renseigner l'utilisateur et lui fournir tout le matériel linguistique dont ce dernier pourrait avoir besoin.

De cette façon,

EXHAUSTIVITÉ (INTERNE) = PRÉSENTATION DE LA TOTALITÉ DES INFORMATIONS LEXICOGRAPHIQUES POUR CHAQUE LEXIE, Y COMPRIS LES RÉFÉRENCES À TOUTES LES LEXIES APPARENTÉES (faites de façon directe ou à travers d'autres lexies).

## 2 Nomenclature du DEC

La détermination de la nomenclature du dictionnaire proposé (= du DEC) s'appuie sur les deux principes suivants :

- le principe de disponibilité (2.1);
- le principe de lexicalité (2.2).

### 2.1 Principe de disponibilité

|| La nomenclature d'un DEC ne comprend que le lexique disponible, c'est-à-dire appartenant à la langue GÉNÉRALE STANDARD CONTEMPORAINE.

Nous appelons *disponible* le lexique dont dispose un locuteur « moyen », c'est-à-dire idéalisé – en fait, créé par la réflexion du lexicologue. Pour préciser quelque peu ce concept important, disons que ce locuteur « moyen » se situe à l'intersection des trois axes déterminant une langue naturelle : 1) axe socio-culturel ou sociolectal, 2) axe géographique ou topolectal et 3) axe chronologique ou chronolectal.

Selon l'*axe sociolectal*, on distingue les variétés linguistiques des classes et des groupes sociaux différents : langue des paysans, langue des ouvriers, langue des aristocrates, langue des médias, langue des voleurs et des criminels, langues de spécialités... Il est clair qu'un dictionnaire ne peut tenir compte de la totalité de ces multiples variétés et doit obligatoirement faire un choix justifié. Ainsi le DEC a ciblé la langue GÉNÉRALE, qu'on définit traditionnellement comme l'usage linguistique des gens ayant atteint un certain niveau de scolarité (niveau de formation universitaire). Notre locuteur « moyen » apparaît donc comme instituteur/professeur, journaliste, avocat, médecin ou autre représentant d'une profession intellectuelle.

Selon l'*axe topolectal*, on distingue les variétés linguistiques des régions différentes – la langue standard opposée aux dialectes et parlers régionaux. Pour des langues parlées dans plusieurs pays, telles que le français, l'anglais, l'allemand, etc., on doit distinguer de plus la variété du pays « central » (France, Angleterre, Allemagne) et les autres variétés nationales (le français suisse, belge, canadien, ...; l'anglais américain, australien, indien, ...; l'allemand suisse, autrichien, ...). Ici également, un choix s'impose. Le DEC se doit de préciser de façon rigoureuse la variété de langue retenue pour la description; cela peut bien être le parler standard littéraire ou un parler local, pourvu que la variété choisie soit systématique et que toutes les indications nécessaires soient présentes. Dans ce manuel, nous ne parlons que d'un DEC du parler STANDARD du pays « central »; dans notre cas, il s'agit du français hexagonal – du franco-français.

Selon l'*axe chronolectal*, on distingue les variétés linguistiques des périodes différentes – la langue de nos jours opposée à la langue des époques révolues. Même s'il est impossible de tracer avec précision les limites de « notre époque », ici encore une fois le lexicologue doit prendre une décision nette. Par exemple, bien que les textes de Molière appartiennent au français moderne, un dictionnaire de français ne peut recenser globalement tout leur vocabulaire tel quel (par exemple, il n'inclut pas SOINS au sens d'efforts, comme dans *Tous nos soins ont été inutiles* [« Dom Juan », III : 3, 1039], etc.). Le DEC établit, peut-être arbitrairement, les frontières de notre époque par l'âge du locuteur idéal : il a environ 50 ans, et nous postulons une continuité de la langue entre lui et ses parents (qui ont environ 25 ans de plus) ainsi qu'entre lui et ses enfants (qui ont environ 25 ans de moins). Par conséquent, dans le cadre du DEC, la langue contemporaine couvre la période des 50 dernières années.

Les professionnalismes /jargonismes, les régionalismes et les archaïsmes qui ne font pas partie de la langue générale standard contemporaine ne seront donc pas décrits dans le DEC; par cela, le DEC s'oppose nettement à un dictionnaire du type trésor, qui, lui, cherche à intégrer TOUTES les lexies attestées au cours des siècles, même les *hapax*, etc. (par exemple, dans la littérature). Cependant, nous ne voulons pas, par là, exclure les lexies professionnelles, régionales et archaïques si elles sont connues et utilisées par le locuteur moyen. Ainsi l'archaïsme AÉROPLANE ou le professionnalisme CORYZA seront listés parce qu'ils appartiennent au lexique commun des francophones. Notre façon de faire est ici identique à celle de beaucoup de dictionnaires, mais nous voulons être plus sélectifs et donc exclure toutes les lexies qui ne respectent pas le principe indiqué. Par exemple, la lexie BRISCARD sera exclue parce qu'elle nous semble ne pas appartenir au lexique disponible du locuteur contemporain (même si certains la reconnaissent). De telles lexies doivent être enregistrées dans un trésor.

Dans la démarche du DEC, la nomenclature ne peut être déterminée qu'en s'appuyant sur les jugements d'un groupe représentatif de locuteurs natifs. On cesse ainsi d'être esclave de l'attestation littéraire. D'une part, le lexicologue du DEC tient compte, en plus des romans et des essais, des écrits journalistiques, du

discours politique, des textes de vulgarisation scientifique et technique, ... – et même des enregistrements oraux; tout cela, sous forme de bases de données textuelles (voir plus loin, Ch. II, 3.2, p. 50). Mais d'autre part, pour le DEC, de telles bases de données ne sont qu'une source primaire d'information : elles fournissent la matière brute au chapitre des acceptions, de la combinatoire, etc., mais même l'attestation multiple d'un phénomène lexical ne suffit pas pour qu'il soit répertorié dans le DEC. Tout « matériel lexicologique » est passé au crible par les lexicologues; de cette façon, la décision finale – inclure une lexie dans le DEC ou l'exclure – reste celle des auteurs du DEC.

## 2.2 Principe de lexicalité

Les dictionnaires courants sont, par tradition, des dictionnaires de MOTS. Dans des tels dictionnaires, on voit, à l'intérieur des articles des mots individuels, ce qu'on appelle des locutions. Par exemple, sous ÂNE1 ('animal domestique...'), le PR donne *têtu comme un âne*, sans aucune explication, et *le coup de pied de l'âne*, avec une explication détaillée (et même l'« étymologie » : on indique que l'expression vient de la fable connue, etc.). Le DEC, qui ne procède pas de cette manière, se fixe pour but de décrire, de la même façon, les lexies de la langue quelle que soit leur nature. Une lexie est une expression qui ne peut être décrite complètement par la grammaire, c'est-à-dire par des règles générales, mais qui nécessite une description individuelle, donc une description par un article de dictionnaire. Comme on l'a déjà dit (section 1.1 du chapitre I, p. 16), une lexie peut être mono- ou multi-lexémique, et elle est donc soit un lexème soit un phrasème.

Un DEC se doit de décrire non seulement les lexèmes de la langue  $\mathcal{L}$  mais encore tous ses phrasèmes (dont le nombre est nettement supérieur à celui des lexèmes), en devenant de ce fait un dictionnaire de MOTS et de PHRASÈMES, donc un dictionnaire de LEXIES.

En suivant ce principe, l'exemple mentionné ci-dessus doit être réanalysé. Tout d'abord, *comme un âne* dans *têtu comme un âne* n'a pas le même sens que dans *comme un âne* tout court. Soit les phrases :

- (6) *Comme un âne, il a dit non.*  
*Il s'est comporté comme un âne.*

En (6), *comme un âne* veut dire 'stupidement' = ('comme un imbécile') et correspond tout à fait au lexème ÂNE2 ('personne stupide...') (pour une discussion plus détaillée des lexèmes ÂNE, voir Ch. IV, 1.2, p. 160). Nous croyons que c'est une comparaison libre, qui n'a aucune raison de figurer dans un dictionnaire. Par contre, dans *têtu comme un âne*, la même expression ne signifie pas 'stupidement' mais intensifie le sens de 'têtu'; à preuve, l'exemple suivant :

- (7) *Paul est très intelligent mais têtu comme un âne.*

L'absence de contradiction démontre que, dans ce contexte, *comme un âne* n'implique pas la stupidité. N'ayant pas de rapport sémantique avec ÂNE2, cette

expression doit apparaître dans l'article de TÊTU (comme la valeur de la fonction lexicale **Magn**).

Quant à l'expression *le coup de pied de l'âne* ('dernière attaque lâche contre un adversaire abattu'), c'est un phrasème complet qui constitue une entrée autonome dans le DEC. Le DEC traitera de la même façon les expressions *passer (sauter) du coq à l'âne* ('changer de sujet de façon brusque et illogique'), qui sont des phrasèmes complets, aux entrées autonomes. Cependant, la locution *l'âne de Buridan* ('l'âne1 qui est mort de faim entre deux bottes de foin parce qu'il ne pouvait se décider laquelle manger en premier') fera partie de l'article de dictionnaire de ÂNE1, puisque son sens inclut le sens de 'âne1'.

Comme on le voit, le DEC ne décrit pas tous les phrasèmes de la même façon. Les **phrasèmes complets** (= les expressions idiomatiques du type 'ARRIVER COMME UN CHEVEU SUR LA SOUPE') et les **quasi-phrasèmes** (= les locutions du type 'DONNER LE SEIN [à N]'), c'est-à-dire les phrasèmes qui ne peuvent pas être complètement décrits en fonction d'au moins un de leurs constituants, forment des entrées indépendantes – tout comme les lexèmes<sup>7</sup>. Les **semi-phrasèmes** (= les **collocations** du type ACCEPTER/DÉCLINER UNE INVITATION ou VIF INTÉRÊT) sont décrits sous l'entrée d'un de leurs constituants – par ce qu'on appelle les fonctions lexicales, voir plus loin, p. 125.

Par conséquent, la nomenclature du DEC comprend, en plus des lexèmes, tous les phrasèmes complets et tous les quasi-phrasèmes. Par exemple, l'expression RAT DE CAVE ('≈ mince bougie') constitue un phrasème complet indubitable avec une définition et une combinatoire particulières. Les dictionnaires courants

7 Le concept de phrasème et la distinction entre les trois types majeurs de phrasèmes sont centraux pour le présent ouvrage ; il est malheureusement impossible d'en parler en détail ici. Nous nous contenterons donc d'esquisser les définitions approximatives indispensables pour la compréhension de notre exposé.

#### Phrasème complet

Nous appelons *phrasème complet* un phrasème AB au signifié 'C' tel que ce signifié n'inclut ni le signifié de A ni celui de B : 'C' ⊃ 'A' & 'C' ⊃ 'B'.

Exemple : le signifié de 'PONT AUX ÂNES' ('banalité connue de tous') n'inclut ni celui de PONT, ni celui d'ÂNE ; c'est un phrasème complet (angl. *idiom*).

#### Semi-phrasème

Nous appelons *semi-phrasème* un phrasème AB au signifié 'AC' ou 'BC' qui inclut le signifié de l'un des constituants, alors que l'autre soit ne garde pas son sens, soit – même s'il garde son sens – n'est pas sélectionné librement.

Exemple : le signifié de PIQUER UN SOMME ('faire un somme') inclut le sens de SOMME ('sommeil court et léger'), mais pas celui de PIQUER ('entamer avec une pointe') ; c'est un semi-phrasème ou une collocation (avec le mot-clé [*un*] SOMME).

#### Quasi-phrasème

Nous appelons *quasi-phrasème* un phrasème AB au signifié 'ABC' qui inclut les signifiés des deux constituants et un surplus imprévisible 'C'.

Exemple : le signifié de CENTRE COMMERCIAL ('centre commercial formé de nombreux magasins et de lieux de services, ayant un parc de stationnement...') inclut les sens de CENTRE ('lieu où diverses activités sont groupées') et celui de COMMERCIAL ('relatif au commerce') plus la composante ('formé de nombreux magasins et de lieux de services...') ; c'est un quasi-phrasème.



le reconnaissent, mais au lieu de refléter explicitement cette autonomie du phrasème, ils rangent RAT DE CAVE dans l'article CAVE. Nous ne pouvons accepter cette façon de faire pour les trois raisons suivantes :

1. Cela entraîne l'introduction d'un article de dictionnaire complet à l'intérieur d'un autre article de dictionnaire, ce qui est repréhensible du point de vue de la logique.
2. Cela pose le problème du « mot d'accueil » : faut-il ranger RAT DE CAVE vraiment sous CAVE ou peut-être sous RAT ? Nous ne pouvons trouver de raisons valables pour décider dans un sens ou dans un autre, pas plus que les auteurs de certains dictionnaires, qui rangent RAT D'HÔTEL sous RAT, mais RAT DE BIBLIOTHÈQUE, sous BIBLIOTHÈQUE.
3. La recherche des phrasèmes dans le dictionnaire imprimé devient ainsi plus difficile.

Pour pallier ces manques, le DEC présente les phrasèmes complets et les quasi-phrasèmes comme des entrées indépendantes et munit chacun d'un article de dictionnaire particulier. Chaque phrasème de ce type est signalé graphiquement par des coins supérieurs :  $\lceil L_1 + L_2 + \dots + L_n \rceil$ . Ainsi, le DEC a, en tant que vedettes autonomes, les phrasèmes complets  $\lceil \text{RAT DE BIBLIOTHÈQUE} \rceil =$  'personne qui passe tout son temps à compiler des livres',  $\lceil \text{RAT DE CAVE} \rceil =$  'longue bougie fine' et  $\lceil \text{RAT D'HÔTEL} \rceil =$  'voleur qui se spécialise dans le vol des biens des clients d'un hôtel dans leur chambre d'hôtel'; ces phrasèmes seront disposés par ordre alphabétique<sup>8</sup>. Suivant le même principe, le DEC a comme entrées  $\lceil \text{BIEN QUE} \rceil$  (conjonction concessive),  $\lceil \text{TABLE DE NUIT} \rceil$  (qui n'est pas une *table*),  $\lceil \text{POT DE CHAMBRE} \rceil$  et  $\lceil \text{VASE DE NUIT} \rceil$ ,  $\lceil \text{DÉCOUVRIR LE POT AUX ROSES} \rceil$ ,  $\lceil \text{COUCHER EN JOUE} \rceil$ ,  $\lceil \text{BRISER LA GLACE} \rceil$ , etc. Même

8 Comme le DEC accorde une même importance au lexème et au phrasème, les mettant côte à côte dans la même liste et la même hiérarchie et les traitant lexicographiquement de la même façon, ce qui diffère fondamentalement de la pratique usuelle, nous croyons important de compléter notre discussion par d'autres exemples des locutions figées et semi-figées comprenant le lexème RAT.

Ainsi,  $\lceil \text{RAT DE CALE} \rceil$  'passager illégal',  $\lceil \text{RAT D'OPÉRA} \rceil$  'jeune danseuse dans un théâtre' et  $\lceil \text{RAT D'ÉGLISE} \rceil$  'homme bigot de religion chrétienne' sont également des phrasèmes complets et auront des entrées et des articles de dictionnaire indépendants dans le DEC. En même temps, les renvois sémantiques seront établis comme suit (sous forme de fonctions lexicales non standard) :

pour  $\lceil \text{RAT DE CALE} \rceil$  – sous PASSAGER ;

pour  $\lceil \text{RAT D'OPÉRA} \rceil$  – sous THÉÂTRE et DANSEUSE ;

pour  $\lceil \text{RAT D'ÉGLISE} \rceil$  – sous RELIGION, BIGOT,  $\lceil \text{GRENOUILLE DE BÉNITIER} \rceil$  et  $\lceil \text{PUNAISE DE SACRISTIE} \rceil$ .

En plus, les renvois de REPÉRAGE pour ces phrasèmes seront donnés, comme on s'y attend, sous les lexèmes pleins qui les constituent, c'est-à-dire sous RAT, CALE, OPÉRA, ... (dans la zone phraséologique).

Quant aux expressions *rat d'égout*, *rat de champs*, *rat d'eau*, *rat gris* et *rat noir*, elles désignent toutes des variétés de rats. Pour cette raison, ces expressions n'auront pas d'articles particuliers, mais seront simplement listées sous RAT, toujours à l'aide de fonctions lexicales non standard. À ce propos, il est important de souligner que les expressions du type *rat de blé* ou *rat musqué* [= ondatra], qui ne désignent pas des variétés de rats mais des muridés semblables, seront décrites comme les phrasèmes complets :

$\lceil \text{RAT DE BLÉ} \rceil$  et  $\lceil \text{RAT MUSQUÉ} \rceil$ .

les proverbes, tels que «SITÔT DIT, SITÔT FAIT», «L'ARGENT NE FAIT PAS LE BONHEUR», «TEL PÈRE, TEL FILS», «APPORTER DE L'EAU À [A<sub>poss</sub> = *mon, ton, son, ...*] MOULIN», ..., sont des entrées individuelles; chacun est également décrit par son propre article de dictionnaire.

Pour faciliter l'accès diversifié et complet du lecteur aux phrasèmes, le DEC a recours à des renvois sémantiques et formels. D'une part, pour l'encodage, c'est-à-dire la production de phrases (= perspective onomasiologique), le DEC donne les liens sémantiques entre toutes les lexies, y compris les phrasèmes. Ainsi, par exemple, on trouve dans les articles de FUSIL, de VISER et de TIRER «faire feu» un renvoi à «COUCHER EN JOUE»; sous SECRET, on a un renvoi à «DÉCOUVRIR LE POT AUX ROSES»; et «APPORTER DE L'EAU À [A<sub>poss</sub>] MOULIN» est référencié sous ARGUMENT. D'autre part, pour le décodage, c'est-à-dire la compréhension de phrases (= perspective sémasiologique), le DEC liste les renvois à un phrasème sous tous ses constituants (sauf, bien entendu, les mots outils) – dans la zone phraséologique de leur article de dictionnaire (Ch. III, 3.7, p. 153). Ainsi, «COUCHER EN JOUE» est référencié sous COUCHER et sous JOUE, «DÉCOUVRIR LE POT AUX ROSES», sous DÉCOUVRIR, sous POT et sous ROSE, et ainsi de suite. (Le DEC peut procéder ainsi parce que chaque phrasème a son article de dictionnaire complet et les indications dans la zone phraséologique des constituants de ce phrasème ne sont que de simples renvois.)

Pour conclure, la nomenclature du DEC se différencie donc des nomenclatures des dictionnaires courants sous deux aspects :

- elle se limite exclusivement au lexique général standard contemporain;
- elle inclut, à côté de lexèmes, des phrasèmes – toutes ces lexies étant traitées de la même façon.

C'est particulièrement ce deuxième aspect qui fait du DEC un dictionnaire de phraséologismes. En effet, non seulement la nomenclature inclut tous les phrasèmes qui doivent être munis d'articles de dictionnaires indépendants, mais aussi tous les articles comprennent tous les cooccurrents, les collocations, les clichés, etc. qui se rapportent à la lexie vedette. Ceci reflète notre conviction que les locuteurs parlent par phrasèmes, que les textes ne sont constitués que de phrasèmes de toutes sortes et que la séparation entre les dictionnaires de mots et les dictionnaires de phraséologismes n'a pas sa raison d'être.

### 3 Documentation du DEC

Deux principes guident le lexicologue du DEC dans la constitution de la documentation linguistique, son point de départ :

- principe de la primauté du locuteur (3.1);
- principe de consultation obligatoire des bases de données textuelles (3.2).

### 3.1 Principe de la primauté du locuteur

Tout dictionnaire est une description, par des lexicologues, d'un lexique – c'est-à-dire d'une réalité OBJECTIVE, existant indépendamment des lexicologues. Cette réalité, pour être proprement saisie et représentée, doit être, avant tout, parfaitement documentée; l'on comprend que la documentation du lexique visé revêt une importance capitale. Il en est ainsi pour tous les dictionnaires, *a fortiori* le DEC.

Pour les dictionnaires traditionnels, la seule source d'informations factuelles sur le lexique normalement acceptée est un fichier. Un tel fichier est un résultat de dépouillement de textes de diverses origines – majoritairement littéraires, avec l'inclusion sporadique d'exemples journalistiques, didactiques ou autres. Aujourd'hui, comme chacun le sait, le fichier manuel est de plus en plus remplacé par une base de données textuelles informatisée; mais rappelons encore une fois que, même si les bases de données ont des avantages marqués par rapport au fichier manuel (voir ci-dessous, 3.2), il n'en demeure pas moins qu'elles représentent la même matière brute que les fichiers : les textes. De plus, on connaît la pratique largement répandue de reprendre un nombre important de données lexicales des dictionnaires précédents, surtout si ces derniers jouissent d'un certain prestige.

Bref, nous pouvons dire que la documentation d'un dictionnaire traditionnel s'appuie surtout sur l'attestation OBJECTIVE : on ne consigne que ce qui est attesté, et on consigne (en exagérant un peu) tout ce qui est attesté, particulièrement si l'attestation est étayée par une autorité littéraire.

Cette façon de faire, qui est d'ailleurs tout à fait naturelle, n'est pas suffisante dans le cadre méthodologique du DEC. En effet, les données textuelles ne sont pas nécessairement fiables car on y trouve quantité d'*hapax legomena*, de figures de style d'auteur, de lapsus et même de simples erreurs, dont les meilleurs auteurs ne sont pas forcément exempts. Quant aux dictionnaires, ils semblent préserver des entrées avec des attestations littéraires depuis de nombreuses décennies; citons le cas de PIEDLB.1 (chez l'animal) dans le PR, illustré des exemples suivants : *Les bœufs creusaient la terre de leurs larges pieds fourchus* [Sand] et *Le grand héron... levant les pieds sur les cailloux du fond* [Genevoix]. Avec de tels exemples et la définition fournie – 'Extrémité inférieure de la jambe (des équidés), de la patte (de quelques mammifères et oiseaux)...' –, on peut facilement déduire que les vaches, les chiens et les cigognes ont des pieds<sup>9</sup> ! Pour éviter de tels abus, il faut absolument mettre à profit l'intuition linguistique du lexicologue, appuyé, comme nous l'avons déjà signalé, par le groupe représentatif de locuteurs natifs.

9 Les expressions du type 'PIED DE COCHON' ('patte de cochon préparé d'une certaine façon') ou 'PIED DE MOUTON' ('espèce de champignon') ne prouvent pas l'existence de pieds chez les animaux en question : ce sont des expressions idiomatiques – des phrasèmes complets. On ne peut pas dire \**Le cochon a un pied blessé* ou \**Ce mouton a des pieds solides*.

|| Dans le cadre d'un DEC, les décisions d'acceptabilité ou de refus des expressions de la langue  $\mathcal{L}$  ne reviennent qu'au locuteur natif de  $\mathcal{L}$ .

Par conséquent, dans notre approche, le locuteur idéalisé forme, avec l'auteur de l'article de dictionnaire, la source fondamentale des données linguistiques. Cela signifie en fait que l'intuition linguistique des locuteurs contemporains est à la base de la démarche lexicologique. Un DEC modélise donc la conscience linguistique des locuteurs d'aujourd'hui, et, dans ce sens, on peut dire qu'il s'appuie, de façon cruciale, sur l'attestation SUBJECTIVE.

Notons que cette attitude ne contredit pas le cadre général dans lequel s'inscrit la recherche linguistique contemporaine. L'approche rigoureusement distributionnaliste se limitant à l'analyse des textes semble être dépassée, et l'introspection, c'est-à-dire l'appel à sa propre conscience linguistique, est une méthode généralement acceptée. Bien entendu, on ne peut se fier de façon aveugle à cette façon de faire, et il faut que les intuitions du lexicologue soient corroborées par les données textuelles et vérifiées par les intuitions des autres locuteurs. Mais il reste néanmoins que le travail lexicologique est sous-tendu par les intuitions linguistiques des locuteurs.

La primauté du locuteur dans l'élaboration du DEC est d'autant plus indispensable que, à la différence des dictionnaires traditionnels, le DEC fournit des EXEMPLES NÉGATIFS, pratique très souvent nécessaire pour étayer la solution proposée. Par exemple, la décision d'avoir deux entrées pour CHANGER intransitif (Ch. III, 2.1, p. 58) doit être justifiée par des exemples du type *\*La mère a changé Colette de coiffure*<sup>10</sup>. De façon analogique, si on dit *la plus grande surprise de ce séjour* et *sa plus grande surprise*, on ne dit pas – même si quelqu'un a eu une grande surprise pendant son séjour quelque part – *\*sa plus grande surprise de ce séjour* [il faut dire *sa plus grande surprise pendant ce séjour* ou *la plus grande surprise de son séjour*, ...]; cet exemple négatif justifie la restriction qui accompagne le régime du lexème SURPRISE1.2 (voir le vol. III du DEC du français publié; dorénavant, nous notons ces volumes comme DEC-1, DEC-2 et DEC-3).

### 3.2 Principe de consultation obligatoire des bases de données textuelles

Bien entendu, le rôle primordial du locuteur dans le DEC n'exclut en rien l'utilisation des données textuelles, qui constituent une source importante de renseignements sur la langue. Mais, parmi les trois sources de documentation

<sup>10</sup> Pendant longtemps, la linguistique – et surtout la lexicologie et la lexicographie – se contentait d'illustrer les affirmations par des exemples positifs, c'est-à-dire par de «bonnes» expressions. Mais de tels exemples sont nettement insuffisants : très souvent, une règle linguistique interdit de «mauvaises» expressions et par conséquent elle appelle des exemples de telles expressions, c'est-à-dire des exemples négatifs. C'est l'école de grammaire générative (Chomsky) qui a révolutionné la linguistique en introduisant la technique des exemples précédés d'un astérisque. Notons en passant que l'utilisation d'exemples négatifs dans les dictionnaires a déjà été recommandée, il y a des décennies, par le lexicographe russe L. Ščerba (1940).

mentionnées ci-dessus (fichiers manuels, bases de données, dictionnaires existants), le DEC en privilégie une seule :

|| Dans le cadre d'un DEC, les données sur les faits pertinents de la langue  $\mathcal{L}$  doivent provenir des bases de données textuelles informatisées.

La raison de ce choix est que les bases de données textuelles possèdent, en comparaison avec les fichiers manuels, un très grand nombre d'avantages dont nous allons indiquer les trois qui sont les plus centraux :

- stockage énorme, presque illimité, de textes de tout genre, y compris des enregistrements oraux;
- représentativité globale et objective, garantie par la quantité même de données (n'importe quel fichier n'évite pas des choix subjectifs déterminés par les intérêts et les capacités d'un chercheur particulier);
- accès immédiat aux ensembles de faits linguistiques recherchés, qui permet de nombreuses comparaisons sous une forme conviviale et en peu de temps – possibilité extraordinaire, appelée à modifier complètement la nature du travail lexicographique.

Cependant, malgré tous les avantages des bases de données textuelles, le locuteur natif reste l'autorité finale sur l'admissibilité / la non-admissibilité de tel ou tel phénomène linguistique. La base de données qu'utilise le lexicologue n'est que son aide-mémoire, ou mieux – son aide-intuition. Dans le choix des unités à répertorier dans un dictionnaire, le lexicologue seul doit décider, en se fondant sur des considérations linguistiques et logiques; aucune base de données, même la plus riche, ne peut le remplacer. En lexicologie, c'est le lexicologue qui a toujours le dernier mot<sup>11</sup> !

Les bases de données textuelles informatisées prennent de plus en plus d'importance dans la recherche lexicologique / lexicographique. Ce processus s'inscrit tout à fait naturellement dans le cadre de la tendance générale contemporaine d'introduire l'automatisation et, surtout, l'informatisation dans tous les domaines de la vie sociale, industrielle et même privée : après la robotique et la bureautique, on entre dans la domotique !

Quant à la linguistique et, plus précisément, à la lexicologie, l'utilisation de machines n'est pas une idée toute neuve. Dès l'apparition des calculateurs à cartes perforées, les lexicologues ont vu l'avantage de la mécanisation. Au début des années 50, on s'est lancé dans des travaux de compilation mécanique de concordances (d'auteurs), de dictionnaires inverses et de listes de fréquence de mots utiles pour la recherche. Il n'est pas dans notre intention de fournir un historique systématique de cette période de *Sturm und Drang*; mais, comme notre thème central est surtout la lexicologie française, nous nous permettons quand même de

11 Kjellmer 1994 illustre un essai d'extraction automatique de collocations d'un corpus informatisé sans intervention décisionnelle d'un lexicologue; ce dictionnaire est rempli d'expressions comme *Mr. Smith, was a member, the abilities, a bad thing*, etc., qui n'ont rien à voir avec les collocations.

rappeler que, déjà en novembre 1957, le colloque « Lexicologie et lexicographie françaises et romanes » (à l'université de Strasbourg) a lancé les travaux de recherche en automatisé lexicographique appliqués au français. C'est ainsi que sont nés, grâce à l'impulsion de Bernard Quemada, *Le Trésor de la langue française*, l'Institut National de la Langue Française [= INaLF] et la revue *Cahiers de lexicologie*, dont le rôle en lexicologie moderne est primordial.

La multiplication des ordinateurs, ainsi que les recherches sur la traduction automatique et les autres aspects du traitement de textes ont apporté de nouveaux développements, notamment dans la constitution et l'utilisation de BASES DE DONNÉES TEXTUELLES INFORMATISÉES [= BDTI]. En effet, l'INaLF a mis sur ordinateur tous les grands textes de la littérature française des origines à nos jours, créant de cette façon une BDTI énorme, qui, depuis plus de vingt ans, sous-tend tous les travaux lexicographiques institutionnalisés en France. À notre connaissance, c'est la plus importante BDTI littéraire pour le français (des BDTI pour des textes techniques et scientifiques, pour la langue orale, etc. abondent).

Des travaux analogues se font dans beaucoup d'autres pays; citons pour mémoire deux BDTI pour l'anglais. Le corpus de l'Université Brown (Providence, RI; États-Unis), constitué de tranches de textes, d'un million de mots chacun, dans des domaines fort différents, a servi pour des calculs de fréquences et la préparation de dictionnaires de cooccurrents. La BDTI de l'Université de Birmingham (Royaume-Uni) couvre la langue orale et écrite (= textes littéraires, journalistiques, techniques...); elle a permis de produire le dictionnaire COBUILD, qui s'appuie entièrement sur le corpus de cette BDTI (Sinclair 1987).

Nous pouvons même dire que tous les dictionnaires du français et de l'anglais récemment parus ont été élaborés à partir de BDTI (par exemple, le dictionnaire anglais-français de Hachette-Oxford et *Longman Language Activator*). Le nombre de BDTI augmente partout au monde et alimente diverses recherches linguistiques.

Notre propre travail sur le DEC du français contemporain profite également de la consultation systématique d'une BDTI – TEXTUM (où ...UM = Université de Montréal). Cette base de données est constituée de textes ENTIERs – littéraires, journalistiques et techniques, dont la taille totale va jusqu'à 500 000 000 mots courants. TEXTUM est, bien entendu, muni d'un logiciel d'interrogation – le système PAT (développé à l'Université de Waterloo, Ontario, pour l'informatisation du *Oxford English Dictionary*)<sup>12</sup>.

12 À titre d'illustration, nous reproduisons ci-dessous une sortie d'interrogation pour les expressions ÂNE et ÂNES :

5859025, ..t ce temps, l' âne - exclu des négociations - n'était déjà que le  
17697554, ..nnées à dos d' âne, à cheval, en vélo tout terrain. Ce soir, à 17  
5905865, ..t un bonnet d' âne au dernier de la classe. Peut-être que lui au  
5858791, ..se dresse, un âne ça s'éduque, on réalise que l'âne et le chev  
5858340, ..s un bonnet d' âne... Car l'âne, selon le maître du lieu, ne  
27709026, ..DERNEAU. - Un âne chevauché d'un pantin à tête de cochon les pré

Le Chapitre II, comme on vient de le voir, est consacré entièrement à l'énonciation et à l'explicitation des huit principes centraux sous-jacents à tout travail lexicologique. Cependant, les principes exposés n'ont pas nécessairement la même importance ni ne sont logiquement indépendants les uns des autres. Pour bien faire, il serait essentiel d'en établir la hiérarchie et de dégager les liens logiques entre certains principes. Pour le moment, nous ne pouvons pas satisfaire ces exigences; nous sommes persuadés, néanmoins, que tous ces principes sont déjà utiles tels quels pour la recherche lexicologique.

(suite note 12)

27712195, ..duction d'un âne dans les locaux... Conseil général Le Finistère  
 22011518, ..n ? Le dos d' âne du passage à niveau des Forges raboté Finis le  
 5853499, ..boeuf et d'un âne et de l'enfant Jésus. Son village est devenu u  
 5537700, ..és camaïeu, l' âne et le bœuf, l'ange et puis notre saint Joseph  
 5858827, ..réalise que l' âne et le cheval s'opposent comme le jour et la nu  
 16461374, .. Sans doute l' âne. Je le trouve fin et intelligent. Votre boisson  
 5857926, ..n bonnet... d' âne. L'occasion est belle, pour l'heureux propriét  
 29394972, .. f, Peau d' âne, La conquête de l'Ouest. Ba  
 31092035, .. du coq à l' âne, Le jeu de l'Oie continue. Il faudra encore  
 5905666, ..e dit. Si l' âne n'avance pas quand son maître le lui demande,  
 5859131, ..la crèche : l' âne n'était qu'un âne peut-être, mais c'est lui qui  
 1602497, ..mémoires d'un âne, ou bien Le général Dourakine. Certains exégèt  
 30223428, ..t voir Peau d' âne ou Cendrillon avec les Marionnettes de Nantes  
 5857442, ..hevaline. L' âne peut vivre dans la cour de votre maison, il n'  
 5859149, ..n'était qu'un âne peut-être, mais c'est lui qui - de Nazareth à  
 10474864, ..ssi exhibé un âne portant les inscriptions FR3 : effet triste !  
 28337836, ..ne, mi-peau d' âne. Quelque part, il y a un secret qui ne serait  
 23961967, ..mais boire un âne qui n'a pas soif. La formation selon GARF I.  
 5858355, ..ne... Car l' âne, selon le maître du lieu, ne s'entête pas bête  
 3422275, ..x d'un étalon âne trop lourd pour elle qui était en train de la  
 5858310, ..Tu n'es qu'un âne... Tu mérites un bonnet d'âne... Car l'âne,  
 22011674, ..tin, le dos d' âne va sauter. Certes, sous les roues, on sentira  
 32585455, ..es, comme les ânes, les bœufs et les Rois mages en bois d'olivier  
 25020028, .. rose. Les ânes vagabondent en liberté, à moins qu'on n'ait b  
 5856642, .. des petits ânes. A Saint-Mars-la-Brière, Claude Hamelin met  
 9665987, .. acheté des ânes, Abélard et Cadichon, et aussi un manège. Quan  
 20129632, .. et plusieurs ânes avaient pris place à l'extérieur pour la gran  
 5856472, .. parler des ânes. Avec Noël, oubliés les ânes d'attelage, de m  
 10524649, .. le chant des ânes, coqs et autres chiens de village, le vol - e  
 5856501, .., oubliés les ânes d'attelage, de monte, de bât pour les randonn  
 5856613, ..ons, tous les ânes deviennent des petits ânes. A Saint-Mars-  
 5856742, .. plus petits ânes du monde. Bambou, l'âne-nain de Yougoslavie  
 5859292, ..s plus petits ânes... du monde chez Claude Hamelin, à Saint..  
 14014185, ..cale pour les ânes) faisaient encore partie du paysage. Et les .  
 5905616, .. de Paul, les ânes ne sont pas aussi bêtes qu'on le dit. Si l'  
 5905005, .. othèques. Les ânes ne sont pas si bêtes. Ses histoires ne sont p  
 5856349, .. e des petits ânes. Pas si bête qu'on ne les croit. C'est en déc  
 16458566, .. il aime les ânes, Romy Schneider et le Père Noël : un pur idéal  
 5857338, .. formel : les ânes sont effectivement, par leur comportement, plu  
 Le numéro qu'on voit au début de chaque ligne est la référence permettant de retrouver l'œuvre et le passage exact d'où la ligne en question est tirée. Ce numéro permet aussi un élargissement du contexte pour mieux saisir le sens de l'expression identifiée.





# CHAPITRE III

## MICROSTRUCTURE

### DU DICTIONNAIRE EXPLICATIF ET COMBINATOIRE

Comme nous l'avons déjà dit, la description du DEC est subdivisée en deux parties. La première partie s'intéresse à la MICROSTRUCTURE du DEC, c'est-à-dire à l'organisation d'un seul article de dictionnaire, organisation qui se répète de façon systématique pour tous les articles. La deuxième partie concerne la MACROSTRUCTURE du DEC, c'est-à-dire l'organisation de l'ensemble du dictionnaire à partir des articles qu'il comprend. Ce chapitre traite de la microstructure; la macrostructure sera considérée dans le chapitre suivant.

Dans le présent chapitre, nous examinons trois questions fondamentales :

- l'explicitation du CONCEPT DE LEXIE, unité de base du DEC (cf. plus haut, Ch. I, 1, p. 15), qui est décrite par un article de dictionnaire (1);
- la DÉLIMITATION des lexies (2);
- et la description d'une lexie dans le DEC, c'est-à-dire la STRUCTURE D'UN ARTICLE DE DICTIONNAIRE du DEC (3).

#### 1 Concept de lexie

Nous traiterons de la lexie en deux sous-sections : il sera d'abord question de l'interdiction de l'enchâssement d'articles de dictionnaire (1.1), puis de la formulation des définitions des concepts pertinents (1.2).

##### 1.1 *Éléментарité des articles de dictionnaire du DEC*

Pour bien clarifier les choses, rappelons que, pour les dictionnaires courants, l'unité de base est un mot pris dans toutes ses acceptions, donc un MOT POLYSÈME; ce mot est décrit par un seul article de dictionnaire pouvant avoir des subdivisions ultérieures. Les règles d'organisation utilisées dans de tels dictionnaires ne visent que les articles de ce type. Or cette approche pose des problèmes puisque des subdivisions indispensables constituent le plus souvent, à leur tour, des articles de dictionnaire complets. On trouve ainsi des articles à l'intérieur d'autres articles et ainsi de suite, donc des articles de dictionnaire gigognes. Ce qui est gênant dans cette façon de faire, c'est que cet enchâssement multiple se réalise plutôt « par accident », très souvent sans application de règles spécifiques pour l'organisation de chaque subdivision, c'est-à-dire, de chaque acception particulière. On connaît, il est vrai, des dictionnaires qui structurent la description des acceptions indivi-

duelles de façon systématique, en suivant des règles assez précises; mais dans de tels cas, ces dictionnaires utilisent *de facto* – en tant qu'unité de base – une acception individuelle et non un mot polysème, sans l'avoir déclaré explicitement.

Un DEC ne peut procéder ainsi : le respect des obligations formelles de rigueur logique oblige le lexicologue du DEC à rédiger des articles de dictionnaire ÉLÉMENTAIRES, c'est-à-dire n'incluant pas d'autres articles de dictionnaire. Le terme *élémentaire* est ici pris, bien entendu, dans son acception quasi mathématique, c'est-à-dire 'non constitué d'autres entités du même type' (semblable à l'acception qu'on voit, par exemple, dans *particules élémentaires*, etc.). Un article de dictionnaire du DEC peut être considéré à juste titre comme un « atome lexicographique ». L'élémentarité d'un article n'exclut en rien son caractère complet, voire exhaustif, ni son caractère fort complexe; la seule exigence est que le lexicologue ne glisse pas à l'intérieur d'un article de dictionnaire d'autres articles qui ne sont pas en situation régulière, surtout sans l'avoir « officiellement » annoncé.

Il en découle que l'unité lexicale décrite par un article de dictionnaire doit être élémentaire à son tour; une telle unité est bien entendu la *lexie*, dont il a déjà été question.

## 1.2 *Lexie, unité de base du DEC*

Plus précisément, comme un article du DEC est élémentaire, il couvre nécessairement une seule acception bien circonscrite d'un mot, soit un *lexème*, ou une seule acception bien cernée d'une expression figée, soit un *phrasème* (plus précisément, un phrasème complet ou un quasi-phrasème, voir note 7, p. 46). Un lexème ou un phrasème constitue une *lexie*, qui est donc l'UNITÉ DE BASE du DEC. La *lexie* est soumise au traitement lexicographique selon les trois aspects suivants : DÉLIMITATION, DESCRIPTION et REGROUPEMENT DES LEXIES. Nous allons discuter des deux premiers aspects tout de suite; quant au troisième, comme il appartient à la macrostructure du dictionnaire, nous le reverrons au Chapitre IV.

Pour mieux ancrer notre discussion ultérieure, donnons les définitions formelles des trois concepts énoncés ci-dessus.

### Définition III.1 : Lexème

|| Un *lexème* est un mot pris dans une seule acception bien déterminée et munie de tous les renseignements qui spécifient totalement son comportement dans un texte.

### Définition III.2 : Phrasème

|| Un *phrasème* est une locution prise dans une seule acception bien déterminée et munie de tous les renseignements qui spécifient totalement son comportement dans un texte.

### Définition III.3 : Lexie

|| Une *lexie* est soit un lexème, soit un phrasème.

## 2 Délimitation d'une lexie

Après avoir examiné les caractéristiques générales du problème de la délimitation d'une lexie (2.1), nous passons à la présentation des critères utilisés pour cette délimitation (2.2).

### 2.1 Généralités

Le lexicologue commence le travail de rédaction d'un article de dictionnaire par le choix intuitif d'une lexie hypothétique L. Plus spécifiquement, il identifie L par un ensemble d'occurrences contextualisées d'un mot ou d'une locution qui lui semblent toutes appartenir à la même lexie. En d'autres termes, il sélectionne un nombre de phrases qui comportent L et telles que L y affiche le même sens et le même comportement syntaxique. Par exemple, si le lexicologue veut considérer le verbe S'OCCUPER, il peut commencer sa réflexion par l'examen des exemples suivants :

- (1) a. *Sylviane s'occupe de vous comme il faut ?*
- b. *Votre foldingue de sœur s'occupera de vous.*
- c. *Elle ne s'occupe pas beaucoup de moi.*
- d. *Cette demoiselle Gertrude s'occupe des invités à sa place.*
- e. *On payait Paule pour qu'elle s'occupe des fichus moutards.*
- f. *Sylviane s'occupera du thé.*
- g. *Depuis trois ans, il s'occupe de la tribu des Mayalis.*

Supposons que le lexicologue en conclut que, dans tous les exemples (1), il y a une même lexie L, dont le sens est à peu près ('X prend soin de Y de la façon dont X doit prendre soin de Y dans la situation en question'). Cette première analyse reflète une intuition encore assez fragmentaire; pour la préciser et la raffiner, il faudra procéder à l'opération rigoureuse de la délimitation des lexies. Pour cela, il faut au préalable établir des critères d'analyse; nous allons le faire immédiatement, puis, ainsi équipés, nous reprendrons l'étude du verbe S'OCCUPER.

La DÉLIMITATION d'une lexie par rapport aux lexies apparentées ayant le même signifiant correspond en fait à la distinction des acceptions d'une expression polysème dans l'approche traditionnelle (cf., par exemple, van der Eijk *et al.* 1995), à la différence que, pour le DEC, la délimitation des lexies se fait, comme nous venons de le dire, sous le contrôle d'un JEU DE CRITÈRES rigoureux.

Comme l'utilisation systématique de divers critères est une caractéristique vitale de la lexicologie explicative et combinatoire, c'est le moment d'en préciser le rôle. (Il est clair que les critères ne sont pas indispensables seulement pour la délimitation des lexies, mais pour toutes les manipulations des matériaux lexicaux; nous en parlons ici tout simplement parce que c'est la première fois que nous en faisons mention.)

C'est exclusivement l'observance rigoureuse des critères préétablis qui garantit le caractère scientifique d'une description. Les critères lexicologiques qu'on verra par la suite facilitent, d'une part, le travail du chercheur, en rendant ce travail plus systématique et en le préservant de certaines erreurs. Dans ce sens, les critères proposés servent à la fois d'outil et de garde-fou. D'autre part, et de façon encore plus cruciale, ces critères permettent aux critiques de trouver des contre-preuves précises et valables applicables aux descriptions proposées; donc les critères lexicologiques font en sorte que les descriptions lexicales dans le DEC soient méthodiquement FALSIFIABLES. Comme on le sait très bien, la falsifiabilité des affirmations scientifiques est dans la science moderne une mesure du caractère scientifique de ces affirmations. Ceci explique pourquoi nous accordons une telle importance à nos critères.

Dans la perspective qui est la nôtre, il faut souligner, et avec vigueur, que les critères lexicologiques doivent être PUREMENT LINGUISTIQUES. Cela signifie, entre autres, que toutes les décisions concernant la délimitation des lexies ne doivent s'appuyer que sur des considérations strictement textuelles, à savoir sur la possibilité ou l'impossibilité de construire des expressions correctes comprenant la lexie étudiée dans telle ou telle construction particulière. Prenons l'exemple de CHANGER, verbe intransitif. On peut se demander si CHANGER dans *changer de chambre* (*d'école, de garnison, ...*) et CHANGER dans *changer de direction* (*de veste, de numéro de téléphone, ...*) doivent être décrits comme un seul lexème ou comme deux lexèmes différents. On pourrait, par exemple, dire que ce sont deux lexèmes puisque les actions 'changer de fauteuil' et 'changer de direction' sont différentes. Mais même si ce raisonnement est en soi tout à fait valable, on ne peut s'y fier, parce qu'on n'a aucun critère clair et stable qui permet de savoir si deux actions quelconques sont différentes ou non. Ainsi, les actions de manger de la soupe, un bifteck, des noix ou de la crème sont physiquement distinctes; pourtant, comme en français toutes les quatre sont désignées par le même lexème – MANGER, un francophone les considère comme une même action. (Cependant cela ne sera pas le cas pour une langue qui a des verbes particuliers pour 'manger de la soupe' ou 'manger des noix'.) Le lexicologue du DEC ne peut pas être satisfait de tels raisonnements : sa réflexion doit être limitée aux faits de la

langue. Ainsi, il dira que les deux CHANGER en question représentent des lexèmes différents puisque seulement le premier, mais pas le second, possède un correspondant transitif causatif. En effet, on dit *changer Jean de chambre* (d'école, de garnison, ...) mais aucunement \**changer Jean de direction* (de veste, de numéro de téléphone, ...). (Remarquons, en passant, que cette dernière impossibilité ne peut s'expliquer sémantiquement : le sens (causer que Jean change de direction (de veste)) ne soulève en soi aucun problème – on peut l'exprimer comme *le faire changer de direction* (de veste)<sup>1</sup>.) Le lexicologue rédigera ici obligatoirement deux articles de dictionnaire différents, l'un décrivant CHANGER dans son usage illustré par *Jean change de fauteuil* (de chambre, d'école, ...), et l'autre, CHANGER dans l'usage illustré par *Jean change de direction* (de veste, de numéro de téléphone, ...). Ce n'est que de cette manière que le DEC pourra fournir, de façon rigoureuse et cohérente, l'information indispensable à l'utilisateur du dictionnaire en vue d'un encodage : dans le premier article CHANGER (*Jean change de chambre*), il y a le causatif correspondant, CHANGER transitif (*Maman change Jean de chambre*), mais pas dans le second. Si l'on néglige la délimitation proposée, nous ne voyons pas comment faire figurer cette information indispensable à un bon usage de la langue.

Passons à présent à la formulation de nos critères. Ils ne constituent en fait qu'une formalisation des intuitions des lexicologues connues depuis fort longtemps; notre « originalité » est que nous les formulons tous ensemble et rigoureusement; nous les prenons très au sérieux et les acceptons comme des règles immuables.

## 2.2 Critères de délimitation des lexies

Comme on le sait, la délimitation des lexies, c'est-à-dire la division des acceptions d'un mot ou d'une locution polysème, est un des problèmes les plus délicats et les plus ardues qui se posent au lexicologue; les décisions prises à ce propos sont toujours lourdes de conséquences. On voit immédiatement qu'il s'agit de la question suivante, fort connue en lexicologie (et en lexicographie pratique !) :

|| Comment différencier entre l'homonymie, la polysémie et le caractère vague des lexies ?

(Pour une étude spécialisée de la question ainsi qu'une riche bibliographie, on peut consulter Moon 1987, Deane 1988 et Tuggy 1993.)

En fait, délimiter des lexies revient à trancher, dans chaque cas particulier, le problème de l'ambigu vs le vague. Commençons par deux exemples banals.

(2) *C'est Jean qui a peint ce plafond.*

<sup>1</sup> Dans cette construction, l'emploi du syntagme nominal au lieu du clitique est gênant :  
(i) \**faire changer Jean de direction.*

- a. *a peint* = 'a couvert le plafond de peinture'; Jean a fait du ravalement;
  - b. *a peint* = 'a couvert le plafond d'images artistiques'; Jean a produit une œuvre d'art.
- (3) *C'est ma tante Juliette.*
- a. *tante* = 'une sœur de ma mère';
  - b. *tante* = 'une sœur de mon père'.

Au premier abord, ces deux exemples semblent parallèles : dans les deux, une expression lexicale – PEINDRE dans (2) et TANTE dans (3) – correspond à deux réalités extralinguistiques différentes. L'usage courant qualifie les deux expressions d'*ambiguës*. Cependant, en lexicologie, la terminologie doit être bien plus fine.

Si par *expression lexicale* nous entendons une unité linguistique qui manifeste des caractéristiques apparentes d'une lexie, les définitions suivantes permettent de mettre en relief la distinction voulue.

#### Définition III.4 : Ambigu

|| Une expression lexicale est *ambiguë* si et seulement si elle correspond alternativement à plus d'une lexie [= soit à L<sub>1</sub>, soit à L<sub>2</sub>, soit à ...].

#### Définition III.5 : Vague

|| Une expression lexicale est *vague* si et seulement si son sens correspond alternativement à plus d'un référent extralinguistique, alors qu'elle-même correspond à une seule lexie.

Comme on le verra plus loin, PEINDRE est ambigu (au sens technique qui vient d'être introduit) : les interprétations (2a) et (2b) sont décrites par deux lexèmes différents. Quant au nom TANTE, ce n'est pas le cas : techniquement parlant, il est vague, mais pas ambigu; il ne constitue qu'une seule lexie (TANTE = 'une sœur de la mère ou du père ou l'épouse de l'oncle').

#### Remarque

L'usage terminologique d'*ambigu* et de *vague* se différencie donc de l'utilisation de ces mots dans le français courant. Ainsi, pour un usager normal, le nom TANTE n'a absolument rien de vague.

Les critères 1 à 5 ci-dessous n'ont d'autre but que de fournir un cadre relativement formel pour orienter l'intuition du lexicologue dans la bonne direction lors de sa recherche des distinctions entre l'ambigu et le vague. Plus spécifiquement, supposons que le lexicologue considère le statut à donner à une expression lexicale L, qu'on peut qualifier, pour les besoins de l'exposé, de lexie potentielle. Pour décider si cette candidate au statut de lexie est vraiment une lexie, le lexicologue doit procéder à deux types de vérifications :

- d'une part, pour chaque lexie potentielle L, il doit vérifier s'il peut la considérer comme une seule lexie ou bien s'il faut la scinder en deux lexies, L<sub>1</sub> et L<sub>2</sub>;
- d'autre part, pour chaque paire de lexies potentielles L<sub>1</sub> et L<sub>2</sub>, il doit vérifier s'il peut les considérer comme distinctes ou bien s'il faut les unir en une seule lexie L.

Les lexies potentielles à étudier sont ainsi passées au crible des cinq critères qui vont être présentés maintenant. Chaque critère permet de prendre une décision spécifique.

Les critères formulés visent une lexie potentielle L ayant le sens (... σ'... σ''...), où (σ') et (σ'') sont des composantes sémantiques.

### Critère 1 : Critère d'interprétation multiple

Si, pour une lexie potentielle L (... σ'... σ''...), on peut construire une phrase qui a deux interprétations uniquement à cause de L telles qu'une de ces interprétations fait intervenir le sens (... σ'...) et l'autre fait intervenir le sens (... σ''...), alors L doit être scindée, de sorte qu'au lieu d'une lexie L, on a deux lexies L<sub>1</sub> (... σ'...) et L<sub>2</sub> (... σ''...), sauf indication contraire du Critère 3 ci-dessous<sup>2</sup>.

Si on ne peut pas construire pour L une phrase à interprétations multiples du type signalé, le Critère 1 est sans effet.

Ce critère vise l'AMBIGUÏTÉ, qu'il permet de déceler en concomitance avec le Critère 3 (ce dernier établissant la distinction entre l'ambigu et le vague).

#### Exemple

Pour la lexie potentielle PROPOSER, la phrase (4) a deux interprétations :

- (4) *Jean leur a proposé un voyage.*
- a. ≈ 'Jean leur a offert un voyage';
  - b. ≈ 'Jean leur a suggéré de faire un voyage'.

On voit que PROPOSER est la seule source possible de cette « duplicité »; donc, en vertu du Critère 1, on doit y voir deux lexèmes. Cependant, avant de prendre une décision finale, il faut encore vérifier si le Critère 3 ne la contredira pas. On constate que le Critère 3 n'empêche pas la scission de PROPOSER; on est donc obligé de distinguer dans le DEC deux lexèmes PROPOSER : PROPOSER1 (pour (4a)) et PROPOSER2a (pour (4b)); voir DEC-1.

Le verbe PEINDRE dans (2) est aussi la seule source d'une double interprétation et, là aussi, le Critère 3 ne s'opposant pas à la scission, on en fera deux lexèmes (PEINDRE Y = 'couvrir Y de peinture pour le colorer ou le protéger' vs

2 Pour bien utiliser ce critère, il faut soigneusement écarter les autres sources possibles d'ambiguïté, et particulièrement, les ambiguïtés syntaxiques du type suivant :

- (i) – *Lis le livre sur le lit !* =
- a. 'le livre qui se trouve sur le lit';
  - b. 'le livre qui traite du lit';
  - c. 'en étant couché sur le lit'.

PEINDRE Y de Z = 'créer des œuvres d'art Z peintes sur Y', puisqu'on ne peut pas dire, sans faire un jeu de mots, \**Jean a peint sa voiture en bleu et son plafond de scènes bibliques*).

Par contre, pour TANTE, même si le Critère 1 recommande la scission, le Critère 3 ne la permet pas (voir plus loin, l'exemple (9), p. 65).

Le Critère 1 est formulé de telle façon que si la prémisse n'est pas remplie, il ne s'applique pas. Cela reflète le fait substantiel suivant : la présence d'une multiplicité d'interprétations doit immédiatement éveiller l'attention du lexicologue; par contre, l'absence d'une telle multiplicité n'entraîne aucune conséquence lexicographique. Le Critère 1 n'est donc valable que dans le sens positif (L le satisfait); autrement, il n'indique rien.

Il n'est pas toujours facile de distinguer une « vraie » multiplicité d'interprétations, c'est-à-dire, l'ambiguïté, d'une simple indétermination, c'est-à-dire du vague. Par exemple, la phrase (5) :

(5) *Le liquide s'est refroidi.*

peut être aussi interprétée de deux façons :

- a. 'Le liquide est devenu PLUS FROID' [*Le liquide s'est un peu refroidi*];
- b. 'Le liquide est devenu FROID' [*Le liquide s'est complètement refroidi*].

Cependant, contrairement à (4), cette double interprétation ne manifeste pas une ambiguïté, mais témoigne du caractère vague de (5). Le verbe SE REFROIDIR, qui est la source de cette double interprétation, renferme les deux sens dans une définition unie *disjonctive* : SE REFROIDIR = 'devenir plus froid ou froid'. Une telle description découle de l'application du Critère 3 ci-dessous. (On trouvera d'autres exemples de définitions disjonctives lors de la discussion des Critères 3-5.)

L'application du Critère 1 à la lexie potentielle S'OCCUPER de l'exemple (1) donne le résultat suivant :

- Pour les phrases de (1a) jusqu'à (1f), on ne décèle aucune ambiguïté; le Critère 1 reste sans effet.
- Par contre, la phrase (1g) est ambiguë : elle signifie soit 'prendre soin de', comme dans *Ce délégué de la Croix-Rouge s'occupe de la tribu des Mayalis*, soit 'étudier, travailler sur', comme dans *Cet ethnologue s'occupe de la tribu des Mayalis*. Le Critère 1 prescrit donc la scission de la lexie potentielle L en deux :  $L_1$  = 'X prend soin de Y de la façon dont X doit prendre soin de Y dans la situation en question' et  $L_2$  = 'X fait de la recherche sur Y'. Notons une troisième lexie  $L_3$  = 'X vit de son travail dans le domaine Y', cf. *Il s'occupe du commerce de laines au Moyen Âge* (S'OCCUPER =  $L_2$  ≈ 'étudie') vs *Il s'occupe de commerce de laines* (S'OCCUPER =  $L_3$  ≈ 'gagne sa vie avec la laine')<sup>3</sup>.

3 L'impossibilité de réunir S'OCCUPER =  $L_2$  et S'OCCUPER =  $L_3$  sous une même lexie est démontrée par le caractère ludique de :

(i) \**Jean s'occupe de la langue bretonne* [=  $L_2$ ] et de *commerce de chevaux* [=  $L_3$ ].

Il s'agit en fait de l'application du Critère 3, voir ci-dessous.



Il faut faire remarquer que le Critère 1, tout à fait comme les autres critères, n'est ni automatique, ni absolu, c'est-à-dire, qu'il ne s'applique pas mécaniquement, sans réflexion, et que le résultat obtenu n'est rien d'autre qu'une indication, et non une prescription.

Le Critère 2 traite des cas où le lexicologue perçoit une différence sémantique entre deux emplois d'une unité lexicale L, cette différence apparaissant plutôt comme mineure et n'entraînant pas de variations prononcées au niveau du comportement syntaxique de L. Plus précisément, la définition ⟨L⟩ ne couvre pas un emploi L', qui, lui, nécessite sa propre définition ⟨L'⟩; cette dernière ne comprend, par rapport à ⟨L⟩, que des précisions sémantiques contraignantes pour l'emploi L'. Le dilemme est donc de décider s'il vaut mieux avoir deux lexies apparentées – L et L' – ou bien postuler une lexie à définition assez générale et vague pour couvrir les deux emplois. Le Critère 2 doit faciliter la prise de décision.

Ce critère se base sur l'opposition « différences sémantiques locales ~ différences sémantiques globales ». Si l'emploi d'une lexie potentielle L' manifeste une différence sémantique par rapport aux emplois « prototypiques » (de L) telle qu'elle n'est pas liée à d'autres différences sémantiques ailleurs dans la définition, alors une telle différence est appelée *locale*; par contre, si elle est liée à d'autres différences sémantiques, alors elle est *globale*.

### Critère 2 : Critère de différence sémantique locale/globale

Si une différence sémantique observée entre deux emplois de la lexie L est locale, elle doit être couverte par une même définition : elle est prise en charge par un *supplément sémantique* décrit par une règle supplémentaire au régime; L ne doit donc pas être scindée.

Si une différence sémantique est globale, elle ne doit pas être couverte par une même définition : elle est prise en charge par deux définitions différentes, c'est-à-dire que L est scindée – de sorte qu'au lieu d'une lexie L, on a deux lexies L<sub>1</sub> et L<sub>2</sub> (= L').

#### Exemples

- (6) a. *Elle vendait des tapis* ⟨du poisson, des livres, ses services comme interprète, ...⟩  
*aux touristes* ⟨aux gens du quartier, aux conseillers municipaux, ...⟩.  
 vs  
 b. *Elle vendait ses caresses* ⟨ses faveurs, ses charmes, son corps, ...⟩ *aux touristes*  
 ⟨aux gens du quartier, aux conseillers municipaux, ...⟩.

Il paraît clair que (6b) est sémantiquement différent de (6a) : en (6a), il s'agit du sens prototypique de VENDRE – 'transmettre les droits de possession permanente [d'une marchandise Y pour...]', alors que en (6b) on voit une différence sémantique : 'transmettre les droits de possession TEMPORAIRE<sup>4</sup> SEXUELLE

4 Dans le cas de *vendre ses faveurs* ⟨ses charmes, ...⟩, il s'agit en fait plutôt de LOUER que de VENDRE, puisque le droit de possession n'est cédé que temporairement. Soulignons que la composante 'possession = posséder' doit être prise ici au sens primaire ≈ 'être propriétaire de' ; il ne s'agit nullement de jouer sur l'autre sens de *posséder* = 'copuler'.

[de la personne impliquée pour...]. Cette différence est pourtant locale : elle n'affecte en rien les autres composantes de la définition (ni, bien entendu, la syntaxe et la cooccurrence lexicale de VENDRE). En respectant le Critère 2, le lexicologue formule une seule définition pour ces emplois, en spécifiant cependant, par une règle sémantique additionnelle, l'emploi particulier dans (6b). (L'analyse lexicologique du verbe VENDRE sera reprise au Chapitre V.)

Pour la formulation du Critère 3, la notion de phrase à cooccurrence compatible et la notion de phrase normale sont capitales.

Nous appelons *phrase à cooccurrence compatible* pour la lexie potentielle L de sens (...  $\sigma'$ ... $\sigma''$ ...) une phrase qui contient L et dans laquelle L se trouve en combinaison

- 1) soit simultanément avec deux lexies L' et L'' telles que L' correspond à ( $\sigma'$ ), et L'' à ( $\sigma''$ ),
- 2) soit avec une lexie L' qui correspond simultanément à ( $\sigma'$ ) et à ( $\sigma''$ ).

Et nous appelons *phrase normale* une phrase qui ne renferme ni jeu de mots ni calembour.

### Exemples

- (7) a. Cas 1 : pour SE REFROIDIR, une phrase à cooccurrence compatible serait  
*Le liquide se refroidit un peu [= L'] et puis complètement [= L'']*.
- b. Cas 2 : pour BELLE-MÈRE 'mère du conjoint ou épouse du père qui a remplacé la mère décédée', une phrase à cooccurrence compatible serait  
*Mes [= L'] deux belles-mères s'entendaient parfaitement.*

Par contre, pour le verbe PRENDRE dans *prendre une décision* et dans *prendre une bière*, on ne trouve pas de phrase à cooccurrence compatible; une phrase comme *\*Jean a pris une décision et trois bières* n'est pas une phrase normale : elle est perçue comme un calembour (c'est une construction appelée *zeugme*, condamnée par la stylistique).

### Critère 3 : Critère de cooccurrence compatible (= critère de Green-APRESJAN)<sup>5</sup>

Si, pour la lexie potentielle L'...  $\sigma'$ ...  $\sigma''$ ..., on peut construire une phrase normale à cooccurrence compatible, alors L ne doit pas être scindée – de sorte qu'on a une seule lexie (= L) avec la disjonction dans sa définition : L'...  $\sigma'$  ou  $\sigma''$ ...

Si on ne peut pas construire une phrase normale à cooccurrence compatible pour L, le Critère 3 est sans effet.

Ce critère vise le VAGUE.

5 Le Critère 3 est une reformulation du critère d'APRESJAN, postulé pour une disjonction des composantes au sein d'une définition lexicographique (APRESJAN 1974 : 85). Le prototype de ce critère avait été proposé par G. GREEN (1969). Ceci nous permet de nommer le Critère 3 *Critère de Green-APRESJAN*.

## Exemples

- (8) BOMBARDER {lancer des bombes [= 'σ'] ou des obus [= 'σ''}]  
 Phrase à cooccurrence compatible :  
*Des avions [= 'L'] et des navires [= 'L''}] bombardaient le port inlassablement.*

Cela entraîne le traitement du verbe BOMBARDER comme un seul lexème – avec la disjonction indiquée dans la définition.

- (9) Pour BELLE-MÈRE, ainsi que pour de nombreux termes de parenté semblables (BEAU-FILS, TANTE, ...), le Critère 3 entraîne également un seul lexème ayant une définition à disjonction :  
*Tous mes beaux-fils étaient là – le fils aîné de ma femme, ainsi que les maris de nos trois filles.*

- (10) a. Considérons le verbe FLAMBER dans les phrases *Son gosier flambait*, *Son visage flambait* et *Ses yeux flambaient*.

Peut-on couvrir les trois sens différents qu'on y voit par un seul lexème avec une définition à disjonction ? Le Critère 3 donne une réponse négative pour le premier sens ('sensation de brûlure') vis-à-vis des deux autres :

- b. \**Son gosier et son visage (ses yeux) flambaient.*

Par contre, pour le deuxième et le troisième sens – ('avoir une teinte anormalement rouge') et ('avoir un éclat anormal') – on peut citer une phrase à cooccurrence compatible :

- c. *Son visage et ses yeux flambaient de fièvre.*

Par conséquent, ces deux sens doivent apparaître au sein d'un seul lexème à disjonction (= FLAMBERVII.b dans le DEC-2, qui s'oppose à FLAMBERV).

Pour revenir à l'exemple de S'OCCUPER, les deux lexèmes potentiels dégagés n'admettent pas une phrase à cooccurrence compatible :

- (11) ?*Jean s'occupe beaucoup de ses petits-enfants et de la lexicologie moderne.*  
 est également un zeugme et représente sans contredit un calembour. Le Critère 3 confirme donc l'existence de deux lexèmes S'OCCUPER.

Le critère de cooccurrence compatible est connu aussi sous une autre forme (Lakoff 1970) :

### Critère de construction avec « AUSSI »

Si, pour la lexie potentielle L de sens ('...σ'... σ''...), on peut construire une phrase de la forme *X fait P avec L*, et *X' fait P avec L aussi* (autant, tout comme, ...), le premier P faisant intervenir ('σ') et le deuxième, ('σ''), alors L ne doit pas être scindée – de sorte qu'on a une seule lexie (= L) avec la disjonction dans sa définition : L ('... σ' ou σ''...).

Ce critère, qui peut être appliqué en utilisant tout synonyme de AUSSI (en anglais, par exemple, l'expression... *and so does X*), est équivalent au Critère de cooccurrence compatible. Soulignons, cependant, que des contraintes grammaticales peuvent empêcher l'application de l'un de ces deux critères. Dans les

exemples cités, le Critère de construction avec « AUSSI » donne les mêmes résultats que le Critère de cooccurrence compatible :

- (7b') – *C'est vrai, si j'épouse Janine, mon pauvre bébé aura une belle-mère; mais moi, j'en aurai une aussi, puisque la mère de Janine vient habiter avec nous.*
- (8') *Des avions bombardaient le port inlassablement, et des navires aussi.*
- (10') a. *Son visage flambait tout comme ses yeux.*  
 b. *\*Son gosier flambait tout comme ses yeux.*

Pour SE REFROIDIR (exemple (5)), le Critère de construction avec « AUSSI » n'est pas applicable.

L'existence de ces deux critères équivalents donne au lexicologue une flexibilité supplémentaire pour la délimitation des lexies.

Les Critères 1-3 fonctionnent dans le cadre de phrases, c'est-à-dire que leur application exige des phrases du type prévu par le critère utilisé. Ce sont des critères de cohérence externe : cohérence entre les articles de dictionnaire et les textes. Par contre, les Critères 4 et 5 fonctionnent dans le cadre de la description lexicographique de la lexie en cause; leur utilisation est basée sur l'idée d'un article de dictionnaire qui évite les redondances et les répétitions fâcheuses. Ces critères sont des tests de cohérence interne des articles de dictionnaire.

#### Critère 4 : Critère de cooccurrence différentielle

Si, pour la lexie potentielle L (...  $\sigma'$ ...  $\sigma''$ ...), on peut dégager deux ensembles disjoints de cooccurents (morphologiques, syntaxiques ou lexicaux) tels que l'un correspond à ' $\sigma'$ ' et l'autre à ' $\sigma''$ ', alors L doit être scindée – de sorte qu'au lieu de L on a deux lexies  $L_1$  et  $L_2$ .

Si on ne trouve pas pour L de cooccurrence différentielle, le Critère 4 est sans effet.

#### Exemples

Comme premier exemple, prenons le verbe ACCEPTER avec les deux sens apparaissant dans les phrases (12a) vs (12b) :

- (12) a. *Jean a accepté son exil (ce mode de vie) sans rechigner [= ne s'y oppose plus].*  
 b. *Jean a accepté ce cadeau (ce poste) sans rechigner [= a effectivement pris].*

Il peut sembler que ACCEPTER peut être décrit en (12a-b) comme un seul lexème L ayant une définition assez générale :

(Réagissant à la situation Y [= ' $\sigma'$ '], en particulier, quand Z offre Y à X [= ' $\sigma''$ '], X fait ce que Y exige de lui sans s'y opposer).

Pendant, l'analyse fait apparaître des cooccurents syntaxiques différents pour les deux sens. Ainsi, premièrement, en (12a), ACCEPTER admet une proposition infinitive :

- c. *Jean a accepté d'être exilé (de vivre comme ça) sans rechigner.*

En (12b), par contre, ACCEPTER n'admet pas d'infinitif sans affecter le sens; quand on dit

d. *Jean a accepté de recevoir ce cadeau.*

on retrouve le premier sens : l'acceptation mentale, sans action correspondante (Jean est seulement d'accord de prendre le cadeau); (12b) implique que Jean a FAIT quelque chose : il a pris le cadeau.

Deuxièmement, en (12b), mais pas en (12a), ACCEPTER admet un actant de plus, introduit par DE :

e. *Jean a accepté ce cadeau de ses parents <ce poste de l'université> sans rechigner.*

Le Critère 4 nous force à distinguer deux lexèmes : ACCEPTER<sub>I.1a</sub> 'réagissant à la situation Y, X fait ce que Y exige de lui ou ne s'y oppose pas' (pour (12a)) et ACCEPTER<sub>I.3</sub> 'Z offrant Y à X, X accepte<sub>I.1a</sub> de prendre Y de Z et prend Y' (pour (12b)); voir DEC-3. Soulignons que la différence syntaxique indiquée n'est pas provoquée au niveau de surface par la simple différence formelle de deux CO<sup>dir</sup> (CADEAU vs ÊTRE EXILÉ) mais tire son origine d'une différence sémantique profonde. Considérons, par exemple, la situation suivante : 'Les parents ont donné à Jean un cadeau, qu'il a pris'. Si, tout en voulant utiliser ACCEPTER, on choisit de présenter cette situation sémantiquement sans référence explicite à une offre, on peut construire ACCEPTER avec l'infinitif mais on ne peut exprimer l'offrant :

f. *Jean a accepté (\*de ses parents) de prendre le cadeau.*

Si, par contre, le locuteur présente la situation en question comme une offre, il doit construire ACCEPTER avec un nom et peut exprimer l'offrant :

g. *Jean a accepté (de ses parents) le cadeau.*

Le Critère 3 ne contredit pas notre décision : une phrase comme 'J'accepte l'exil et ce cadeau en souvenir' (proposée par J.-L. Descamps) n'est évidemment qu'un jeu de mots.

En tant que deuxième exemple, considérons les deux S'OCCUPER dégagés dans la discussion du Critère 1, p. 62. Il semble bien qu'il n'y ait pas de cooccurrence différentielle importante entre ces lexies :

(13) a. Les mêmes formes morphologiques sont possibles :

*Occupe-toi*  $\left\{ \begin{array}{l} \text{du thé} \\ \text{de littérature} \end{array} \right\};$

*Il s'est occupé*  $\left\{ \begin{array}{l} \text{du thé} \\ \text{de littérature} \end{array} \right\}$  *avec ferveur.*

- b. Les mêmes modifications sont possibles :

*Il s'occupe*  $\left\{ \begin{array}{l} \text{du thé} \\ \text{de littérature} \end{array} \right\}$  *avec ferveur*

*Il s'est occupé*  $\left\{ \begin{array}{l} \text{du thé} \\ \text{de littérature} \end{array} \right\}$  *corps et âme.*

Cependant, nous pouvons quand même indiquer une différence dans la cooccurrence syntaxique :

- c. L'emploi de l'article : *Il s'occupe du thé* = 'Il prend soin du thé' vs *Il s'occupe de thé* = 'Il travaille dans l'industrie et le commerce du thé/Il étudie le thé'.

Quant au Critère 3, il confirme cette division des sens : \**Il s'occupe de thé et de ses enfants* (cf. par contre, *Il s'occupe du thé et de ses enfants*).

### Critère 5 : Critère de dérivation différentielle

Si, pour la lexie potentielle  $L \langle \dots \sigma' \dots \sigma'' \dots \rangle$ , on peut dégager deux ensembles disjoints de dérivés (morphologiques ou lexicaux) tels que l'un correspond à ' $\sigma'$ ', et l'autre à ' $\sigma''$ ', alors  $L$  doit être scindée – de sorte qu'au lieu de  $L$  on a deux lexies  $L_1$  et  $L_2$ .

Si on ne trouve pas pour  $L$  de dérivation différentielle, le Critère 5 est sans effet.

#### Exemples

Soit ÉLEVER dans les deux acceptions illustrées dans (14) :

- (14) a. *Toute sa vie, Jeanne a élevé des enfants* [par exemple, Jeanne est une jardinière d'enfants].  
 b. *Toute sa vie, Jeanne a élevé des cochons* [par exemple, Jeanne est une fermière].

Dans (14b), ÉLEVER a comme dérivé le nom d'action ÉLEVAGE (*Jeanne s'occupe d'élevage de cochons*) et le nom d'agent ÉLEVEUR/ÉLEVEUSE (*Jeanne est éleveuse de cochons*). Cependant, pour (14a), ces dérivations sont impossibles (\**Jeanne s'occupe d'élevage d'enfants*; \**Jeanne est éleveuse d'enfants*); au besoin, on utilise les dérivés lexicaux (= supplétifs) ÉDUCATION, FORMATION, INSTITUTEUR, «JARDINIÈRE D'ENFANTS», etc.

Le Critère 5 permet de distinguer les deux acceptions ('élever des enfants' et 'élever des animaux') de la troisième (à caractère technique) qu'on voit dans (14c) :

- c. *Toute sa vie, Jeanne a élevé du vin* [Jeanne est viticultrice];

même si ÉLEVEUR/ÉLEVEUSE *de vin* est possible, on n'a pas de \*ÉLEVAGE *de vin*. Le dégagement de l'acception 'élever du vin' est confirmé par le Critère 3 (cooccurrence compatible) :

- d. \**Toute sa vie, Jeanne a élevé du vin et des cochons.*

De la même façon, le Critère 5 s'applique au verbe PRENDRE dans *prendre un médicament* et *prendre une bière*. Dans le premier cas, la nominalisation existe (*une prise de médicament*), mais pas dans le deuxième (*\*une prise de bière*). Cependant, il s'agit ici du seul décalage que nous connaissons, et il peut être traité comme exception. (Des exceptions disparates ne posent pas de problème pour un DEC; les Critères 4 et 5 essaient d'éviter les exceptions SYSTÉMATIQUES, qui, en fait, représentent d'autres régularités.) De plus, le Critère 3 va à l'encontre de la scission du verbe PRENDRE dans les contextes donnés; en effet, on peut dire facilement : *Tu prends trop de vin et de médicaments en même temps* ou *J'ai pris une aspirine et un café très fort*; ce sont des phrases normales. Nous avons ainsi une nouvelle démonstration de la puissance du Critère 3 vis-à-vis des Critères 4 et 5 : le Critère 3 veille à la conformité de la description lexicologique par rapport à la réalité linguistique, alors que les Critères 4 et 5 ne sont responsables que pour l'esthétique de cette description.

La délimitation des lexies est au fond concomitante avec la description de celles-ci; en fait, les deux opérations s'effectuent quasiment en parallèle. Une délimitation bien-fondée des lexies assure les assises voulues pour leur description, tandis qu'une description suffisamment détaillée et prenant en charge toutes les données disponibles confirme la délimitation proposée – ou, au contraire, force à une révision de cette dernière.

En spécifiant les critères de délimitation des lexies et en en illustrant l'application, nous avons réglé le problème de l'ambigu et du vague; nous allons maintenant aborder le problème de la description lexicographique des lexies.

### 3 Description d'une lexie = structure d'un article de dictionnaire

Après les mises au point nécessaires (3.1), nous examinerons en détail le contenu d'un article du DEC, en procédant zone par zone (3.2-3.7).

#### 3.1 Généralités

La description d'une lexie dans le DEC n'est pas autre chose qu'un article de dictionnaire. Tout article du DEC, quelle que soit la nature de la lexie décrite, doit être traité selon le même moule, c'est-à-dire en suivant rigoureusement un article patron, qui détermine la démarche. Tout article de dictionnaire aboutit donc à la même structure formelle.

Rappelons qu'un article de dictionnaire du DEC correspond toujours à une SEULE lexie L, qui, elle aussi, correspond à un SEUL article de dictionnaire.

**Définition III.6 : Vedette**

|| La lexie L décrite par un article de dictionnaire s'appelle *lexie vedette* ou *vedette* de cet article.

Dans la littérature spécialisée, on trouve souvent le terme synonyme *entrée* (de l'article de dictionnaire).

Pour éviter toute confusion, rappelons encore que la lexie vedette peut être soit un lexème, comme ÉLÉPHANT, ACCEPTER, ROUGE, etc., soit un phrasème, c'est-à-dire un syntagme composé de plusieurs lexèmes, comme « METTRE EN JOUE », « ÉLÉPHANT DE MER », « EN VOIR DES TOUTES LES COULEURS », « ON N'EST JAMAIS SI BIEN SERVI QUE PAR SOI-MÊME », etc.

**NB :** Ce ne sont, bien entendu, pas tous les phrasèmes qui constituent des lexies vedettes. Les collocations (= semi-phrasèmes), où un lexème L conserve totalement son sens et ses propriétés syntaxiques, sont inscrites sous ce lexème L, dans la zone de combinatoire lexicale (restreinte), voir Ch. III, 3.5, p. 125. Ainsi, l'expression MÉMOIRE D'ÉLÉPHANT, qui désigne une espèce de mémoire, sera traitée – par fonctions lexicales – sous la vedette MÉMOIRE.

Un article de dictionnaire du DEC se compose de trois zones majeures :

- 1) zone phonologique (= le signifiant de la lexie vedette L) ;
- 2) zone sémantique (= le signifié de L) ;
- 3) zone de combinatoire (= le syntactique de L).

Cette dernière zone se divise tout naturellement en quatre parties :

- 3.1) la sous-zone de combinatoire morphologique (partie du discours, types de déclinaison / conjugaison, formes non réalisables, formes irrégulières, etc.);
- 3.2) la sous-zone de combinatoire stylistique (marques d'usage du type **spéc.**, **litt.**, **fam.**, **arch.**, etc., qui indiquent le type de contexte textuel dans lequel la lexie peut s'insérer);
- 3.3) la sous-zone de combinatoire syntaxique (classe et sous-classes syntaxiques, constructions typiques, régime, etc.);
- 3.4) la sous-zone de combinatoire lexicale restreinte.

À ces trois zones majeures, qui sont une réflexion directe de la nature sémiotique de la lexie (en tant que signe linguistique), s'ajoutent deux zones auxiliaires, importantes du point de vue dictionnaire :

- 4) zone d'illustrations ;
- 5) zone phraséologique.

Un article de dictionnaire du DEC est donc structuré en fonction de ces zones. Notons cependant que l'information relevant des sous-zones 3.1, 3.2 et 3.3 reçoit un traitement particulier. La caractérisation morphologique de la lexie vedette ainsi que les marques d'usage (3.1 et 3.2) sont indiquées avant la zone sémantique – immédiatement après le numéro lexical. Cette façon de faire correspond à



l'usage en cours dans les dictionnaires et a une justification pédagogique : ces indications aident le lecteur dans l'identification de la lexie. Nous n'avons pas remis en question cette tradition, même si elle brise la logique de la structuration des articles du DEC, essentiellement parce que la description morphologique des lexies et les marques d'usage n'ont pas fait l'objet d'une étude approfondie dans le cadre du DEC. La même remarque s'applique aux propriétés syntaxiques des lexies qui ne relèvent pas du régime (une partie de la sous-zone 3.3). Elles n'ont jamais été décrites dans le DEC de façon systématique et apparaissent de façon informelle dans des NB de la zone syntaxique. Les sous-zones dont il vient d'être question ne sont pas discutées ci-dessous<sup>6</sup>.

Nous procédons maintenant à la description d'un article de dictionnaire du DEC zone par zone.

### 3.2 Zone phonologique

La zone phonologique, où le dictionnaire spécifie la forme phonologique du signifiant de la lexie vedette accompagnée, s'il y a lieu, de remarques sur les problèmes d'orthoépique et d'orthographe, est assez bien traitée dans les dictionnaires usuels du français et ne présente pas de difficulté théorique majeure. Pour cette raison, elle ne sera pas considérée dans ce livre.

Nous aimerions quand même signaler que la plupart des dictionnaires existants négligent complètement les données prosodiques<sup>7</sup>. Ces dernières ont pourtant une importance lexicographique indiscutable; rappelons, à titre d'exemple, le cas de l'adjectif VRAI, traité au Ch. I, 3.3, p. 31. Un autre exemple évident est l'adverbe MERCI (BIEN) :

(15) A : *Tu veux un autre verre ?*

↗  
B<sub>1</sub> : *Merci !* [= 'Oui']

↘  
B<sub>2</sub> : *Merci !* [= 'Non']

[Les flèches indiquent l'intonation montante ou descendante.]

6 Il existe en fait une raison méthodologique pour laquelle nous n'avons pas pu présenter de façon satisfaisante les caractéristiques morphologiques et syntaxiques des lexies consignées dans le DEC du français. Le DEC ne peut être bien connecté qu'aux descriptions morphologiques et syntaxiques de la langue qui s'inscrivent dans le même cadre linguistique théorique que le DEC : c'est-à-dire la théorie Sens-Texte. Dans notre démarche, le dictionnaire et la grammaire d'une langue doivent être parfaitement accordés. Comme nous ne disposons pas d'une grammaire formelle Sens-Texte du français, nous devons nous contenter, pour le moment, d'approximations. On voit ici une très bonne illustration de ce que nous avons mentionné au Ch. I, 3.2, pp. 28-29 : le lexicologue doit être en même temps un grammairien.

7 Il faut absolument mentionner ici les efforts de Ju. Apresjan, qui a attiré l'attention des lexicologues sur l'importance de ce problème et a proposé un modèle de représentation (Apresjan 1990a).

La prosodie distingue aussi le phrasème « ON NE PEUT PAS SE PLAINDRE » (Les choses sont relativement convenables) et l'expression libre *On ne peut pas se plaindre* (Il est interdit de se plaindre). Le phrasème se prononce sur une intonation rectotonale, sans mise en relief accentuelle, alors que l'expression libre porte l'accent de force sur PAS et manifeste une intonation montante-descendante :

(16) a. À l'armée, on ne peut pas se plaindre [= 'la vie n'est pas si mauvaise'].

vs.

b. À l'armée, on ne peut pas se plaindre [= 'on n'a pas le droit de se plaindre'].

### 3.3 Zone sémantique

La zone sémantique d'un article de dictionnaire comprend deux sous-zones : la définition *lexicographique* de la lexie vedette L et les *connotations* de L. De cette façon, le sémantisme de L, c'est-à-dire le « contenu sémantique » que les locuteurs associent à L – son signifié – est présenté, dans un DEC, dans deux compartiments différents de l'article de dictionnaire de L. Cette division reflète la présence de deux types clairement distincts d'éléments sémantiques : les uns, de loin les plus importants puisque toujours présents, forment la partie DÉNOTATIONNELLE du sens de L et constituent sa définition (3.3.1); les autres, superfétatoires, absents de la plupart des lexies (mais accompagnant les lexies les plus usuelles) forment la partie CONNOTATIONNELLE du sens de L et sont décrits comme ses connotations (3.3.2, p. 111).

#### 3.3.1 Définition lexicographique

La discussion de la définition lexicographique dans le DEC passe par les six points suivants :

- présentation de la définition lexicographique,
- actant sémantique,
- règles d'élaboration de la définition lexicographique,
- critères de formulation,
- structure et contenu,
- liste de vérification.

(Pour un complément de discussion sur le problème de la définition lexicologique/ lexicographique, toujours fort débattu, nous pouvons conseiller Robinson 1954, Rey 1965, Rey-Debove 1966 et Chaurand & Mazière 1990. Plus particulièrement, dans notre optique, les articles Apresjan 1979 : 512-158 et Wierzbicka 1992 et 1992/93 sont d'importance primordiale.)

### 3.3.1.1 Présentation de la définition lexicographique

La DÉFINITION LEXICOGRAPHIQUE présente de façon formelle le SENS, OU SIGNIFIÉ (DÉNOTATIONNEL), de la lexie vedette L. (Le sens dénotatif est aussi appelé *sens situationnel* ou *sens propositionnel*.)

L'approche Sens-Texte, qui constitue le cadre de référence du DEC, est basée sur la thèse suivante :

|| La plupart des propriétés de comportement d'une lexie sont sous-tendues ou même carrément déterminées par son sens dénotatif.

On comprendra donc aisément que la définition lexicographique est considérée comme le noyau central de l'article de dictionnaire. En effet, il est impossible de traiter en profondeur le régime syntaxique ou la cooccurrence lexicale de L avant d'avoir formulé son sens. (Comme nous ne considérons ici que le sens DÉNOTATIONNEL, nous nous permettons d'omettre cet adjectif.)

La définition lexicographique d'une lexie L n'est pas autre chose qu'une représentation du sens de L, soit une représentation sémantique [= RSém]. Dans la théorie Sens-Texte, une RSém s'écrit obligatoirement sous la forme d'un réseau. Un réseau sémantique est un objet formel constitué de points (= nœuds) reliés par des flèches (= arcs); les nœuds sont étiquetés de noms d'éléments sémantiques, et les arcs sont munis de numéros de relations « prédicat ~ arguments ». Il n'est pas dans notre intention de développer ici l'ensemble du concept de RSém (voir, par exemple, Mel'čuk 1988 : 52 *ssq.*); nous l'avons mentionné parce que cette RSém est tout à fait indispensable dans notre approche théorique. Pour des raisons qui seront explicitées ci-dessous, nous n'utilisons pas le formalisme des RÉSEAUX SÉMANTIQUES pour décrire le sens des lexies. Nous avons recours, à la place, à des définitions rédigées en français, qui sont semblables aux définitions que l'on trouve habituellement dans les dictionnaires courants. Précisons tout de suite qu'il s'agit en fait d'un français « traité » : le langage définitoire est contraint et exclut un certain nombre de lexies françaises. Il est aussi obligatoirement désambiguïté – par l'utilisation de numéros lexicaux (voir p. 87, après la Règle 4). Il admet encore en principe des combinaisons pouvant violer certaines normes stylistiques. La question du langage définitoire sera reprise de façon plus détaillée plus loin, 3.3.1.3, p. 78 *ssq.* ainsi que 89-90.

Notre choix de la définition rédigée signifie en fait que, pour une représentation du sens dans le DEC, nous avons privilégié la LINÉARITÉ, caractère essentiel du langage naturel, en opposition à la MULTIDIMENSIONNALITÉ du réseau, objet logique. Ce choix se justifie au moins par les deux raisons suivantes, l'une pratique, l'autre linguistique.

La première raison de notre préférence est tout simplement la difficulté de production et de perception des réseaux. D'une part, les réseaux sémantiques sont difficiles à reproduire typographiquement; d'autre part, ils sont délicats à interpréter par un usager sans formation spécialisée. Les obstacles techniques et

commerciaux, joints aux conditions pédagogiques et psychologiques, nous ont amenés à surseoir à la création d'un dictionnaire à réseaux. Même si un tel dictionnaire est, à notre avis, une réalisation indispensable, nous croyons qu'il reste encore, pour le moment, un idéal hors d'atteinte.

La deuxième raison est, quant à elle, conditionnée par la nature même du langage humain – par sa linéarité. En effet, un réseau sémantique, bien adapté à l'analyse et à l'interprétation logiques, n'est pas directement accessible à l'intuition linguistique du locuteur. Le locuteur qui juge de l'équivalence entre une expression française et le réseau censé représenter le sens de cette expression se voit obligé de reformuler le réseau sous forme d'un (quasi-)texte linéaire pour permettre à son intuition de bien saisir le contenu linguistique. Cette reformulation est d'autant plus naturelle que l'intuition linguistique, en général, n'est soutenue que par des textes : on ne parle pas en réseaux (Polguère 1992 : 130-136) !

Le DEC de l'avenir devra avoir une double représentation du sens pour chaque lexie, soit une définition linéaire répondant à l'intuition linguistique des usagers et un réseau correspondant facilitant l'analyse logique. Ainsi, pour le lexème PROFESSEUR (comme *professeur de math au lycée Victor-Hugo / professeur de math de Jean*; voir la discussion détaillée plus loin, 3.3.1.3, p. 91 *ssq.*) on pourrait avoir les représentations de sens parallèles suivantes :

**sous forme de définition rédigée**

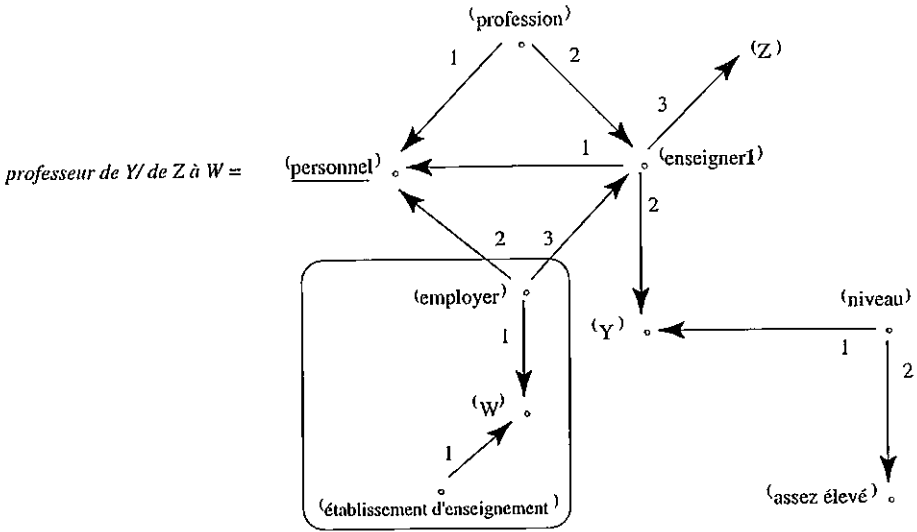
*professeur de Y / de Z (à W)* = 'Personne qui, [en tant que] profession, enseigne<sup>I</sup> Y d'un niveau assez élevé à Z (étant employée par un établissement d'enseignement W)'.

[Dans une définition, une composante entre parenthèses est optionnelle. Dans le cas présent, la composante 'étant employée par un établissement d'enseignement W' est optionnelle car on appelle *professeur de langue <de piano>* même quelqu'un qui enseigne<sup>I</sup> les langues ou le piano de façon privée.]

Pour donner une meilleure perception du sens décrit ci-dessus, citons maintenant la définition de 'enseigner<sup>I</sup>', qui apparaît dans la définition de 'professeur' :

*X enseigne<sup>I</sup> Y à Z* = 'X, cause volontairement que Z apprenne<sup>III.1b</sup> Y, en transmettant, systématiquement à Z des connaissances (portant sur) Y ou des techniques (portant sur) Y'.

sous forme de réseau sémantique



[Dans un réseau, le nœud dominant est souligné; dans le cas présent, c'est le sens 'personne'. Cela signifie que la composante sémantique 'personne' « résume » notre réseau : un professeur est essentiellement une 'personne qui...' – avec les qualificatifs indispensables. La partie encadrée est optionnelle.]

Il est même possible de penser à une application informatique, où un algorithme ferait des transformations automatiques entre les définitions linéaires et les réseaux sémantiques (voir, à ce propos, Ch. VI, 3.3, p. 236). Cependant, pour le moment, nous devons être plus modestes et nous contenter encore de définitions rédigées.

En résumé, une *définition lexicographique* dans le DEC est une expression de la forme

$$A = \langle B \rangle$$

où *A* – le *défini* – est la lexie vedette *L* incluse dans ce que nous appelons la *forme propositionnelle* (expression à variables), et *\langle B \rangle* – le *définissant* – est une description du sens de *L*, c'est-à-dire de *\langle L \rangle*, faite dans un métalangage lexicographique, qui est du français « traité ».

3.3.1.2 *Concept fondamental de la définition lexicographique : actant sémantique*

Pour bien comprendre la suite de notre exposé, le lecteur doit avoir à l'esprit un concept sémantique fort important : l'*actant sémantique* d'une lexie. Ce concept est, à son tour, défini par les concepts de *prédicat sémantique* et d'*argument de*

**prédicat sémantique.** Nous ne pouvons pas définir ici ces deux derniers concepts; nous les prenons pour acquis et nous renvoyons le lecteur intéressé aux manuels de logique (par exemple, Lepage 1991 : 115 *ssq.*) ou de sémantique (Lyons 1977 : 147 *ssq.*). Nous rappellerons, quand même, que *grosso modo* un prédicat (pris comme terme logico-sémantique) est un sens qui a des « trous » pour recevoir d'autres sens; un sens prédicatif est un sens « liant » – il réunit d'autres sens en des configurations sémantiques tout comme un tube de jonction réunit les pôles d'une tente pour former le squelette porteur de la tente. Les prédicats sémantiques désignent des actions, des événements, des processus, des états, des propriétés, des relations, etc. – en un mot, des faits qui impliquent nécessairement des participants. Si un prédicat (sémantique) P représente un fait ayant, par exemple, trois participants, les arguments  $A_1, A_2, A_3$  de P correspondent aux participants de ce fait, ce qui s'écrit comme suit :  $P(A_1, A_2, A_3)$ .

Citons quelques exemples :

- Le sens de DONNER est un prédicat à trois arguments : *quelqu'un* [1] *DONNE quelque chose* [2] à *quelqu'un* [3].
- Le sens de la préposition SOUS est un prédicat à deux arguments : *quelque chose* [1] [*est*] *sous quelque chose* [2], tout comme le sens d'ÉPOUX (*quelqu'un* [1] [*est*] *ÉPOUX de quelqu'un* [2]) et de POIDS (*quelque chose* [1] *a un POIDS de...* [2]); un autre exemple est le phrasème <sup>1</sup>TIRER SA RÉVÉRENCE <sup>1</sup> [= 'quitter N?'] (*quelqu'un* [1] *TIRE SA RÉVÉRENCE à quelqu'un* [2] : *M. Tartampion a tiré sa révérence à l'Université*).
- Le sens de SOMMEIL ou de DORMIR est un prédicat à un argument : *quelqu'un* [1] *DORT*; d'autres prédicats à un argument sont : TÊTE (*TÊTE de quelqu'un* [1] = *quelqu'un* [1] *a une TÊTE*), GRAND [taille humaine] (*quelqu'un* [1] *est GRAND*), <sup>1</sup>CASSER SA PIPE <sup>1</sup> [= 'mourir'] (*quelqu'un* [1] *CASSE SA PIPE*), etc.

Si en logique on parle d'un prédicat et de ses arguments, en linguistique et, plus particulièrement, en lexicologie, on utilise – pour désigner les mêmes choses – plutôt les termes *sens prédicatif* et *actants sémantiques*. Le concept d'actant sémantique est indispensable pour la description des lexies à sens prédicatif, c'est-à-dire des lexies dont le sens est un prédicat sémantique. Cependant, pour définir l'actant sémantique, il faut d'abord préciser la façon d'écrire la définition pour une lexie à sens prédicatif. Pour ce faire, nous formulons la contrainte suivante :

- || Soit une lexie L dont le sens est un prédicat à  $n$  arguments :  $\langle L(A_1, A_2, \dots, A_n) \rangle$ .  
 || La définition de L doit alors inclure nécessairement  $n$  variables.

### Définition III.7 : Actant sémantique

- || Nous appelons *actant sémantique* [= ASém] de la lexie L une expression qui correspond à un argument du prédicat  $\langle L(A_1, A_2, \dots, A_n) \rangle$ ; cette expression est soit un sens, soit une variable dans la définition de L.

En explicitant les exemples ci-dessus, nous pouvons dire que DONNER (dans, par exemple, *Jean donne une fleur à Marie*) a trois ASém : X – celui qui donne, Y – celui qui reçoit, et Z – ce qui passe de X à Y; ÉPOUX n'a que deux ASém : X – celui qui est l'époux, et Y – celle dont X est l'époux; GRAND a un seul ASém : X – celui / ce qui est grand.

Les quatre précisions suivantes nous semblent de rigueur :

- Toutes les lexies n'ont pas des ASém. Conformément à la définition donnée, seulement les lexies dont le sens est un prédicat possèdent des ASém. Ainsi, EAU, SOLEIL, FLEUR, SABLE, HOMME, CHIEN, ... n'ont pas d'ASém.
- Pourtant, la plupart des lexies d'une langue  $\mathcal{L}$  ont des ASém. Comme le montrent nos exemples, ce ne sont pas uniquement des verbes (même si on a tendance à le croire, à cause de la confusion possible des ASém avec des compléments de tout genre). Plus précisément, les ASém apparaissent nécessairement dans tous les verbes, tous les adjectifs et adverbes, toutes les prépositions et conjonctions, et dans la majorité des substantifs. Ainsi, possèdent des ASém les noms d'action, d'événement, de processus, etc. (EXÉCUTION de  $Y$  par  $X$ , MORT de  $X$ , CHANGEMENT de  $X$  en  $Y$ , ...), les noms des parties du corps (NEZ de  $X$ ), les noms de fonctions et de relations ( $X$  est le MINISTRE de  $Y$  du pays  $Z$ ,  $X$  est un FRÈRE de  $Y$ ), les noms d'institutions et d'artefacts (ÉCOLE où  $X$  enseigne  $Y$  à  $Z$ , CUILLÈRE avec laquelle  $X$  mange  $Y$ ), etc. En fait, les seules lexies d'une langue quelconque dénuées d'ASém sont les noms propres (personnes, lieux, ethnies, ...), les noms de certaines substances (liquides, solides, poudres, ...), de certains phénomènes naturels (éléments géographiques, météorologiques, ...), et de certaines espèces naturelles (animaux, oiseaux, fleurs, ...), etc. Il est souvent impossible d'assigner aux sens de tels types de lexies une situation qui présuppose des participants clairement identifiés, de telles situations étant indispensables pour qu'il y ait actants sémantiques.

**NB :** Nous nous devons de signaler ici que l'affirmation ci-dessus ne s'applique évidemment pas à la TOTALITÉ des substances, des phénomènes naturels et des espèces. En effet, une substance qui est importante pour la société humaine (le pétrole, l'uranium, etc.) crée une structure actancielle appropriée (à quoi on l'utilise, qu'est-ce qu'on fait d'elle, ...). Ceci est également vrai pour certains phénomènes naturels et certaines espèces.

- Le nombre d'ASém varie, en fonction des langues, d'un à six (ce dernier chiffre est un constat; il n'a aucune valeur théorique); comme exemples de lexies à 4, 5, et 6 ASém, nous pouvons citer VENDRE ( $X$  vend  $Y$  à  $Z$  pour la somme  $W$ ), LOUER ( $X$  loue  $Y$  à  $Z$  pour la somme  $W$  pendant la période  $T$ ) et EXILER ( $X$  exile  $Y$  de  $L_1$  à  $L_2$  pour  $Z$  pendant la période  $T$ ). Soulignons que la présentation de tous les ASém est indispensable pour préserver le sens voulu dans la définition. Ainsi, si, dans la définition de LOUER, on supprime l'ASém  $T$  (= la période de location), le résultat n'est plus 'louer' : c'est 'vendre';

si, dans la définition de VENDRE, on supprime l'ASém W (= la somme d'argent), le résultat n'est plus 'vendre' : c'est 'donner'; etc.

- Un ASém de L n'est pas forcément présent à la surface dans toute phrase où L apparaît. La phrase *Pierre a vendu sa voiture* présuppose qu'il y a un acheteur [= Z] et que Pierre a reçu une somme d'argent [= W] pour sa voiture, même si l'acheteur et la somme ne sont pas exprimés syntaxiquement. Nous reviendrons sur la relation entre les ASém et leur expression syntaxique plus loin, dans la section 3.4, p. 117.

### 3.3.1.3 Règles d'élaboration de la définition lexicographique

Pour que les définitions d'un DEC répondent aux exigences de rigueur logique, il est absolument indispensable de respecter, dans la rédaction d'une définition lexicographique, certaines RÈGLES D'ÉCRITURE. Dans la rédaction de définitions, nous distinguons quatre aspects concernant :

- 1) le défini particulier,
- 2) le définissant particulier,
- 3) l'ensemble des définissants dans le DEC,
- 4) la relation entre le défini et le définissant particuliers.

Les règles d'élaboration de définitions que nous proposons ci-dessous correspondent à ces quatre aspects :

- 1) la Règle de forme propositionnelle concerne la forme du défini;
- 2) la Règle de décomposition et la Règle de bloc maximal concernent la forme du définissant;
- 3) la Règle de standardisation concerne l'ensemble des définissants;
- 4) la Règle de substituabilité concerne le rapport entre un défini donné et son définissant.

L'ensemble de ces règles détermine en fait formellement une langue de définition.

#### Règle 1 : Règle de la forme propositionnelle

Pour des raisons logiques évidentes, les actants sémantiques, qui apparaissent inévitablement dans le définissant, doivent aussi figurer dans le défini. Par conséquent, contrairement aux autres dictionnaires, le défini dans un DEC ne peut pas être une lexie seule. Pour une lexie prédicative L, le défini doit obligatoirement être cette même lexie L munie de VARIABLES qui renvoient aux ASém de L. (Les variables sémantiques sont symbolisées par des lettres majuscules.)

Ainsi, pour définir REPROCHER (par exemple, dans *Marie reproche son absence à Pierre*), il faut en fait définir

*X reproche Y à Z;*



pour CHANGER (dans *Marie a changé son fils d'école*), le défini est en réalité *X change Y de Z*;

pour SOMMEIL (dans *Le sommeil de Marie a été perturbé*), le défini est *sommeil de X*; etc.

Les définis sont représentés par des expressions appelées *formes propositionnelles*.

Les formes propositionnelles ne concernent, bien entendu, que des lexies prédicatives, c'est-à-dire, comme on l'a déjà dit, des lexies dont le sens est un prédicat sémantique.

La Règle de la forme propositionnelle se formule donc comme suit :

- || Pour une lexie prédicative L, le défini doit être une forme propositionnelle où les variables représentent les actants sémantiques de L.

Nous ne discuterons pas ici de l'utilité de l'emploi, dans la définition, des variables actanciennes. Nous nous permettons quand même de faire remarquer que l'utilisation des formes propositionnelles en tant que définis lexicographiques est une nouveauté; à notre connaissance, aucun autre dictionnaire n'a encore choisi cette voie. (Une seule exception est le dictionnaire Collins Cobuild 1987, où les définitions sont introduites par des expressions du type suivant (souligné) : « Si vous interrogez quelqu'un, vous le questionnez à fond et pendant un certain temps dans le but d'obtenir de lui des informations » [notre traduction]. Pourtant, l'usage des pronoms pour représenter des actants nous semble assez délicat, surtout s'il s'agit de plusieurs actants du même type, comme dans « quelqu'un communique quelque chose à quelqu'un dans le but d'inciter quelqu'un à faire quelque chose à quelqu'un ». Une représentation par des variables algébriques nous semble beaucoup plus claire.)

## Règle 2 : Règle de décomposition<sup>8</sup>

Affirmer que le défaut fondamental de tous les dictionnaires est la circularité des définitions est depuis toujours un truisme. En effet, on définit, par exemple, (i)  $A = B + C$ ; B, à son tour, est défini comme (ii)  $B = D + E$ ; et ensuite, (iii)  $D = A + F$ . En substituant  $A + F$  à D dans (ii), on obtient (iv)  $B = A + F + E$ ; et par conséquent  $A = A + F + E + C$ . Une absurdité qui stipule que A est égal à lui-même augmenté d'autres choses !

Les lexicologues se sont toujours attaqués à ce problème; la chasse aux cercles vicieux dans les définitions lexicographiques est devenue un des passe-temps

8 En ce qui a trait aux décompositions sémantiques, nous nous appuyons plus particulièrement sur les travaux de Zholkovsky, de Boguslawski, d'Aprésjan et surtout de Wierzbicka, voir Wierzbicka 1972 et 1980.

favoris de tous les théoriciens de la lexicographie. Pourtant, on ne constate dans la pratique dictionnaire que peu de changements sous cet aspect.

Le DEC a comme un de ses objectifs principaux d'éliminer à jamais cet état des choses. Pour atteindre cet objectif, le lexicologue du DEC doit observer de la façon la plus stricte la Règle de décomposition, qui se formule comme suit :

|| La définition de la lexie vedette L doit être faite par des lexies  $L_1, L_2, \dots, L_n$  (en ce sens que  $\langle L \rangle = \langle L_1 \rangle + \langle L_2 \rangle + \dots + \langle L_n \rangle$ ) qui sont chacune SÉMANTIQUEMENT PLUS SIMPLE que L.

Nous disons que la lexie  $L_1$  est *sémantiquement plus simple* que la lexie L si et seulement si on a besoin de  $L_1$  comme définissant de L alors qu'on ne peut utiliser L pour définir  $L_1$ . Ainsi, PROFESSEUR est défini comme 'personne qui, [en tant que] profession, enseigne1 Y à Z...', de sorte que le définissant comprend les éléments – les sens de lexies – 'personne', ('en tant que'), 'profession', 'enseigner1', etc. Nous sommes persuadés qu'en les définissant à leur tour, nous n'aurons pas besoin d'avoir recours à 'professeur' et même qu'un tel recours s'avèrera impossible. En conclusion, les lexèmes PERSONNE, PROFESSION et ENSEIGNER1 sont considérés *sémantiquement plus simples* que PROFESSEUR.

Un deuxième exemple rendra encore plus claire cette notion fort importante « être *sémantiquement plus simple* [que...] ». Qu'est-ce qui est *sémantiquement plus simple*, 'homme' ou 'femme' (en tant qu'êtres sexués) ? Nous croyons, avec Wierzbicka (1972 : 44 *ssq.*), que le sens 'femme' peut être défini sans la mention de 'homme', à savoir, exclusivement par la 'capacité de donner naissance'; cependant, le sens de 'homme' ne peut se définir sans la mention de 'femme', puisque 'homme' doit être défini par la 'capacité de causer qu'une femme donne naissance'. Le sens de la lexie HOMMEII 'personne de sexe masculin' inclut ainsi le sens de la lexie FEMMEI 'personne de sexe féminin', mais pas l'inverse; FEMMEI est donc *sémantiquement plus simple* que HOMMEII.

Soulignons que la condition « être *sémantiquement plus simple* [que...] » ne doit absolument pas être interprétée comme PSYCHOLOGIQUEMENT plus simple ou plus disponible (du point de vue du locuteur). Ainsi, pour reprendre l'exemple ci-dessus, le lexème PROFESSION n'est nullement plus simple ou plus disponible psychologiquement que PROFESSEUR; bien au contraire, il est de loin plus complexe (plus abstrait, plus difficile à cerner). Il n'en reste pas moins que, pour la définition, c'est PROFESSION qui est sous-jacent à PROFESSEUR, et non l'inverse.

Nous ne voulons aucunement prétendre que, étant donné deux lexies  $L_1$  et  $L_2$  *sémantiquement liées*, il est toujours facile de déterminer laquelle est *sémantiquement plus simple*, donc répondre à la question de savoir si  $\langle L_1 \rangle \supset \langle L_2 \rangle$  ou si  $\langle L_2 \rangle \supset \langle L_1 \rangle$ . Tout d'abord, ce n'est pas une alternative stricte, car il se peut que  $\langle L_1 \rangle$  et  $\langle L_2 \rangle$  se trouvent en intersection :  $\langle L_2 \rangle \cap \langle L_1 \rangle \neq \Lambda$  [le symbole  $\Lambda$  représente l'ensemble vide]. Prenons, par exemple, TABLE et GUÉRIDON. Les dictionnai-

res consultés définissent 'guéridon' comme 'petite table ronde ayant un seul pied central, ...'; d'après eux, 'guéridon'  $\supset$  'table'. Cependant, dans un registre de meubles, on ne mettra pas  *cinq tables*  pour désigner quatre tables et un guéridon; on n'enverra pas quelqu'un chercher un document  *sur cette table près de la fenêtre* , s'il s'agit d'un guéridon, etc. Un guéridon n'est donc pas une table, et aucun de ces deux lexèmes n'est définissable par l'autre.

Et même, s'il y a inclusion entre deux lexies comparées, la direction de cette inclusion peut ne pas devenir évidente à la première étape de leur analyse. Ainsi, pour ASTRONOME et ASTRONOMIE, on peut bien dire que 'astronome' est 'personne qui [en tant que] profession fait de l'astronomie' (donc, ASTRONOME est défini par ASTRONOMIE) ou, à l'inverse, que 'astronomie' est 'science que font [en tant que] profession les astronomes' (ASTRONOMIE étant définie par ASTRONOME). Les deux affirmations sont factuellement correctes (aucune entorse à la langue); mais, du point de vue lexicographique, c'est seulement la première qui est bonne. En effet :

- 1) Si 'astronome' est 'personne qui [en tant que] profession fait de l'ASTRONOMIE', alors à la deuxième étape de l'analyse, 'astronomie' est définie comme 'science des corps célestes'; tout retour à 'astronome' est ainsi évité.
- 2) Si, par contre, 'astronomie' est définie comme 'science que font les ASTRONOMES), à la deuxième étape, on est forcé de définir 'astronome' comme 'personne qui [en tant que] profession fait l'étude des corps célestes'. Mais dans ce cas, on obtient, par substitution,

'astronomie' = 'science que font les personnes qui [en tant que] profession font l'étude des corps célestes'.

Par réduction évidente, 'science que font les personnes qui [en tant que] profession font l'étude des X' est tout simplement 'science des X'. Par conséquent, nous pouvons écrire

'astronomie' = 'science des corps célestes'.

On retrouve ainsi la première affirmation, c'est-à-dire que 'astronome' doit se définir comme 'personne qui [en tant que] profession fait de l'astronomie'.

La difficulté dans ce cas particulier est créée par le fait que la relation morphologique entre ASTRONOMIE et ASTRONOME est inverse par rapport à leur relation sémantique :  *astronome*   $\subset$   *astronom+ie*  [ASTRONOME est morphologiquement plus simple], alors que 'astronome'  $\supset$  'astronomie' [ASTRONOMIE est sémantiquement plus simple]. L'orientation de la dérivation morphologique complique donc l'interprétation sémantique correcte.

Il n'existe aucune méthode automatique garantie pour déterminer la simplicité sémantique relative; c'est seulement en procédant par essais successifs que le lexicologue peut aboutir à la bonne formulation de la définition recherchée. Néanmoins, malgré l'absence de techniques faciles à appliquer, l'adhésion à la notion « être sémantiquement plus simple [que...] » est tout à fait cruciale dans le DEC.

Si la règle ci-dessus est rigoureusement respectée, toute définition lexicographique représente une DÉCOMPOSITION – c'est-à-dire une configuration de composantes plus simples – du sens de la lexie définie. Trois conséquences importantes en découlent, dont la première est d'ordre pratique, et les deux autres sont d'ordre théorique.

- Il est interdit de définir par des synonymes : un synonyme n'assure pas une décomposition sémantique du défini. Ainsi, on ne peut définir RIPOSTER [*Elle riposta en éclatant d'un rire fou*] par 'faire une riposte' malgré la synonymie 'riposter1' = 'faire une riposte', parce que 'faire une riposte' n'est pas une décomposition de 'riposter1' : le verbe FAIRE est vide ici (= un verbe support), et 'riposte' est sémantiquement égal à 'riposter1'. Cependant, cette interdiction n'exclut pas l'emploi d'un synonyme moins riche de L comme composante générique dans la définition de L. Par exemple, RIPOSTER1 [*La personne X riposte à la personne Y par propos ou gestes Z*] se définit comme 'X répondI.2 aux propos ou gestes de Y – que X croitII adressés à X et lui portant atteinte – par des propos ou gestes Z que X adresse à Y dans le but de porter atteinte à Y'. Cette définition couvre au moins les emplois suivants :

- (17) a. *À leur proposition de se rendre, le maréchal riposta par un mot devenu célèbre.*  
 b. *– Va ranger ta chambre, dit maman, et ne riposte pas !*  
 c. *– Riposter à un policier par un bras d'honneur, c'est de l'audace !*

On voit bien que 'riposter1' est un cas particulier de 'répondreI.2' et que ce dernier est donc un synonyme moins riche du premier; donc l'emploi de 'répondreI.2' dans la définition de 'riposter1' est plus que justifié : c'est en fait la seule possibilité.

- Les cercles vicieux, la défaillance la plus redoutée, critiquée et ridiculisée du lexicologue, sont automatiquement éliminés par la décomposition.
- Les décompositions du sens des lexies d'une langue, poussées aussi loin que possible, amènent inévitablement aux primitifs sémantiques, le rêve absolu du lexicologue. En effet, la décomposition sémantique n'utilise, à chaque étape, que des éléments plus simples que ceux à définir, de sorte qu'au niveau  $n + 1$  de décomposition tous les éléments sémantiques sont plus simples que les éléments du niveau  $n$ . D'autre part, le nombre d'éléments sémantiques disponibles pour la définition (c'est-à-dire le nombre de significations des lexies de la langue en question) est FINI : tout simplement parce que le stock lexical d'une langue est fini. Il en découle qu'en poursuivant les décompositions assez longtemps, le lexicologue se retrouve infailliblement avec un résidu irréductible – des significations lexicales indéfinissables par d'autres significations lexicales de la même langue (il n'en reste aucune de plus simple). Ces éléments indéfinissables sont des PRIMITIFS SÉMANTIQUES (cf. Mel'čuk 1989).

Il n'est pas dans notre intention de développer ici toute la question fascinante des primitifs sémantiques, et nous nous limiterons aux trois remarques suivantes.

Premièrement, l'étude moderne des primitifs sémantiques a été entamée par A. Wierzbicka; la consultation de ses ouvrages est donc indispensable : Wierzbicka 1972, 1980, Goddard & Wierzbicka 1994. Le lecteur intéressé par la question des primitifs sémantiques y trouvera une discussion bien argumentée.

Deuxièmement, l'affirmation « Les primitifs sémantiques ne sont pas définissables » ne doit pas être prise au sens général; ils sont indéfinissables seulement linguistiquement, c'est-à-dire par des éléments sémantiques de la même langue. Ainsi, le sens de 'ne... pas' (= négation) est en toute vraisemblance un primitif en français : il nous semble impossible de définir le sens de l'expression lexicale NE... PAS par d'autres expressions lexicales plus simples du français. Pourtant, en logique, la négation est définie sans difficulté :

|| La négation  $\neg$  est une opération telle que si A est une proposition vraie, alors  $\neg A$  est une proposition fausse, et vice versa.

C'est une définition parfaite; néanmoins elle ne peut figurer dans un dictionnaire de français au même titre que les définitions formulées antérieurement. D'autres exemples de définitions dépassant le cadre linguistique pourraient être celles de 'eau' comme ' $H_2O$ ', de 'lumière' comme 'ondes électromagnétiques ayant une fréquence données', celle de 'chat' comme '*felis felis*', etc. De telles définitions se rapportent à la chose dénotée, c'est-à-dire au référent, et non pas au sens du signe linguistique.

Troisièmement, les primitifs sémantiques définis comme nous l'avons fait ci-dessus sont spécifiques à la langue; en fait, on doit parler des primitifs sémantiques du français, du chinois, du swahili, ... Cela n'exclut en rien l'identité au moins partielle des ensembles de primitifs sémantiques pour des langues particulières, problème que nous ne creuserons pas davantage.

**NB :** Nous aimerions souligner ici que les primitifs sémantiques dont nous parlons ne correspondent pas infailliblement ou nécessairement à ce qu'on appelle des primitifs de la pensée humaine. Nous nous sentons tout à fait incompetents pour traiter des primitifs de comportement, de perception, de réflexion, etc. et de leurs liens avec le sémantisme des lexies. Notre seul objectif est de bien trancher entre les primitifs psychologiques et nos primitifs sémantiques.

### Règle 3 : Règle du bloc maximal

Si l'application de la règle de décomposition fait que le MINIMUM d'éléments qu'une définition contienne est deux<sup>9</sup>, on peut se demander quel en est le MAXIMUM. Le

9 Sauf, bien entendu, des renvois à des synonymes absolus : BICYCLETTE = 'vélo'.

nombre maximal d'éléments dans un définissant dépend, de façon évidente, de la profondeur de la décomposition. Si celle-ci est la plus profonde possible, chaque définition sera obligatoirement composée de primitifs sémantiques. Cela correspond à un idéal scientifique : la description des sens des lexies d'une langue se fait par un métalangage réduit, unifié et totalement contrôlable, métalangage dont les éléments lexicaux ne sont que des primitifs. Pourtant, dans le cadre plus étroit de notre tâche spécifique, celle de fournir un dictionnaire de la langue, cette démarche – décomposition maximale – a au moins trois inconvénients.

- 1) Il faut d'abord disposer de l'ensemble bien établi de primitifs, et un tel ensemble n'est pas encore disponible (même Wierzbicka a modifié à quelques reprises ses hypothèses de départ et continue à développer l'inventaire des primitifs : de 13 à 55 en 1994). Avant de rédiger un dictionnaire, on serait donc forcé d'accomplir une tâche préalable, celle de déterminer les primitifs sémantiques.
- 2) Une définition formulée à l'aide seulement de primitifs est très longue et extrêmement complexe, ce qui la rend difficile à manier autant par l'auteur que par le lecteur. De plus, une telle définition rend l'évaluation intuitive par des locuteurs plus que délicate. (Par exemple, les définitions de certains lexèmes dans Wierzbicka 1985 atteignent deux pages – et cela, avec l'utilisation de plusieurs éléments intermédiaires, c'est-à-dire non réduits aux primitifs. Ainsi, la définition de BIRDS occupe deux pages [180-181], tout comme celle de LEMONS [310-311] et beaucoup d'autres.)
- 3) Avec la décomposition maximale, les rapports sémantiques entre les lexies ne sont pas directement visibles dans leur définition; par exemple, on ne verrait pas que le sens 'professeur' inclut celui de 'enseigner1', car 'enseigner1' serait remplacé par des primitifs.

Pour pallier tous ces inconvénients, nous posons la notion de *bloc maximal*. Elle implique que la décomposition doit être la moins profonde possible, c'est-à-dire que toute définition lexicographique doit utiliser les éléments définissants LES PLUS ENGLOBANTS disponibles dans la langue. De façon formelle, si l'on a la définition (i)  $A = B + C + D + E$  et que la langue possède une lexie F telle que  $F = B + C$ , alors la définition (i) doit être réécrite comme (ii)  $A = F + D + E$ ; F forme ici le bloc maximal.

La Règle du bloc maximal se lit alors comme suit :

|| La définition de la lexie vedette L doit être faite par des lexies  $L_1, L_2, \dots, L_n$  telles qu'aucune configuration de ces lexies ne peut être remplacée par une seule lexie sémantiquement équivalente.

Ainsi, la définition (i) n'est pas acceptable puisqu'elle ne respecte pas la Règle de bloc maximal : la configuration  $B + C$  peut être remplacée par F; en effet, la règle en cause stipule que si, dans une définition, une configuration lexicale peut être remplacée par une seule lexie, un tel remplacement est obligatoire.

Comme on peut le voir, cette règle impose une décomposition strictement graduelle : un sens est toujours représenté par des « constituants sémantiques IMMÉDIATS ».

Reprenons notre exemple de PROFESSEUR. Comparons deux définitions équivalentes : 1) la définition citée plus haut, qui est conforme à la Règle du bloc maximal, et 2) une nouvelle définition, qui ne la respecte pas :

- 1) 'personne qui<sup>1</sup> en tant que<sup>1</sup> profession ENSEIGNE1 à Z Y d'un niveau assez élevé (étant employé par un établissement d'enseignement W)';
- 2) 'personne qui<sup>1</sup> en tant que<sup>1</sup> profession CAUSE VOLONTAIREMENT QUE Z APPRENNE<sup>III.1b</sup> Y, EN TRANSMETTANT SYSTÉMATIQUEMENT À Z DES CONNAISSANCES (PORTANT SUR) Y OU DES TECHNIQUES (PORTANT SUR) Y, Y étant d'un niveau assez élevé (cette personne étant employée par un établissement d'enseignement W)'

La seule différence entre la première et la seconde définition est que, dans la seconde, le sens 'enseigner1' est décomposé :

(X enseigne1 Y à Z) = (X cause volontairement que Z apprenne<sup>III.1b</sup> Y, en transmettant systématiquement à Z des connaissances (portant sur) Y ou des techniques (portant sur) Y).

Ce n'est que la toute première étape de décomposition d'enseigner1; en effet, en rejetant la Règle du bloc maximal, on se verrait obligé de décomposer de la même façon 'volontairement', 'apprendre<sup>III.1b</sup>' et tous les autres éléments du définissant d'ENSEIGNER1. Et ce n'est pas encore tout : les mêmes considérations sont valables pour PERSONNE et PROFESSION dans le définissant de PROFESSEUR. La décomposition complète, jusqu'au niveau des primitifs, donnerait une définition très longue et extrêmement complexe; en réalité, une telle définition ne serait pas maniable : l'intuition linguistique serait entièrement détraquée. La psychologie a montré que l'intelligence humaine procède, pour la résolution de n'importe quelle tâche, en divisant cette tâche en blocs majeurs, subdivisés à leur tour, etc., c'est-à-dire en échelonnant les difficultés. En introduisant la Règle de bloc maximal, nous ne faisons que suivre, dans ce cas particulier d'élaboration de définitions lexicographiques, la tendance générale à diviser les problèmes en unités maniables (cf. le précepte 2 de Descartes<sup>10</sup>).

Maintenant, pour ne pas succomber à l'arbitraire dans une décomposition par étapes, nous sommes obligés de décomposer LE MOINS POSSIBLE (la décomposition maximale amenant fatalement aux primitifs, et une décomposition entre les deux introduisant l'arbitraire). En d'autres termes, les définitions du DEC doivent être les moins profondes possibles ! Mais comment le lexicologue détermine-t-il la profondeur minimale souhaitable dans chaque cas ? Ses décisions seront justi-

10 «Le second [précepte est] de diviser chacune des difficultés que j'examinerais en autant de parcelles qu'il se pourrait et qu'il serait requis pour les mieux résoudre» [«Discours de la méthode», p. 138, *Œuvres et Lettres*, 1953, Paris : Gallimard, Bibliothèque de la Pléiade].

fiées par le lexique de la langue  $\mathcal{L}$  qu'il décrit : il doit choisir pour son définissant les lexies les plus spécifiques au sens le plus englobant qui existent en  $\mathcal{L}$ .

Il faut signaler avec force que l'application de la Règle de bloc maximal ne change en rien le contenu de la définition et n'affecte point sa logique, puisqu'il s'agit de substitutions équivalentes : le bloc maximal choisi est par définition égal à sa décomposition. Cependant, l'utilisation systématique de blocs maximaux entraîne un gain didactique notable, à savoir, une simplification formelle, parfois très importante, de la définition. Une présentation simplifiée de la définition assure une bien meilleure perception de la structure interne du sens, permettant, par là même, une mise en action immédiate de l'intuition linguistique et de la logique subconsciente du lexicologue. Au contraire, son intuition linguistique, qui est fondamentale dans notre approche comme source principale des données lexicographiques, peut être complètement bloquée par une décomposition trop morcelée. (Rappelons qu'en mathématiques, les transformations équivalentes, qui ne font que mettre une expression sous une autre forme, sont souvent indispensables pour la solution des problèmes.)

#### Règle 4 : Règle de standardisation

Étant donné la nature formelle et la logique rigoureuse du DEC, il va de soi que la définition des lexies vedettes doit être faite dans un métalangage uniformisé, soumis à des contraintes explicites, applicables de façon homogène à l'ensemble du lexique. De façon plus formelle, cette exigence peut être exprimée par la règle suivante comprenant deux contraintes :

- || La définition doit être faite de façon à éviter, dans le définissant, a) les termes AMBIGUS et b) les termes SYNONYMES.

La première contrainte signifie que chaque élément utilisé dans un définissant doit véhiculer toujours un seul et même sens : il doit être le sens d'une lexie bien définie. Quant à la seconde, elle impose que chaque sens utilisé dans un définissant est toujours exprimé par un seul et même élément, c'est-à-dire qu'il est exprimé dans tous les définissants par la même lexie.

Pour respecter la première contrainte, celle de la « non-ambiguïté », il faut désambiguïser tous les éléments lexicaux utilisés dans les définitions en les munissant de *numéros lexicaux distinctifs*, qui en spécifient l'acception voulue. L'utilisation de numéros distinctifs est une pratique universelle des dictionnaires; cependant, aucun ne les utilise dans ses propres définitions, ce qui rend celles-ci souvent extrêmement ambiguës. Il est vrai que, pour un locuteur natif, cette ambiguïté peut être facilement levée par le contexte de la définition et surtout par les exemples. Mais cela est déjà beaucoup moins évident pour un étranger et devient tout à fait insoluble pour un dispositif logique, tel un ordinateur. Et comme le DEC est un dictionnaire qui ne fait aucunement appel à l'intuition hasardeuse d'un utilisateur, il n'est pas possible de tolérer de pareilles ambiguïtés.



Le DEC utilise donc des numéros lexicaux, qui reflètent les DISTANCES SÉMANTIQUES entre les lexies. La distance sémantique entre les lexies  $L_1$  et  $L_2$  est mesurée par deux paramètres considérés ensemble :

- la TAILLE de la composante sémantique commune à  $L_1$  et  $L_2$  (plus cette taille est grande, plus  $L_1$  et  $L_2$  sont proches);
- la RÉGULARITÉ de la distinction sémantique entre  $L_1$  et  $L_2$  (plus élevé est le nombre de paires de lexies où la même distinction apparaît dans la langue en question, plus  $L_1$  et  $L_2$  sont proches).

Dans cette étude, nous utilisons QUATRE niveaux de numéros lexicaux (nous ne fournissons ici qu'une caractérisation fort approximative; voir encore Ch. IV, 1.4, p. 166) :

- des exposants numériques pour les lexies homonymes : VOLER<sup>1</sup> ('se déplacer dans l'air') ~ VOLER<sup>2</sup> ('prendre de façon subreptice et illégitime'). De telles lexies n'ont aucun sémantisme en commun et appartiennent à des vocables différents;
- des chiffres romains pour les lexies dont la partie sémantique commune est assez forte mais dont la distinction sémantique n'est pas régulière en langue : COMBLERI (L'eau comble le caniveau) ~ COMBLERII (La joie le comble);
- des chiffres arabes pour les lexies dont la partie sémantique commune est forte et dont la distinction sémantique est assez régulière en langue : COMBLERII.1 (La joie comblait son cœur) ~ COMBLERII.2 (Ses soupirs comblaient les vides dans la conversation);
- des minuscules latines pour les lexies dont la partie sémantique commune est très forte et dont la distinction sémantique est très régulière en langue : COMBLERI.a (L'eau comble le caniveau) ~ COMBLERI.b (Jean comble le caniveau d'eau).

Si les exposants sont utilisés pour marquer l'homonymie, les autres types de numéros lexicaux indiquent des degrés différents de polysémie, en distinguant les lexies d'un même vocable.

Pour illustrer le problème de l'ambiguïté dans les définitions, examinons une définition courante du lexème HAUTEURI.1 (*la hauteur d'une tour*). Relevée telle quelle dans le *Petit Robert*, la définition de HAUTEURI.1 se lit comme suit :

HAUTEURI.1 = 'dimension dans le sens vertical, de la base au sommet'<sup>11</sup>

Cette définition est d'une ambiguïté multiple, étant donné que DIMENSION a, dans le même dictionnaire, 6 acceptions, SENS en a 3, BASE, 11 et SOMMET, 3 (l'adjectif VERTICAL n'est pas ambigu); cette définition pourrait donc s'interpréter de  $(6 \times 3 \times 11 \times 3 =) 594$  façons différentes !

<sup>11</sup> Cette définition n'est pas tout à fait satisfaisante : par exemple, une table ou une voiture, tout en ayant une hauteur, n'ont ni base ni sommet. Cependant, ce détail n'est pas pertinent ici.

Selon notre démarche, cette définition devrait être réécrite comme suit :

HAUTEUR<sub>I.1</sub> = 'dimension<sub>I.2</sub> dans le sens<sup>2</sup> vertical, de la base<sub>I.1</sub> au sommet<sup>1</sup>'.

Les numéros lexicaux distinctifs sont empruntés toujours au même dictionnaire – qui les a bien notés, mais qui ne les utilise pas dans ses définitions (voir Rey 1990 : 52).

La présentation de définitions avec les éléments du définissant désambiguïsés permet une substitution automatique de ces éléments par leur propre définition – en vue d'une vérification formelle. Même pour un utilisateur ordinaire, cette façon de rédiger la définition est d'une grande aide : elle lui facilite de façon indiscutable la recherche du sens exact de la lexie vedette. En plus, nous croyons que cette écriture de définition ne gêne en rien (ou, en tout cas, pas trop) la lisibilité par des humains. Pour un ordinateur, c'est le seul choix possible.

Avouons qu'en pratique, lors de la rédaction du DEC, nous sommes obligés de surseoir de nombreuses fois à l'application stricte de cette contrainte – pour la simple raison que nous devons souvent utiliser en tant qu'éléments des définissants des lexies que nous n'avons pas encore traitées. (Les lexies traitées, elles, sont obligatoirement utilisées avec leur numéro distinctif.) Dans des cas où une distinction sémantique entre lexies non traitées s'avère fondamentalement pertinente, la polysémie pouvant créer une confusion fatale, nous utilisons des numéros provisoires, que nous empruntons à un bon dictionnaire (par exemple, au *Nouveau PR* pour le français). Au fur et à mesure de la progression de nos travaux, cette mesure transitoire sera de moins en moins nécessaire et finira par être totalement éliminée.

Pour respecter la seconde contrainte de la Règle 4, celle de la « non-synonymie », il faut déterminer, pour chaque sens utilisé dans une définition, une et une seule lexie qui devra l'exprimer dans toutes les définitions. Cependant, il est encore plus difficile de reconnaître des sens identiques sous des formes bien différentes que des termes identiques ayant des sens différents; la seconde contrainte crée donc encore plus de problèmes sérieux que la première. Par exemple, MONTRE est souvent défini comme 'dispositif POUR savoir l'heure', MARTEAU, comme 'outil DE percussion', COUTEAU, comme 'instrument SERVANT À couper', CUILLÈRE, comme 'ustensile QUI SERT À porter la nourriture à la bouche'. Il n'est pas immédiatement évident que *pour*, *de* et *servant* (*qui sert*) expriment le même sens dans le contexte des définitions données. Ce sens, une fois établi, doit être spécifié par une seule expression, qui sera la seule utilisée dorénavant. Le choix de cette expression est soumis, à son tour, aux trois conditions suivantes :

- 1) L'expression retenue doit être la MOINS AMBIGUË POSSIBLE pour permettre le fonctionnement harmonieux de l'intuition linguistique du lexicologue. Ainsi, dans la définition de MONTRE ci-dessus, le terme POUR, qui exprime bien le sens voulu, n'est quand même pas à retenir car il est trop ambigu. Même désambiguïsé par un numéro lexical, il ne devient pas beaucoup plus facile à traiter : tout locuteur aura des difficultés à identifier le sens exact de POUR<sub>II.3</sub>.

La même remarque s'applique à DE dans la définition de MARTEAU, à un degré plus élevé encore. Quant à SERVANT À/QUI SERT À, bien que moins ambiguës lexicalement, ces expressions ont une forme trop polyvalente : elles signifient à la fois 'servant actuellement', 'pouvant servir' et 'destiné à servir'.

- 2) L'expression retenue doit être la MOINS IDIOMATIQUE POSSIBLE, c'est-à-dire qu'elle ne doit pas comporter de nuances supplémentaires.
- 3) L'expression retenue doit posséder la PLUS GRANDE MOBILITÉ SYNTAXIQUE POSSIBLE : elle doit être utilisable dans le plus grand nombre de contextes.

Par conséquent, pour les quatre lexies citées, comme elles sont toutes des noms d'artefacts, nous préférons l'expression *destiné à*; les définitions reformulées se liront donc comme suit :

MONTRE	=	'dispositif	DESTINÉ à montrer l'heure...'
MARTEAU	=	'outil	DESTINÉ à frapper...'
COUTEAU	=	'instrument	DESTINÉ à couper...'
CUILLÈRE	=	'ustensile	DESTINÉ à porter...'

En effet, l'expression *destiné à* est nettement moins ambiguë que ses concurrents; elle nous paraît plus analytique et moins idiomatique que, par exemple, les prépositions; et elle est, en syntaxe, extrêmement souple, puisqu'elle peut régir autant des verbes (*destiné à subir...*) que des noms (*destiné au chauffage*) et des phrases (*destiné à ce que X fasse...*).

Remarque sur la langue de la définition dans le DEC (« métafrançais »)

Il faut absolument signaler ici que le respect rigoureux de la contrainte « une expression pour un sens donné » peut provoquer des violations de cooccurrences lexicales de la langue décrite et, par conséquent, produire des définitions qui ne sont pas irréprochables du point de vue de la qualité de la langue. Comme on le sait, la lexicographie traditionnelle récuse toute définition non conforme aux canons stylistiques, ceci étant un de ses principes fondamentaux inviolables. Nous suivons aussi cette pratique dans tous les cas où il est possible de le faire sans entrer en conflit avec les exigences de la standardisation présentées ci-dessus, mais nous donnons préséance à la cohérence logique de la définition plutôt qu'à l'esthétique linguistique de celle-ci. On peut le faire puisque la langue de définition est une MÉTALANGUE par rapport à la langue décrite et une métalangue ne peut éviter d'être distincte de la langue objet. Par conséquent, même si les définitions du DEC semblent être écrites en français, elles sont rédigées en fait en « métafrançais », c'est-à-dire dans une langue lexicographique ARTIFICIELLE créée dans ce but. Nous nous devons de signaler que nous ne nous écartons pas sciemment du français « normal » sans nécessité, mais nous ne nous refusons pas de le faire en cas de besoin. Ainsi, pour traiter le concept de causation, qui est nécessaire dans la définition de la plupart des verbes transitifs français, nous utilisons l'élément 'causer que', qui enfreint les normes habituelles. Soit les deux phrases suivantes :

- (18) a. *La lumière a baissé.*  
 b. *André a baissé la lumière.*

En (18a), on a *X baisse* = 'X devient plus faible'; en (18b), *BAISSER* devra être défini comme *Y baisse* *X* = 'Y cause que X baisse'. Même si notre choix de 'causer que' dérange quelque peu le lecteur, nous croyons que cette expression est de loin préférable à, par exemple, 'faire en sorte que'. Cette dernière expression est, d'une part, trop idiomatique et, d'autre part, ne passe pas dans tous les emplois. Par exemple, considérons la phrase *Le divorce a beaucoup changé Pierre*; sa description sémantique 'Le divorce a fait en sorte que Pierre a beaucoup changé' n'est pas du tout plus acceptable que 'Le divorce a causé que Pierre a beaucoup changé'.

Une langue de définition standard et uniformisée a toujours été un objectif convoité par le lexicologue/lexicographe; il n'en est pas autrement pour nous. Cependant, notre démarche est radicalement différente de l'approche habituelle. Dans la lexicographie, on connaît des tentatives de réduire la langue de définition le plus possible. On la construit avec un nombre limité de lexies déterminées par diverses considérations : fréquence, disponibilité, primitivité sémantique, etc. Ce nombre oscille entre une trentaine de primitifs sémantiques chez Wierzbicka, environ 1600 « mots fondamentaux » dans Gougenheim 1961 et 2130 dans *Longman Dictionary of Contemporary English* (1987). Nous, par contre, nous pensons exploiter dans les définitions toute la richesse lexicale de la langue; c'est-à-dire que pour nous la langue de définition, bien qu'uniformisée, n'est aucunement minimale. Cela ne veut pourtant pas dire que notre langue de définition soit complètement identique au français normal. C'est du français – mais la Règle de standardisation en fait du français « traité », qui, par rapport au français normal, montre les quatre caractéristiques suivantes :

- 1) toutes les lexies ambiguës utilisées dans les définitions sont désambiguïsées à l'aide de numéros distinctifs, excluant ainsi de la langue de définition les expressions ambiguës;
- 2) en principe, un sens est noté par une même expression sélectionnée par le lexicologue, excluant ainsi de la langue de définition les expressions synonymes;
- 3) certaines expressions ont vu leur sens et/ou leur syntaxe légèrement modifiés par le lexicologue – pour mieux les adapter aux besoins sémantiques et logiques (cf. *causer que* ci-dessus), le nombre de telles composantes « traitées » étant quand même fort limité;
- 4) la cooccurrence lexicale normative du français n'est pas nécessairement respectée dans la langue de définition.

### Règle 5 : Règle de substituabilité

Cette règle d'élaboration des définitions du DEC est, à notre avis, la plus fondamentale. En effet, le DEC se base sur la thèse, d'ailleurs universellement

admise, que la lexie et sa définition, c'est-à-dire le défini et le définissant, sont rigoureusement égaux dans leur « substance » sémantique. La seule façon de vérifier cette égalité est, comme on le sait depuis l'Antiquité, la substitution réciproque du défini et du définissant dans un contexte. D'où la Règle de substituabilité suivante :

- || La lexie vedette L et sa définition (L) doivent être réciproquement substituables dans tous les énoncés, sans modification du sens exprimé.

La Règle de substituabilité n'est qu'une reformulation du principe vaillamment défendu par A. Wierzbicka depuis des décennies et qu'elle a érigé en principe de base de toute étude sémantique. Nous souscrivons entièrement à cette prise de position – de sorte que toutes les définitions du DEC sont testées par la substitution du défini et du définissant dans les deux directions : le défini doit admettre le remplacement par le définissant, et le définissant, par le défini.

De telles substitutions détruisent quasi inévitablement le caractère naturel du texte (cf. ci-dessus les remarques au sujet du métafrançais), mais il y a toujours PRÉSERVATION ABSOLUE DU SENS : aucun changement de nuances n'est admis. Si un changement de sens a lieu, la définition est inexacte. Reprenons notre exemple de PROFESSEUR, en donnant, cette fois-ci, la définition complète :

*professeur de Y/de Z (à W) = 'personne qui, [en tant que] profession, enseigne1 à Z Y d'un niveau assez élevé (étant employé par un établissement d'enseignement W)*

*[professeur de math (au lycée Victor-Hugo); notre professeur de math]*

La composante de la définition entre parenthèses est une *composante optionnelle* (voir plus loin), c'est-à-dire qu'elle n'est pas nécessairement toujours actualisée. En effet, un professeur n'est pas nécessairement employé par un établissement : par exemple, un *professeur de langue (de musique)* peut bien être son propre employeur et exercer à domicile.

Si, dans la phrase (19a)

(19) a. *Enfin, le nouveau professeur de chimie est arrivé.*

on substitue à PROFESSEUR le définissant proposé, on obtient (19b) :

b. *Enfin, la nouvelle personne qui, [en tant que] profession, enseigne1 la chimie (étant employée par un établissement d'enseignement) est arrivée.*

La phrase (19b) n'est pas tout à fait normale – en ce sens qu'aucun francophone ne l'emploierait spontanément. Cependant, (19a) et (19b) sont rigoureusement synonymes : la première véhicule exactement les mêmes informations que la seconde. Nous voyons donc que dans l'exemple (19a) notre défini est parfaitement remplaçable par le définissant.

Procédons maintenant à la substitution inverse. Soit la phrase (19c) :

c. *André est une personne qui, [en tant que] profession, enseigne1 la chimie de façon excellente (étant employé par un établissement d'enseignement).*

Si nous remplaçons la partie soulignée de (19c), qui correspond au définissant de PROFESSEUR, par le défini, nous obtenons

d. *André est un professeur de chimie excellent.*

Ces substitutions réciproques montrent qu'au moins pour les contextes examinés, notre définition de PROFESSEUR est bonne.

**NB :** Signalons que les substitutions dont il est question se font en effectuant en même temps les adaptations syntaxiques nécessaires : accords pertinents, changements de servitudes grammaticales, etc.

Le test de substitution doit vérifier, pour une définition donnée, la NÉCESSITÉ de chaque composante du définissant et la SUFFISANCE de l'ensemble de ses composantes. (Cette façon de faire est, bien entendu, similaire aux conditions nécessaires et suffisantes qu'on utilise en mathématiques.) Examinons d'abord la condition de nécessité.

Cette condition stipule qu'aucune composante ne peut figurer dans le définissant de L si sa nécessité n'est pas justifiable, les justifications possibles étant des trois types suivants :

- La composante ('σ) est justifiable, 1) si sans elle une mauvaise substitution devient possible ou bien une bonne substitution devient impossible et 2) si sans elle une description exhaustive de la cooccurrence de L n'est pas possible.

Ainsi, la composante ('Y] d'un niveau assez élevé) est nécessaire, puisque sans elle le définissant de PROFESSEUR permettrait de dire *\*professeur d'école maternelle* [= *maître / maîtresse*] ou *\*professeur de conduite automobile* [= *moniteur*]. Avec cette composante, de telles expressions sont rejetées, puisque les «matières» de la maternelle ou la conduite automobile ne sont pas considérées comme étant 'd'un niveau assez élevé'. La composante ('étant employé par un établissement d'enseignement'), même si elle est optionnelle et peut donc être supprimée dans certains contextes, est cependant nécessaire pour assurer la cooccurrence du lexème PROFESSEUR dans des contextes comme *professeur à (\*de, \*avec) l'Université Lumière*, avec les expressions *donner des cours, faire les examens, noter les travaux, participer à des jurys*, etc.

- La composante ('σ) est justifiable, si elle n'est pas redondante, c'est-à-dire si elle n'est pas déjà implicitement incluse dans une autre composante.

Ainsi, on pourrait penser que, dans le définissant de PROFESSEUR, Y devrait être étoffé par la composante 'connaissances et/ou aptitudes'; cependant, 'enseigner1' contient déjà cette composante, puisqu'on n'enseigne que des connaissances et/ou des aptitudes. Notons que, si dans le cas précédant, l'absence d'une composante nécessaire mène à des erreurs langagières, dans ce cas-ci, la présence d'une composante redondante (= non nécessaire) n'affecte en rien l'exactitude de la définition. Une telle composante n'est à la rigueur qu'encombrante. (Une composante redondante peut même être parfois pédagogiquement utile, en facilitant la compréhension de la définition.)

- La composante (σ) est justifiable, si elle est nécessaire pour une autre composante (justifiée indépendamment) en tant que son actant sémantique. Ainsi, dans le définissant de PROFESSEUR, la composante 'enseigner<sup>1</sup>' exige, d'après sa propre définition, trois actants sémantiques : l'ASém 1 est 'personne', l'ASém 2 est 'Y', et l'ASém 3 est 'à Z'. La définition de PROFESSEUR doit donc comprendre au moins ces trois ASém.

**NB :** La variable Z ne peut être réalisée à la surface que par des noms référentiels, désignant des personnes spécifiques. Ainsi, on ne dit pas \**professeur de jeunes filles* / \**d'adultes*, alors que les expressions comme *le professeur de Jean* / *mon professeur* sont tout à fait courantes.

Quant à la condition de suffisance, la situation est radicalement différente. Un seul exemple bien choisi prouve la nécessité d'une composante donnée, alors que, pour prouver la suffisance de l'ensemble des composantes proposées, le lexicologue doit s'assurer de l'INEXISTENCE de contre-exemples – ce qui est, bien entendu, impossible. L'analyse de nombreux exemples tirés de domaines variés était le définissant en cause sans jamais prouver, au sens strict, son exactitude.

Ainsi, même si l'exemple (19) corrobore la définition de PROFESSEUR ci-dessus (p. 91), il ne constitue aucunement une preuve logique de sa validité; bien plus, aucun exemple et même aucun ensemble d'exemples ne constitue une telle preuve ! En effet, la substituabilité (à cause de la condition de suffisance) doit tenir pour TOUTS les contextes, alors qu'il est impossible de les vérifier tous. Un seul contre-exemple PROUVE l'inexactitude d'une définition lexicographique, mais des milliers d'exemples positifs ne prouvent rien (même si de nombreux exemples positifs créent une certitude). Par conséquent, le lexicologue, ayant vérifié un maximum de contextes, doit se satisfaire d'une probabilité forte de la justesse de sa définition – plutôt que de viser la vérité logique. Sa tâche est bien remplie s'il démontre que toutes les composantes de son définissant sont vraiment nécessaires; cependant, il n'a aucune possibilité de démontrer logiquement que toutes les composantes nécessaires y sont en fait présentes.

Néanmoins, la vérification de la substituabilité réciproque du défini et du définissant dans le plus grand nombre de contextes reste la démarche centrale de la recherche lexicographique. Toutes les définitions du DEC doivent se plier au test de substitution. Cette vérification est incontournable, et toute définition qui ne satisfait pas à ce test doit être modifiée. Ainsi, contrairement à PROFESSEUR, qui peut être défini comme 'PERSONNE qui... ' (voir (20a)), ÉTUDIANT ne peut être défini par 'PERSONNE qui... ' (comme le font les dictionnaires usuels) à cause de (20b) :

- (20) a. 'Jeanne est une personne qui, [en tant que] profession, enseigne<sup>1</sup> bien les maths (étant employée dans un établissement d'enseignement)' ⇔ *Jeanne est un bon professeur de math.*
- b. 'Jeanne est une personne qui étudie bien les maths dans un établissement d'enseignement supérieur' ⇔ *Jeanne est un bon \*étudiant en math.*

L'exemple (20b) démontre immédiatement ce qui viole la Règle de substitution dans la définition présumée d'ÉTUDIANT : la composante sémantique 'personne' n'est pas suffisante. Alors que la lexie PROFESSEUR s'applique aux hommes et aux femmes également (donc inclut la composante 'personne' sans distinction de sexe), la lexie ÉTUDIANT (au singulier) ne s'applique qu'aux hommes; par conséquent, la composante générique de sa définition est 'personne de sexe masculin' plutôt que 'personne' tout court. La définition refaite d'ÉTUDIANT doit donc se lire comme suit :

*étudiant en Y de Z* = 'PERSONNE DE SEXE MASCULIN qui étudie la matière Y dans un établissement d'enseignement supérieur Z'  
(*étudiant en math de l'université de Rennes*).

Et parallèlement, pour les femmes, il faut avoir une autre définition :

*étudiante en Y de Z* = 'PERSONNE DE SEXE FÉMININ qui étudie la matière Y dans un établissement d'enseignement supérieur Z'  
(*étudiante en math de l'université de Rennes*).

Cette dernière définition est tout à fait compatible avec (20b).

Ces définitions entraînent les trois remarques suivantes.

1. On pourrait penser qu'en vertu de la Règle du bloc maximal (= Règle 3, p. 84), il faille remplacer, dans la définition du lexème ÉTUDIANT, la composante 'personne de sexe masculin' par 'homme'. Or cette substitution n'est pas équivalente car la composante 'de sexe masculin' est faible dans 'étudiant' et ne l'est pas dans 'homme' (voir Remarque 3 ci-dessous pour le concept de composante faible). Cette différence de statut se manifeste de la façon suivante : le pluriel *les étudiants* peut désigner un groupe incluant des femmes, le pluriel *les hommes* ne le peut pas. L'application de la Règle du bloc maximal n'est donc pas possible. Le cas du lexème ÉTUDIANTE est tout à fait différent car la substitution 'personne de sexe féminin' ↔ 'femme' est vraiment équivalente : le pluriel *les étudiantes* ne désigne que des femmes. La Règle du bloc maximal doit donc être appliquée. (Si nous ne l'avons pas fait dans la définition de ÉTUDIANTE présentée ci-dessus, c'est uniquement afin de faciliter la comparaison avec la définition de ÉTUDIANT).
2. La forme *étudiante*, qu'on devrait utiliser dans (20b), représente une lexie autonome ÉTUDIANTE. Cela n'oblige cependant en rien à faire un traitement à part de cette lexie dans le dictionnaire : le français possède un moyen régulier dérivationnel (= morphologique) pour exprimer le sexe féminin; il suffit donc de marquer la lexie de base ÉTUDIANT comme admettant cette dérivation (tout à fait comme pour PARISIEN ~ PARISIENNE, LION ~ LIONNE, ÂNE ~ ÂNESSE, etc.). Lors de la dérivation, le suffixe féminin remplace la composante 'personne de sexe masculin' par 'personne de sexe féminin' = 'femme' dans la définition de la lexie de départ.



3. Le fait que le pluriel d'ÉTUDIANT s'applique aussi à un groupe mixte de personnes des deux sexes doit être pris en charge par la définition. Pour le faire, il faut spécifier, dans l'article de dictionnaire ÉTUDIANT, que la pluralisation de la lexie change sa définition, en éliminant la composante 'de sexe masculin'. Cela est vrai pour les lexies du type mentionné ci-dessus (*Parisiens, lions, ânes*, etc. au pluriel désignent les groupes mixtes de deux sexes) mais aucunement pour toutes les lexies du français : *frères* au pluriel ne désigne que des hommes. De plus, la composante 'de sexe masculin' est éliminée aussi dans les contextes génériques, dans des comparaisons, etc. :

- (21) a. *Un étudiant doit travailler beaucoup.*  
 b. – *Tu ne peux plus vivre en étudiant !*  
 c. *carte d'étudiant, vie d'étudiant, ...*

En (21), la lexie ÉTUDIANT désigne autant les hommes que les femmes (*Elle menait une vie d'étudiant* [\**d'étudiante*], etc.).

Pour rendre compte de ces faits, il faut indiquer, dans la définition d'ÉTUDIANT, que la composante 'de sexe masculin' est ce qu'on appelle une *composante faible* : elle est mise en veilleuse dans certains contextes. (Dans FRÈRE et GARÇON, par exemple, cette composante est forte – elle n'est jamais éliminée, tout à fait comme dans HOMME, ce que nous avons indiqué ci-dessus<sup>12</sup>.) La composante faible sera marquée par des parenthèses en gras munies de l'indice spécifiant les conditions qui annulent la composante en question. Ainsi, la définition complète d'ÉTUDIANT s'écrit comme suit :

*étudiant en Y de Z* = 'personne (de sexe masculin)<sub>gén</sub> qui étudie la matière Y dans un établissement d'enseignement supérieur Z',

où « (de sexe masculin)<sub>gén</sub> » indique que la composante 'de sexe masculin' est abolie dans des emplois génériques ainsi qu'au pluriel. Nous ne développerons pas davantage cette question, si intéressante soit-elle.

#### 3.3.1.4 Critères de formulation de la définition lexicographique

Les règles d'élaboration de définitions lexicographiques présentées en 3.3.1.3 ont pour objectif essentiel d'assurer, en conformité avec la logique sous-jacente, la parfaite CORRECTION FORMELLE d'une définition telle que la théorie l'impose; autrement dit, ces règles garantissent la bonne formation des définitions. Si une définition lexicographique n'est pas formellement correcte – elle ne répond pas à au moins une des règles formulées – nous ne pouvons ni la critiquer, ni la justifier, ni même l'améliorer. Par contre, si elle est correcte, elle remplit une condition nécessaire pour être utilisable, sans pour autant satisfaire à une

12 On remarquera qu'en espagnol, par exemple, la composante sexuelle est faible même dans les sens 'frère' et 'garçon' : *hermanos*, littéralement 'frères', signifie 'frères et sœurs', et *niños*, littéralement 'garçons', désigne les petits garçons et les petites filles ensemble.

condition de suffisance : elle doit en plus être vraie, c'est-à-dire CONFORME AUX FAITS de  $\mathcal{L}$ . En effet, le but de tout dictionnaire est d'avoir des définitions vraies; les définitions formellement correctes mais factuellement fausses ne servent à rien.

C'est pour cette raison que nous proposons, à côté des règles « formelles » de correction de nos définitions, une série de critères « substantiels », c'est-à-dire se rapportant au contenu. Nous avons déjà présenté cinq critères de ce type en 2.2, p. 61 *ssq.*, en relation avec le problème de la délimitation des lexies. À présent, il s'agit des critères dont la fonction est de fournir au lexicologue l'aide indispensable pour déterminer le choix des composantes sémantiques à inclure dans la définition de la lexie délimitée.

Commençons par le problème de choix des composantes dans la définition d'une lexie  $L$  qui désigne un objet physique  $O$ . Quelles sont les propriétés de cet objet qui doivent être reflétées dans la définition de  $L$  ? Certaines propriétés de  $O$  sont *constitutives* – en ce sens que si elles ne sont pas reflétées dans la définition de  $L$ , cette définition devient trop vague :  $L$  pourrait être utilisée à mauvais escient, pour désigner d'autres référents que  $O$ . Ainsi, la solidité de la glace ou l'invisibilité de l'air sont des propriétés constitutives de ces substances : quelque chose de LIQUIDE ne sera pas appelé GLACE en français, de même un gaz VISIBLE respiré par des êtres fantastiques sur une planète fantastique n'est pas de l'AIR. Ce sont les propriétés constitutives des objets qui contrôlent l'utilisation correcte des lexies correspondantes.

La difficulté apparaît quand la propriété en question, sans être constitutive, semble cependant intrinsèque à  $O$ . C'est, par exemple, le cas de la couleur blanche du sucre, du sel et du riz. Faut-il inclure dans les définitions des lexies SUCRE, SEL et RIZ la composante 'blanc' ? Le critère suivant permet de trancher ce type de difficulté.

### Critère 6 : Critère de pertinence linguistique

Soit  $\langle \sigma \rangle$  une composante potentielle pour la définition de la lexie  $L$ ; elle reflète une propriété non constitutive, mais intrinsèque du référent de  $L$ .

|| La définition de  $L$  doit inclure la composante  $\langle \sigma \rangle$  si et seulement si on peut trouver dans la langue étudiée une autre lexie  $L_1$  qui est formellement liée à  $L$  et dont la définition comprend  $\langle \sigma \rangle$  de façon indiscutable.

La lexie  $L_1$  qui doit étayer la pertinence linguistique de la composante potentielle  $\langle \sigma \rangle$  dans la définition de  $L$  peut être liée à  $L$  d'une des trois façons suivantes :

- $L_1$  est une lexie faisant partie du même vocable<sup>13</sup> que  $L$  ( $L_1$  et  $L$  se trouvent dans une relation de polysémie);

<sup>13</sup> Pour le concept de vocable, voir la Définition IV.4, ch. IV, 1.1, p. 159.

- L<sub>1</sub> est une lexie dérivée de L (L<sub>1</sub> et L se trouvent dans une relation de dérivation morphologique);
- L<sub>1</sub> est un phrasème incluant L (phraséologie).

En fait, dans les trois cas ci-dessus, il s'agit d'assurer, dans le DEC, le lien sémantique explicite entre deux lexies perçues par des locuteurs comme sémantiquement apparentées (voir *pont sémantique*, Ch. IV, Définition IV.1, p. 157).

## Exemples

### Polysémie

- (22) L = NUAGEI 'grand amas de substance blanche... QUI CACHE (PARTIELLEMENT) LE CIEL...';

⟨σ⟩ = la composante '...qui cache (partiellement) le ciel...';

L<sub>1</sub> = NUAGEII.3 'sentiment ou événement déplaisant... qui obscurcit<sup>2</sup> le fait X [comme un nuageI cache (partiellement) le ciel...]' [comme dans... *mais quelques nuages troublèrent leur lune de miel*].

L'existence, en français, de NUAGEII.3 démontre la pertinence linguistique de la composante ⟨σ⟩ = '...qui cache (partiellement) le ciel...' dans la définition de NUAGEI.

En effet, le lien sémantique entre NUAGEII.3 et NUAGEI est absolument évident pour un francophone. Selon la Règle de pont sémantique, ce lien doit être explicitement indiqué dans la définition de NUAGEII.3 par comparaison avec NUAGEI; comme il faut en plus spécifier la base de la comparaison, cela nous amène à la composante en question.

### Dérivation

- (23) ORAGEI 'perturbation atmosphérique VIOLENTE...' [= L] : la composante 'violent' [= ⟨σ⟩] est indispensable à cause de ORAGEUXII [= L<sub>1</sub>] 'impliquant une violence...' (*discussion (séance) orageuse*).

### Phraséologie

- (24) NEIGELa 'substance BLANCHE froide...' [= L] : l'indication de la couleur [= ⟨σ⟩] est pertinente pour la définition du lexème NEIGELa à cause du phrasème [BLANC] 'COMME NEIGE' [= L<sub>1</sub>]. Par contre, les définitions de SUCRE et de RIZ ne mentionnent pas la blancheur, même si ces deux substances sont blanches – puisqu'aucune expression ne met en jeu ces lexies avec une signification de blanc : le français n'a rien comme \**blanc comme sucre*, \**blancheur rizière*, etc.<sup>14</sup>. Mais la définition de SEL doit faire mention de 'blanc', étant donné l'expression 'POIVRE ET SEL' 'dont les cheveux / poils sont noirs ou bruns mêlés de blanc' [*Jean est poivre et sel*]. Ajoutons en passant que pour NEIGELa, le test de polysémie confirme la nécessité de la composante

14 Il n'en est pas de même dans d'autres langues ; ainsi, en russe, on dit *saxarnye zubki* 'quenottes de sucre', de sorte que la définition du lexème russe SAXAR 'sucre' doit inclure la composante 'blanc'. On voit clairement que notre approche est strictement lexicologique et non encyclopédique.

'blanche' : en effet, le français possède le lexème NEIGEII ('cocaïne en poudre'), le lien sémantique étant la blancheur.

Le Critère de pertinence linguistique permet au lexicologue de prendre une décision motivée dans les cas où il s'interroge sur la nécessité d'une composante sémantique non obligatoire. Appartient-elle au sémantisme de la lexie examinée ou relève-t-elle, tout simplement, des propriétés encyclopédiques du référent ? Le critère en cause force l'utilisation de considérations purement linguistiques : l'inclusion, dans la définition de L, d'une composante non obligatoire doit être justifiée soit par la polysémie, soit par la dérivation, soit encore par la phraséologie de L.

Pour conclure, indiquons que ce critère ne s'applique qu'à la définition des lexies concrètes (= *realia*) désignant des objets et des substances. Cela découle du fait que le Critère de pertinence linguistique vise des propriétés observables des objets réels.

Si le critère exposé ci-dessus est plutôt spécifique (lexies concrètes seulement), les trois critères suivants sont de nature plus générale. Ils se basent sur les possibilités combinatoires de L pour déterminer l'existence de certaines composantes sémantiques dans la définition de L. En particulier, ils prennent en charge la cooccurrence avec des modificateurs qualificatifs, avec des quantificateurs et avec la négation.

### **Critère 7 : Critère de cooccurrence avec modificateurs qualificatifs**

|| La définition de L doit refléter explicitement la cooccurrence de L avec les modificateurs qualificatifs.

« Refléter explicitement » signifie ici 'comporter une composante sémantique qui est en mesure d'accepter le sens des modificateurs en question'.

#### **Exemples**

##### *Modificateur adjectival*

(25) APPLAUDISSEMENT : puisque le nom *applaudissements* (au pluriel) reçoit tout naturellement les modificateurs adjectivaux du type *nourris* (*frénétiques*, *clairsemés*), où l'adjectif fonctionne comme opérateur d'intensification/d'atténuation, sa définition doit comprendre une composante en mesure d'accepter cette qualification. Une définition avec une telle composante (en petites majuscules) se lira comme suit :

*Applaudissement(s) de X à Y pour Z* = 'Battements de mains par X en signe d'approbation par X de Z de Y dont LA FORCE ET/OU LA FRÉQUENCE est/sont proportionnelle(s) au degré de cette approbation'.

La présence de la composante (... la force et/ou la fréquence...) facilement intensifiable/atténuable reflète la cooccurrence de APPLAUDISSEMENT non seulement avec les adjectifs cités mais avec d'autres adjectifs du même type (*de*

rare applaudissements, quelques applaudissements, ...) et, en plus, avec des noms comme SALVE, TONNERRE, TEMPÊTE (d'applaudissements).

### Modificateur adverbial

- (26) BATTREII (infliger une défaite) 'comme dans *Jean a battu Pierre au tennis*': puisque le verbe *battreII* peut avoir, comme opérateur d'intensification, le modificateur adverbial à *plate couture*, sa définition doit comprendre une composante intensifiable. Par conséquent, nous ne pouvons pas, par exemple, utiliser la composante '[le] dessus' et définir BATTREII par 'avoir le dessus sur...' (comme on le fait souvent dans des dictionnaires), étant donné que la composante '[le] dessus' n'est pas intensifiable : \**dessus complet* (finale, totale), \**dessus éclatant* (retentissant), etc. Par contre, la composante 'défaite' est naturellement intensifiable : *défaite complète* (finale, totale), *défaite éclatante* (retentissante), etc. Nous la retenons donc pour la définition.

L'observation systématique du Critère de cooccurrence avec modificateurs qualificatifs aide à la fois à éviter la perte de composantes indispensables dans les définitions<sup>15</sup> et à choisir la bonne composante. Ainsi, dans l'exemple (25), si l'on ne prend pas en compte la cooccurrence avec les adjectifs intensificateurs/atténuateurs, on risque de ne pas introduire la composante indispensable 'force/fréquence proportionnelles à l'approbation'. De même, dans l'exemple (26), il serait impossible de choisir entre les composantes '[le] dessus' et 'défaite'.

Les cas analysés ci-dessus ne sont évidemment que des illustrations du procédé à utiliser. Les modificateurs du type considéré ci-dessus – les valeurs de la fonction lexicale *Magn*, voir plus loin, 3.5.1-2, p. 125 *ssq.* – ne sont pas les seuls à prendre en compte; le lexicologue doit considérer TOUTES sortes de modificateurs. Cependant, notre choix d'illustrations n'est pas entièrement fortuit : les cooccurents lexicaux restreints de la lexie L, ayant des liens intimes avec des composantes de la définition de L, sont d'une part immédiatement disponibles et d'autre part permettent de dégager plus facilement le sémantisme de L. La correspondance entre la définition de L et sa cooccurrence lexicale restreinte, c'est-à-dire les fonctions lexicales, est un des principes directeurs dans l'élaboration et la présentation des articles de dictionnaire du DEC.

### Critère 8 : Critère de cooccurrence avec quantificateurs

|| La définition de L doit refléter explicitement la cooccurrence de L avec les quantificateurs – surtout avec les numéraux, si L est un nom, et avec l'expression *n fois*, si L est un verbe.

La pluralisation des noms est traitée comme un cas particulier de quantification.

<sup>15</sup> De telles pertes se produisent couramment dans les dictionnaires existants. Ainsi, aucun, à notre connaissance, n'indique la proportionnalité de l'intensité de l'applaudissement avec l'approbation.

La vérification de la cooccurrence avec les quantificateurs doit se faire au moyen des contextes spécialement construits pour mettre en relief les liens qui peuvent exister entre la quantification/la pluralisation et le sens.

### Exemple

Comparons les quatre noms de légumes suivants :

AIL ~ OIGNON ~ CAROTTE ~ CHOU

Tous les quatre sont pluralisables, de sorte que, si nous nous limitons à cette constatation, il n'en découle rien pour la structure de leur définition. Pourtant, on dit *Apporte-moi des oignons / des carottes / des choux*, mais pas *\*des ails* (<*\*des aulx*>); la pluralisation de AIL (comme dans *Les ails* (<*Les aulx*>) *du Mexique sont très diversifiés*) ne se fait donc pas dans les mêmes contextes que celle des trois autres noms considérés. De plus, à la différence des trois autres noms, AIL ne se combine pas avec des numéraux :

- (27) *Apporte-moi un* <trois> *oignon(s) / une* <trois> *carotte(s) / un* <trois> *chou(x)*  
 mais  
*\*Apporte-moi un ail* <trois ails (<aulx>);

il faut dire *une tête d'ail* et *trois têtes d'ail*.

Par conséquent, les quatre noms en question ne peuvent se définir de la même façon – par exemple, par ‘plante qui...’. En vertu du Critère de cooccurrence avec quantificateurs, nous devons, tout d’abord, séparer AIL des trois autres noms. En effet, AIL a deux acceptions et devra être décrit par deux lexèmes : le premier – ‘condiment qui...’, et le deuxième – ‘plante qui produit ce condiment qui...’; OIGNON, CAROTTE et CHOU seront plutôt définis différemment comme ‘légume qui...’.

Mais les trois noms OIGNON, CAROTTE et CHOU, malgré leur parenté sémantique évidente, montrent certaines différences de comportement dans les contextes « quantifiés » :

- (28) a. – *Manger de l'oignon* <des oignons> / <\*de la carotte> *des carottes / du chou* <des choux>, *c'est bon pour la santé*.  
 b. *J'aime bien l'oignon* <<sup>1</sup>les oignons> / <<sup>1</sup>la carotte> *les carottes / le chou* <<sup>1</sup>les choux>.  
 c. *Dans mon jardin, l'oignon* <les oignons> / *la carotte* <les carottes> / *le chou* <les choux> *pousse(ent) bien*.

Ces différences requièrent une explicitation sémantique, ce qui nous amène à dégager les trois acceptions suivantes, qui correspondent aux exemples ci-dessus :

- a'. 'Unité de légumes U qui... [taille, forme, couleur, consistance, goût, ...]'  
 [= (27)]  
 b'. 'Matière comestible de U...' [= (28a-b)]  
 c'. '(Ensemble des) plante(s) destinée(s) à produire U...' [= (28c)]

Ce schéma de définition doit être appliqué de façon systématique à tous les noms de légumes. Dans chaque acception, on indiquera la pluralisation possible : ainsi, dans l'acception du type **b'** ('Matière comestible de U...'), OIGNON a les deux nombres, CAROTTES est un *plurale tantum*, et CHOU est plutôt un *singulare tantum*, avec le pluriel qui n'est tout de même pas impossible.

Le Critère de cooccurrence avec quantificateurs, en plus de permettre de distinguer les lexèmes d'un même vocable, contribue également au choix de la composante sémantique générique de la définition. Ainsi, dans l'acception du type **a'** ('Unité de légumes U...'), la composante 'unité' permet la pluralisation, l'emploi de l'article UN et la combinatoire avec les numéraux (*des unités, une unité, trois unités*).

### Critère 9 : Critère de cooccurrence avec négation

|| La définition de L doit refléter explicitement la façon dont L se combine avec la négation.

Plus spécifiquement, le lexicologue doit étudier en profondeur l'expression *non L* pour dégager les cas où le sens de *non L* n'est pas une simple négation (au sens logique) de L. Si les phrases *Jean n'a pas mangé* et *Cette maison n'est pas en bois* sont des négations « pures », il n'en est pas toujours ainsi : l'ajout syntaxique d'une particule négative peut être lié à des effets de sens assez particuliers. Nous en donnerons deux exemples.

#### Exemples

Comme première approximation, la définition du nom VEUVE peut se formuler comme suit :

(29) *X est veuve* ≈ 'X est une femme en âge d'être mariée qui a perdu son mari'.

['X a perdu son mari' = 'Le mari de X est mort'; nous utilisons ici la paraphrase avec PERDRE pour des raisons syntaxiques et communicatives.]

Pendant, cette définition n'est pas tout à fait satisfaisante puisque, avec cette définition, la négation de la phrase *Zin-Lin n'est pas veuve* devrait signifier que Zin-Lin N'EST PAS une femme. Pourtant, tout en niant le fait que le mari de Zin-Lin est mort, cette phrase, qui ne peut référer ni à un homme, ni à une fillette, continue à affirmer que Zin-Lin est une femme en âge d'être mariée. Pour refléter cette particularité dans la définition, il faut que la composante 'femme en âge d'être mariée' ait un statut différent de celui de la composante 'qui a perdu son mari', car seule la dernière composante peut subir la négation. Par conséquent, la composante 'femme en âge d'être mariée' doit être rattachée à la variable en tant que modificateur :

*X est veuve* = 'X, qui est une femme en âge d'être mariée, a perdu son mari'.

Une définition ainsi reformulée assure la description voulue :

*X n'est pas veuve* = 'X, qui est une femme en âge d'être mariée, n'a pas perdu son mari'.

Comme on le voit, le Critère de cooccurrence avec négation permet d'établir la composante sémantique qui doit apparaître dans la définition en tant que *contrainte sur une variable*, plutôt que de faire partie du « corps » de la définition.

Considérons maintenant la phrase (30) :

(30) *Papa ne m'a pas permis d'aller au bal.*

Malgré la négation, elle affirme que 'je voulais aller au bal' (ou, au moins, 'papa croyait que je voulais aller au bal'); la composante sémantique de PERMETTRE correspondant au 'desir de Z de Y-er' ne peut pas être affectée par la négation. En conformité avec le critère posé, cette particularité doit être reflétée dans la définition. Nous le faisons en donnant à la composante en question le statut de *présupposé*. La partie présuppositionnelle d'une définition lexicographique du DEC est donc marquée pour la soustraire à l'action de la négation; la notation proposée est le symbole « || », qui sépare la présupposition du reste du définissant. Comme résultat, on a pour PERMETTRE la définition suivante :

*X permet à Z de Y-er* = 'Sachant ou croyant que Z veut faire l'action Y que Z ne doit pas faire contre la volonté de X, || X communique à Z que Y ne serait pas contraire à la volonté de X et, par conséquent, X donne à Z la liberté de faire Y' (DEC-2).

En comparant la formulation des deux définitions, on se rend facilement compte qu'une contrainte sur une variable est en fait logiquement équivalente à un présupposé (et, bien entendu, *vice versa*). Par conséquent, on peut les reformuler comme suit :

*X est veuve* = 'X étant une femme en âge d'être mariée, || X a perdu son mari'.

*X permet à Z de Y-er* = 'X, qui sait ou croit que Z veut faire l'action Y que Z ne doit pas faire contre la volonté de X, communique à Z que Y ne serait pas contraire à la volonté de X et, par conséquent, X donne à Z la liberté de faire Y'.

Cela signifie qu'il existe deux façons équivalentes d'écrire les présupposés; le choix de l'une ou de l'autre dépend de considérations pédagogiques (aisance de lecture).

### 3.3.1.5 Structure et contenu de la définition lexicographique

Parmi les divers problèmes complexes généraux que soulève la définition lexicographique, nous ne retiendrons ici que les deux suivants : la hiérarchisation (3.3.1.5.1) et l'ordonnancement (3.3.1.5.2) des composantes; ce sont hors de tout doute deux problèmes parmi les plus importants de la définition.



### 3.3.1.5.1 Hiérarchisation des composantes sémantiques dans la définition lexicographique

Le texte d'une définition du DEC (rédigée en français) se présente comme un énoncé HOMOGENE, constitué de lexies françaises munies de numéros lexicaux distinctifs. Pourtant, il n'en est rien : en réalité, une définition du DEC a un caractère feuilleté, c'est-à-dire qu'elle comprend plusieurs couches de composantes aux propriétés et comportements logiques fort différents. Nous avons déjà évoqué cette question à plusieurs reprises, au moment où nous avons établi des composantes GÉNÉRIQUES ou des composantes FAIBLES, etc., montrant ainsi l'existence de types différents de composantes sémantiques. Ici, nous faisons le point sur des notions déjà examinées dans les diverses analyses précédentes.

Comme il s'agit d'un secteur de recherches assez nouveau dans le cadre du DEC, nos propositions sont à considérer avec toute la prudence nécessaire. Ceci dit, nous distinguons six types de composantes sémantiques :

– **Composante ordinaire.** Ce type de composante représente la grande majorité des composantes sémantiques dans le DEC en général et dans chaque définition en particulier; nous en avons longuement discuté dans ce chapitre. Ici nous nous contenterons d'attirer de nouveau l'attention sur le fait fondamental suivant :

|| Une définition comprenant la composante sémantique 'σ' hérite automatiquement TOUTES les propriétés sémantiques et syntaxiques de 'σ'.

Cette affirmation découle naturellement de la Règle 5, p. 91 – la Règle de substituabilité absolue, qui s'applique à la fois aux textes et aux définitions lexicographiques.

Au niveau sémantique, cela signifie, par exemple, que TABOURET ne peut se définir comme 'chaise sans dossier', étant donné que CHAISE se définit comme 'siège avec dossier'; la substitution de cette dernière définition à la composante 'chaise' dans la définition de TABOURET produirait une contradiction évidente (TABOURET doit se définir par 'siège'). De la même façon, TRICYCLE ne doit pas se définir comme 'bicyclette à trois roues', mais comme 'cycle à trois roues'; etc.

Au niveau syntaxique, une composante 'σ' à sens prédicatif amène inéluctablement tous ses actants dans la définition où elle figure. Ainsi, PAYERI.2a (dans *payer ses dettes, les taxes, l'échéance, le loyer*) signifie approximativement 's'acquitter d'une obligation financière Y'. Cependant, 'obligation', qui est un prédicat au sens sémantique, a ses propres trois actants : 'X a une obligation Y envers Z', de sorte que tous les trois doivent apparaître dans la définition (partielle) de PAYERI.2a :

*X paie Y à Z = 'X s'acquitte d'une obligation financière Y que X a envers Z...'*

Comme on peut facilement le constater, cet actant Z est bien exprimable en français : *Jean a payé ses dettes à ses créanciers, Jean paie le loyer à son propriétaire toujours à temps*, etc.

Bien entendu, ce qui vient d'être dit au sujet de l'héritage sémantico-syntaxique des propriétés des composantes ordinaires est tout à fait valable pour toutes les composantes sémantiques.

– **Composante générique.** La composante générique de la définition de la lexie L réfère au genre prochain (au sens logique) auquel appartient le signifié de L. Ainsi, CHAISE, TABOURET, BANC, FAUTEUIL, CANAPÉ, etc. sont des sièges, et la composante générique dans leur définition est 'siège'. La composante générique dans la définition de L doit être choisie de façon à assurer pour L la description de la cooccurrence lexicale libre de L. Considérons, par exemple, le lexème NIVEAU.I.1 (comme dans *Le niveau de la Seine monte* ou *Le liquide dans le réservoir a atteint le niveau de la cote d'alerte*). Au premier abord, il semble qu'on puisse le définir comme 'plan<sup>2</sup>1b.c. horizontal imaginaire dans lequel se situe la surface d'un liquide...'. Cependant, 'plan<sup>2</sup>1b.c.' ne convient aucunement comme composante générique pour NIVEAU.I.1. Si on dit facilement *vérifier* (mesurer, indiquer, jauger, rétablir) le niveau, on ne peut pas dire \**vérifier* (\*mesurer, \*indiquer, \*jauger, \*rétablir) le plan; par contre, quelque chose est situé dans un plan, mais pas \*dans un niveau (dans le sens qui nous préoccupe ici). Ces exemples démontrent que 'niveau.I.1' n'est pas un cas particulier de 'plan<sup>2</sup>1b.c.'; c'est plutôt un cas particulier d'un paramètre, à savoir, la hauteur. En effet, avec 'hauteur.I.2'<sup>16</sup> en tant que composante générique de NIVEAU, tous les problèmes de cooccurrence lexicale libre que nous venons de voir disparaissent – parce que la cooccurrence de HAUTEUR.I.2 est plus ou moins valable pour NIVEAU.I.1 : *une hauteur époustouflante* ~ *un niveau époustouflant* [\*un plan époustouflant], *être à la hauteur de 3 m* ~ *être au niveau de 3 m*, *atteindre la hauteur de 3 m* ~ *atteindre le niveau de 3 m*, etc. Comme résultat, la définition de NIVEAU.I.2 aura la composante générique 'hauteur.I.2' :

*Niveau.I.1 de X / de Y = 'Hauteur.I.2 de Y du plan<sup>2</sup>1 horizontal imaginaire dans lequel se trouve la surface de la matière liquide X...'*

Il serait souhaitable de marquer la composante générique explicitement dans la définition ; on peut le faire par des PETITES MAJUSCULES, par exemple.

– **Composante faible.** Comme nous l'avons déjà montré (ce chapitre, 3.3.1.3, p. 78), certaines composantes des définitions peuvent s'estomper dans des conditions précises; ce sont des composantes faibles. Tel est le cas de la composante 'sexe masculin', citée pour l'analyse du lexème ÉTUDIANT (p. 94).

– **Composante optionnelle.** La composante 'employée par un établissement d'enseignement' dans la définition du lexème PROFESSEUR (voir p. 74) est optionnelle. En général, 'professeur' implique par défaut l'exercice dans un établissement d'enseignement; cependant, ce n'est pas obligatoire : comme nous l'avons déjà signalé, on peut bien être *professeur de langues* (de math) à domicile.

<sup>16</sup> *Hauteur.I.2 de X / de Y par rapport à Z = 'Distance de Y entre X et Z, X étant au-dessus de Z' (La hauteur du toit par rapport à notre fenêtre est de 5 m).*

Pour refléter ce fait, on marque la composante 'employée par un établissement d'enseignement' comme optionnelle. Sauf mention explicite du contraire, PROFESSEUR désigne une personne enseignant dans un établissement d'enseignement; mais s'il le faut, cette composante peut être facilement mise à l'écart, c'est-à-dire qu'elle peut être NEUTRALISÉE par un contexte contradictoire explicite (Aprésjan 1979 : 514 [composantes variables], Zaliznjak 1987 [composantes instables], Iordanskaja & Mel'čuk 1990 : 337).

Nous marquerons une composante faible ou optionnelle par la mise entre parenthèses, avec indication possible (par un index) des conditions contextuelles de neutralisation.

– **Contraintes sur les variables.** À côté des composantes sémantiques ordinaires, qui apportent chacune du contenu sémantique, les définitions lexicographiques doivent encore faire appel à des composantes caractérisant les actants sémantiques, c'est-à-dire contraignant les variables correspondantes. Par exemple, dans *X paie Y en Z à W pour U* (*Jean a payé 100 \$ en argent liquide à Marie pour ce pantalon*), la variable Y doit être contrainte : 'Y, QUI EST UNE SOMME D'ARGENT'. Ainsi, si le bien Y transféré par X à W pour U n'est pas de l'argent (ou son équivalent), le verbe PAYERI.1a (DEC-2) ne peut pas être utilisé. La phrase *J'ai \*payé pour ce chien cinq bouteilles d'eau de vie* représente un abus de sens. Il faut dire normalement *J'ai donné cinq bouteilles..., J'ai échangé <troqué> ce chien contre cinq bouteilles..., J'ai eu ce chien pour cinq bouteilles..., etc.* On voit donc la nécessité et le rôle de la contrainte indiquée ci-dessus. Elle fonctionne comme un filtre qui limite l'ensemble des valeurs permises de la variable concernée – plutôt que d'apporter du sémantisme à la définition. La composante '[X], qui est une femme en âge d'être mariée' dans l'exemple (29) en 3.3.1.4, p. 101, est bien entendu une contrainte du même genre qui joue le même rôle.

Les deux contraintes citées ne ciblent que des variables individuelles. Il existe, cependant, des contraintes s'exerçant sur plusieurs variables à la fois. Ainsi, l'énoncé *Je suis certain que M. Tremblay persécute Jean* ne peut être approprié que si le locuteur sait que M. Tremblay a un certain pouvoir sur Jean; dans le cas contraire, même si ces actions restent les mêmes (harcèlement injuste sans relâche), on ne peut les qualifier de *persécutions*.

Les contraintes sur les variables actantielles ne sont en fait des conditions de bon usage (des lexies) qui font appel non seulement à la représentation sémantique de départ, mais aussi à toutes sortes de connaissances dont le locuteur dispose, y compris des connaissances du monde.

Il n'est pas toujours évident de savoir s'il y a nécessité d'établir des contraintes. Un cas typique concerne les cris des animaux. L'approche habituelle des dictionnaires est de les définir comme suit (prenons l'exemple du verbe BARRIR) : « En parlant de l'éléphant, pousser un barrissement », BARRISSEMENT étant « cri de l'éléphant ». En effectuant la substitution et en adaptant la définition à notre formalisme, nous obtenons :

*X barrit* = 'X, qui est un éléphant, émet le cri typique de l'éléphant'.

Pour éliminer la redondance produite par la répétition de 'éléphant', on peut supprimer soit la contrainte sur X, soit la caractéristique du cri. Le choix se fait à partir des exemples suivants :

(31) a. *Cet animal barrit – tout à fait comme un éléphant.*

b. « *C'est un salaud !* », *barrissait André brandissant la hache.*

Ces phrases banales montrent clairement que X n'est pas nécessairement un éléphant : cela peut être un autre animal ou une personne. Il faut donc enlever la contrainte; la définition finale se lira ainsi :

*X barrit* = 'X émet le cri typique de l'éléphant'.

Les contraintes se reconnaissent par leur position syntaxique : elles sont toujours des modificateurs attachés aux variables.

– **Partie présuppositionnelle.** Comme nous venons de discuter des présuppositions dans les définitions lexicographiques à la toute fin de la sous-section 3.3.1.4, nous nous contenterons ici de fournir un autre exemple. Le verbe AIDER1 présuppose une certaine activité de la part de la personne aidée. On ne peut dire *Jean a aidé Marie à laver la vaisselle* que si Marie lavait effectivement la vaisselle; de même, la phrase *Jean n'a pas aidé Marie à laver la vaisselle* continue, malgré la négation générale, à affirmer que Marie lavait la vaisselle. Cette particularité sémantique est exprimée dans notre définition d'AIDER1 par une composante présuppositionnelle :

*X aide Y à Z-er par W* = 'Y faisant ou essayant de faire Z,|| X utilise ses ressources W dans le but de faciliter Z pour Y'.

### 3.3.1.5.2 Ordonnement des composantes sémantiques dans la définition lexicographique

De façon générale, l'ordre des composantes sémantiques dans une définition est déterminé par leur nature et celle des liens avec d'autres composantes. Cependant, il peut exister des cas où, au premier abord, l'ordonnement des composantes semble arbitraire. Pourtant, il n'en est rien. La Règle d'ordonnement par pertinence sémantique, qui fixe l'ordre des composantes dans des cas de ce type, apporte la décision voulue.

#### Règle 6 : Règle d'ordonnement par pertinence sémantique

|| La composante exprimant une caractérisation ( $\sigma_1$ ) sémantiquement plus pertinente qu'une autre caractérisation ( $\sigma_2$ ) doit précéder la composante exprimant ( $\sigma_2$ ).

La plus/moins grande pertinence d'une caractérisation sémantique se manifeste dans la langue, entre autres, par l'ordre linéaire des adjectifs épithètes d'un même nom. L'adjectif portant la caractérisation plus pertinente se trouve plus près du nom qu'un autre adjectif portant une caractérisation moins importante. Ainsi,

si l'on ne tient pas compte des effets de style et de l'expressivité communicative (contraste, mise en relief, ...), on dira obligatoirement *un repas délicieux pas cher* plutôt que *'un repas pas cher délicieux*; pour le sémantisme de REPAS, les caractéristiques ('être délicieux', 'être nutritif', 'être gastronomique'), etc. sont plus pertinentes que 'être cher' (même si dans la vie réelle le coût peut être très pertinent !). De la même façon, on a *un livre français passionnant*, et non pas *'un livre passionnant français* (pour un texte, la langue est sémantiquement plus pertinente que l'impact subjectif sur le lecteur), *bois sec poussiéreux*, et non pas *'bois poussiéreux sec* (pour du bois, être sec est sémantiquement plus pertinent qu'être couvert de poussière), etc.

À titre d'illustration, appliquons ce principe à la définition d'une lexie comme DOIGTI.a ('partie terminale de la main ...'). Un doigt possède trois caractéristiques qui nous semblent inhérentes : il est allongé, articulé et mobile; quelle est leur pertinence sémantique relative ? Il n'est pas difficile d'imaginer des doigts extrêmement courts : la phrase *Il avait des doigts si courts qu'il avait du mal à tenir la pomme* ne choque pas du tout. La caractéristique ('être articulé') peut se perdre lors d'une métaphorisation. Ainsi, les doigts d'un appareil peuvent ne pas être articulés : il n'y a pas d'inconvénient à appeler *doigts* les cinq pièces non articulées en métal d'un mécanisme de réglage. Cependant, un doigt qui est non mobile de façon inhérente est inconcevable; ainsi, les cinq pièces fonctionnelles du râteau ne peuvent pas être des *doigts* et ne sont pas appelées ainsi (un râteau et une fourche n'ont que des *dents*). Il s'ensuit que dans la définition de DOIGTI.a, les caractéristiques en cause doivent se trouver dans l'ordre suivant : '[partie] mobile articulée allongée' – respectant ainsi le principe énoncé.

**NB :** Dans un réseau sémantique, il n'est évidemment pas possible d'utiliser la Règle 6, puisque les éléments d'un réseau ne sont pas soumis à l'ordre linéaire. Dans la représentation sémantique véritable, il faudra introduire une technique appropriée explicite pour marquer le degré de pertinence sémantique.

### 3.3.1.6 Une liste de vérification lexicographique [= une check-list]

Comme nous venons de le voir, la tâche du lexicologue/lexicographe lors de l'élaboration de la définition d'une lexie vedette L dans le DEC est extrêmement complexe. Très souvent, il faut réexaminer le tout, faire des retours en arrière, découper certaines composantes ou en ajouter d'autres, changer des détails (ce qui peut avoir des répercussions imprévisibles), s'adapter à de nouvelles données contextuelles, etc. Il y a donc risque de se perdre dans le labyrinthe de principes, règles et critères. Pour pallier ce danger, nous pensons utile de proposer une liste de problèmes à traiter lors de l'élaboration des définitions : une LISTE DE VÉRIFICATION LEXICOGRAPHIQUE.

L'idée d'une telle liste n'est pas neuve; elle a déjà été présentée, notamment, dans Hudson 1988, pour ce qui est de l'ensemble des types d'informations à

inclure dans un article de dictionnaire et pour ce qui est de la structure de ce dernier<sup>17</sup>. Quant à notre liste de vérification ci-dessous, adaptée de l'article Dostie *et al.* 1992 : 195-196, 1) elle ne porte que sur la définition et ainsi se focalise entièrement sur le sémantisme de L, et 2) elle consiste uniquement en une série de questions qui guident le lexicologue dans la recherche d'une définition adéquate. De plus, ce n'est qu'une ébauche s'appuyant sur les développements présentés jusque-là. De telles listes de vérification seraient des plus utiles pour toutes les zones et les sous-zones de l'article de dictionnaire; pour l'instant, cependant, nous devons nous satisfaire de la présente tentative.

Dans la liste ci-dessous, les questions suivent un ordre logique, sans toutefois former une hiérarchie stricte; une telle hiérarchie nécessiterait une recherche additionnelle.

### LISTE DE VÉRIFICATION LEXICOGRAPHIQUE

#### – Pour toute lexie L

1. L est-elle un prédicat (au sens sémantique) ?  
[Si oui, le défini pour L doit être une forme propositionnelle.]
2. Si L est un prédicat, combien a-t-elle d'actants sémantiques ?  
[La réponse détermine la structure et la formulation du défini et du définissant.]
3. Quelle est la composante générique du définissant de L ?  
[La réponse détermine l'orientation générale du définissant; ainsi, si CHANGER est 'COMMENCER à être différent', la composante générique 'commencer' en fait un verbe inchoatif, et cette propriété doit s'avérer dans tous les contextes.]
4. La définition de L comporte-t-elle des composantes optionnelles ou faibles ?  
[Si oui, elles doivent être identifiées par des parenthèses.]
5. La définition de L comporte-t-elle des disjonctions<sup>18</sup> ?  
[La question se rapporte au problème de la division des lexies.]

---

17 Il nous faut mentionner ici une des premières tentatives de formaliser la procédure du traitement lexicographique des données de départ en vue de l'élaboration d'articles de dictionnaire. Il s'agit de Descamps *et al.* 1976, où l'on trouve à la page 1363 un algorithme détaillé pour le classement des contextes et leur distribution entre les lexies.

18 La disjonction a une valeur particulière pour la définition lexicographique. Elle permet de réunir sous un même lexème deux sens différents lorsque le lexème connaît des emplois vagues en regard de ces deux sens (cf. Apresjan 1974, surtout page 85, et Mel'čuk 1988 : 36).

6. Quelles sont les contraintes sémantiques sur les actants de L ?  
[La question impose une vérification pointue de toutes ces contraintes; par exemple, l'actant Y (= destinataire) de LETTRE au sens de missive doit nécessairement être humain (y compris des collectivités, des organismes civils, des organisations, etc.) : *une lettre à Jean* (à la nation, au ministère, au journal, ...); Y ne peut pas être un ordinateur – \**une lettre à l'ordinateur principal*, même si l'on dit facilement *un message à l'ordinateur principal*.]
7. Quelles sont les relations sémantiques entre les actants de L ?  
[La question attire l'attention sur le problème illustré par l'exemple suivant : comparons les lexies LETTRE (*lettre de X à Z sur Y*) et INTERDIRE (*X interdit à Z de faire Y*). Dans LETTRE, il n'y a aucune relation sémantique entre les actants X (= expéditeur) et Z (= destinataire); en fait, n'importe qui peut envoyer une lettre à n'importe qui. Mais dans INTERDIRE, il y en a une entre X et Z : si X *interdit* à Z de faire Y, il faut que X ait sur Z le pouvoir socialement reconnu de donner le droit de faire Y; ainsi, on dira *Le professeur a interdit à l'étudiant de le tutoyer*, mais ce même fait venant de la part d'un étudiant sera décrit comme *L'étudiant a défendu* (\**interdit*) *au professeur de le tutoyer*. On ne peut utiliser INTERDIRE pour un agent socialement inférieur au destinataire. Lors de l'élaboration d'une définition, il faut donc prêter une attention spéciale aux relations sémantiques possibles entre les actants de la lexie définie; dans la plupart des cas, il s'agit soit de relations du type 'X est socialement supérieur à Y' (nécessaire pour la définition de INTERDIRE) ou 'X appartient à la sphère personnelle de Y', nécessaire, par exemple, pour la définition de REPROCHER. (On ne peut *reprocher* qqch. qu'à une personne avec laquelle on a établi des liens; on peut reprocher un comportement nonchalant à sa femme ou à un collègue, mais certainement pas à un simple passant.)]
8. Comment se répartissent les composantes sémantiques entre posé et présupposé ?  
[La question est liée, entre autres, au Critère 9 – cooccurrence avec la négation.]
9. Quelle est la position du locuteur vis-à-vis des faits et des actants mis en jeu (approbation/désapprobation, implication personnelle/distanciation, etc.) ?  
[Par exemple, la différence sémantique entre *les Américains* et *les Amerloques* s'exprime par la composante 'le locuteur désapprouve', en sorte que AMERLOQUE = 'Américain que le locuteur désapprouve'. Cette question oblige le lexicologue à tenir compte des composantes exprimant l'attitude du locuteur quant au sens « objectif » de la lexie.]

10. Comment L se comporte-t-elle face à la négation, l'interrogation et l'enchaînement argumentatif ?

[La question est impliquée dans la détermination de la structure logique de la définition, c'est-à-dire qu'elle permet de distinguer entre le présupposé et le posé, de dégager les contraintes sur les variables, etc.]

11. Quelles sont les relations temporelles entre les faits mis en jeu par le sémantisme de L ?

[La question traite des événements qui précèdent ou suivent d'autres événements ou encore sont concomitants avec eux. Ainsi, le lexicologue peut se demander si le sens de FACTURE 'document présenté par X à Y qui indique la somme Z que Y doit payer I.1a à X pour la marchandise ou les services W vendus I.1 par X à Y' inclut l'indication de postériorité entre les événements 'présenter', 'payer I.1a' et 'vendre I.1'. Ici, la relation de postériorité existe entre la présentation et le paiement de la facture, mais pas entre sa présentation et la vente, qui peut avoir lieu avant, après ou bien en même temps que la facture est présentée.]

– Pour une lexie verbale L

12. L peut-elle être employée performativement et, dans l'affirmative, quelles en sont les conséquences pour les cooccurrents ?

[Comparez *Je vous remercie infiniment* [emploi performatif de REMERCIER] vs *Je l'ai remercié \*infiniment* [emploi non performatif]; ou encore *Je vous remercie \*chaleureusement* vs *Je l'ai remercié chaleureusement*.]

13. L est-elle factive ?

[Une lexie factive présuppose la véracité de son complément phrastique. Ainsi, l'expression 'FAIRE SAVOIR' est factive puisque *J'ai fait savoir la date d'arrivée du père à Marie* entraîne sans équivoque que Marie connaît la date d'arrivée du père (à moins qu'elle l'ait oubliée), alors que ANNONCER non factif se comporte différemment. La phrase *J'ai annoncé la date d'arrivée du père à Marie* peut être continuée avec *mais elle n'a pas reçu mon message*.]

14. L est-elle un verbe d'état, d'activité, d'accomplissement ou d'événement (cf. classification aspectuelle de Vendler) ?

15. L admet-elle le passé composé (au sens de l'accompli) et le passé simple ?

16. Comment L se combine-t-elle avec des éléments qui sont des révélateurs des caractéristiques aspectuelles (*depuis, cesser de, être en train de, etc.*) ?

[Comparez *J'ai cessé de fumer* vs *\*J'ai cessé de me coucher*; on notera, cependant, que avec le sens itératif, la construction devient possible : *J'ai cessé de me coucher à trois heures du matin*.]

17. L admet-elle le passif et le réfléchi ?



– Pour une lexie nominale L

18. L doit-elle être définie au singulier ou au pluriel ?

[Comparez SKI, qui doit être défini au pluriel : l'équipement en cause est une paire de skis et pas un ski unique; de la même façon, nous définissons YEUX plutôt qu'ŒIL, puisque l'organe de la vue, ce sont les yeux, et non l'œil. Voir p. 40, définition des noms ethniques.]

19. L admet-elle les deux nombres grammaticaux (singulier ET pluriel) ?

20. L admet-elle les modificateurs numériques ?

[Comparez *\*trois neiges* bien que le pluriel *les neiges* existe.]

Nous ne prétendons aucunement avoir épuisé toute la question de la définition, mais nous estimons que les renseignements et les explications fournis permettent de considérer la notion de définition comme opératoire dans notre cadre. Nous allons donc passer maintenant à l'étude de la connotation.

### 3.3.2 Connotation lexicographique

Comme nous l'avons signalé au début de la sous-section 3.3, p. 72, de façon générale, le sémantisme d'une lexie L n'est pas toujours complètement couvert par sa définition; L peut évoquer certains éléments de sens qui n'appartiennent pas à sa définition mais qui, pourtant, déterminent des usages figurés de L. Ainsi, le français dira pour une thèse universitaire sans valeur *C'est du vent !*, ce qui suppose l'existence du phrasème [C'EST DU VENT]<sup>1</sup> (le texte dont il est question et qui est censé avoir du contenu est sans contenu de valeur) (*Ce qu'il a écrit sur les équations de gravitation, ce n'était que du vent; Il m'a promis le bonheur, mais c'est du vent; Ses exploits au Népal, c'est du vent*). Tout locuteur du français voit immédiatement le lien sémantique entre ce phrasème et le lexème VENTI.1 ('phénomène atmosphérique – déplacement de l'air...') : « Mais oui, on dit *C'est du vent !* parce que le vent est quelque chose de léger, de fugace, sans substance palpable ». Le lexicologue se doit donc de refléter, dans l'article de dictionnaire de VENTI.1, cette intuition. Cependant, il ne peut pas le faire en ayant recours à la définition. La composante 'quelque chose de léger ou de fugace' ne doit pas être incluse dans la définition du lexème VENTI.1 pour au moins deux raisons. D'une part, en conformité avec la Règle de substituabilité, cette composante n'est pas sémantiquement nécessaire, étant donné qu'une définition de VENTI.1 qui ne la renferme pas remplit quand même la condition de substituabilité mutuelle absolue. D'autre part, il n'existe pas de cooccurrent de VENTI.1 qui nous obligerait à inclure la composante en question dans la définition : les Critères 1-5 ne l'exigent pas non plus. Pourtant, on sent que VENTI.1 implique des éléments sémantiques comme 'absence de contenu...', 'fugacité...', et c'est pour tenir compte de tels éléments sémantiques, tout à fait réels mais qui ne relèvent pas de la définition de la lexie étudiée, que nous utilisons le concept de connotation lexicographique.

Le terme *connotation* connaît des emplois fort divers, parfois même contradictoires. Pour éviter la confusion, nous rejetons deux interprétations répandues de ce terme. Premièrement, nous excluons de notre perspective la connotation logico-philosophique, qu'on oppose traditionnellement à la dénotation et qui se rapporte aux distinctions entre le sens (= l'intension) et le référent (= l'extension). Deuxièmement, nous refusons également l'interprétation lexicographique traditionnelle, qui fait relever de la connotation tout élément sémantique dit « subjectif », en opposant ainsi le sens connotatif au sens « objectif ». Cet usage est généralement illustré par l'exemple classique anglais FAMOUS 'célèbre = connu de tout le monde pour quelque chose de bon' vs NOTORIOUS 'tristement célèbre = connu de tout le monde pour quelque chose de mal' : on dit que ces deux adjectifs s'opposent par leurs connotations, à savoir, 'bon' vs 'mal'. Quant à nous, de telles distinctions appartiennent simplement à la définition.

Nous proposons, reprenant l'idée d'Apresjan (1974 : 67-68), un concept de connotation plus étroit et plus technique.

### Définition III.8 : connotation lexicographique

|| Nous appelons *connotation lexicographique* d'une lexie L une caractéristique sémantique 'σ' que L attribue à son référent et qui n'appartient pas à la définition lexicographique de L, c'est-à-dire que (L) ⊄ 'σ'.

Une lexie L peut avoir plusieurs connotations; elles sont réunies et placées dans une sous-zone spéciale de la zone sémantique, immédiatement après la définition de L. Par exemple, le lexème VENTI.1 a quatre connotations; nous allons les présenter ci-dessous :

## VENT

### Définition

I.1. *Vent de X* = 'Phénomène atmosphérique – déplacement horizontal perceptible de l'air<sup>1</sup> venant de X, X étant un lieu géographique ou un des 'points cardinaux<sup>b</sup>'.

### Connotations

- 1) 'rapidité'
- 2) 'changement d'orientation'
- 3) 'force motrice'
- 4) 'absence de contenu'

Ces connotations sont nécessaires pour rendre explicites les liens sémantiques qu'un locuteur français établit entre le lexème VENTI.1 et certaines expressions phraséologiques. Par exemple, on dit *partir, arriver, accourir, disparaître, ... en coup de vent; s'habiller, faire ses devoirs, manger, ... en coup de vent*. De ces exemples, se dégage le phrasème 'EN COUP DE VENT'<sup>1</sup> :

## 「EN COUP DE VENT」

### Définition

[X(P)] 「en coup de vent」 = 「[X a accompli l'action P] très RAPIDEMENT」.

La composante (X a accompli l'action P) est nécessaire pour rendre compte des contrastes du type suivant : \**Il a travaillé en coup de vent* [une action mais non accomplie] ~ *Il a fait son travail en coup de vent*; \**Il a couru en coup de vent* ~ *Il est accouru en coup de vent*; \**Il mange en coup de vent* ~ *Il a mangé en coup de vent*; \**Les papiers ont brûlé en coup de vent* [accompli, mais pas une action] ~ *Il a brûlé les papiers en coup de vent* ~ *Les papiers ont été brûlés en coup de vent*. Notons que cette composante n'est qu'une contrainte sémantique imposée sur la variable P, voir la sous-section 3.3.1.5.1, p. 105.

Pour un francophone, il existe un lien sémantique évident (voir *pont sémantique* au Ch. IV, 1.1, p. 157) entre «ventL.1» et «en coup de vent」: c'est l'élément sémantique «rapidité»; le DEC reflète ce fait par la première connotation de VENTI.1. Cette même connotation sert également de pont entre VENTI.1 et le phrasème 「FILER COMME LE VENT」.

La rapidité est donc une caractérisation sémantique que le lexème VENTI.1 attribue au vent, mais qui ne fait pas partie de la définition de ce lexème. En effet, en ajoutant «rapide» à la définition de VENTI.1, nous rendrions cette définition incompatible avec des phrases toutes naturelles du type de (31) :

(31) *Dans l'air épais, le vent soufflait à peine, et les feuilles paraissaient immobiles.*

De telles phrases montrent que la rapidité, tout en étant un trait caractéristique du vent en français, n'est pas une composante sémantique définitoire de VENTI.1.

Les autres connotations s'expliquent de la même façon. Ainsi, la connotation 2 (changement d'orientation) est justifiée, entre autres, par l'existence du phrasème 「LE VENT TOURNE <CHANGE>」:

「*Le vent tourne <change>*」 = 「L'attitude sociale face à ce dont il est question CHANGE D'ORIENTATION」.

[*Il y a encore deux ans la mode était aux mini-jupes; maintenant le vent tourne.*]

Les autres phrasèmes qui sont liés à VENTI.1 par la même connotation sont 「ALLER OÙ LE VENT SOUFFLE」, 「SAVOIR PRENDRE LE VENT」, 「TOURNER À TOUT VENT」, etc.

À l'appui de la connotation 3 (force motrice), nous pouvons citer 「AVOIR LE VENT EN POUPE」, 「ALLER AU GRÉ DES VENTS」, 「CONTRE VENTS ET MARÉES」, etc. À titre d'exemple, citons la définition d'un de ces phrasèmes pour montrer la présence de la composante (force) :

X「*a le vent en poupe*」 = «X, qui poursuit un but social, bénéficie d'appui d'une FORCE sociale qui le fait avancer vers ce but」.

[*Le parti réformateur a maintenant le vent en poupe.*]

La connotation 4 (absence de contenu) assure le lien avec les phrasèmes du type « C'EST DU VENT ! », « FAIRE DU VENT », « AVOIR DU VENT DANS LA TÊTE », etc. :

« C'est du vent ! » = « Le texte dont il est question et qui est censé avoir du contenu est sans contenu de valeur ».

[*Les promesses électorales du parti réformateur, ce n'était que du vent ! Sa thèse, c'est simplement du vent. Ce qu'il raconte de ses exploits, c'est du vent.*]

Tout comme la connotation de rapidité, les trois autres connotations ne peuvent être incluses dans la définition de VENTI.1 – car il n'y a pas de contradiction dans les syntagmes *les vents constants* et *un vent faible*, ce qui exclut CHANGEMENT et FORCE comme composantes définitoires; quant à ABSENCE DE CONTENU, il n'y a aucun moyen de le rattacher à la définition.

Comme on l'a vu dans les exemples ci-dessus, il n'est pas toujours immédiatement évident de décider si un élément sémantique « $\sigma$ » porté par la lexie L est une composante de la définition de L ou forme une de ses connotations. Pour aider le lexicologue dans sa réflexion à ce sujet, nous proposons les deux critères suivants.

### Critère 10 : Critère de sens antonymique

Soit une lexie L et l'élément sémantique « $\sigma$ » « problématique »; plus précisément, on cherche à savoir si « $\sigma$ » appartient ou non à la définition de L. La notation « $\bar{\sigma}$ » représente un sens antonymique de « $\sigma$ ».

|| Si l'adjonction à la lexie L d'un modificateur avec le sens « $\bar{\sigma}$ » crée une contradiction logique, alors « $\sigma$ » est une composante de la définition de L; dans le cas contraire, « $\sigma$ » est une connotation de L.

En fait, c'est ce critère que nous avons utilisé dans la discussion des exemples précédents. Ainsi, si L = VENTI.1 et « $\sigma$ » = «changeant (de direction)», alors « $\bar{\sigma}$ » = «constant»; comme l'expression *un vent constant* ne crée pas de contradiction logique, le sens «changeant (de direction)» ne relève pas de la définition et est décrit comme connotation.

### Critère 11 : Critère de « mauvais état »

Soit une lexie L référant à un objet destiné à être utilisé d'une certaine façon, et l'élément sémantique « $\sigma$ » « problématique » qui décrit un aspect important de l'utilisation du référent REF de L [= REF(L)]. Par exemple, pour BRASI.1a «bras humains», « $\sigma$ » = «action physique».

|| Si la phrase signifiant «REF(L) est en mauvais état» entraîne «REF(L) fait « $\sigma$ » mal ou pas du tout», alors « $\sigma$ » est une composante de la définition de L; dans le cas contraire, « $\sigma$ » est une connotation de L.

À titre d'exemple, considérons les lexèmes TÊTEI.1a ('partie supérieure du corps'<sup>II.1a</sup> d'une personne...) et CŒURI.1a ('organe principal de la circulation sanguine d'une personne...'). Le français associe au premier le sens 'organe de la raison' et au deuxième, le sens 'organe des sentiments' : *Sa tête disait oui, mais son cœur n'y était pas*. Est-ce que les éléments sémantiques 'organe de la raison' et 'organe des sentiments' doivent être traités de façon identique ? En d'autres mots, appartiennent-ils tous les deux aux définitions ou aux connotations ? L'application du Critère 11 nous donne le résultat suivant : 'organe de la raison' fait partie de la définition de TÊTEI.1a, mais 'organe des sentiments' est une connotation de CŒURI.1a. En effet, la phrase :

(32) *J'ai des maux de tête.*

peut se compléter par ... *et je ne peux plus penser* (<et j'oublie tout>). Si nous y appliquons le Critère 11, L = TÊTEI.1a et 'σ' = 'organe de la raison', nous voyons qu'il est respecté; en effet, la phrase (32) signifie que ma tête est en mauvais état et elle entraîne tout naturellement la continuation proposée. Par conséquent, 'σ' = 'organe de la raison' doit être une composante de la définition de TÊTEI.1a.

Par contre, la phrase :

(33) *J'ai une maladie de cœur.*

ne peut se compléter par ... *\*et je ne peux plus aimer* (<... \*et je suis devenu indifférent aux passions>). En effet, le mauvais état du cœur n'entraîne nullement l'incapacité d'aimer, par conséquent, il y a entorse au critère 11 (L = CŒURI.1a et 'σ' = 'organe des sentiments'); le sens 'σ' = 'organe des sentiments' est donc une connotation de CŒURI.1a.

Un exemple plus complexe pourrait être le lexème ÉPAULE. Comme première approximation, nous pouvons dire que ÉPAULE de X se définit comme 'partie supérieure horizontale du corps'<sup>II.1d</sup> de X située à côté du cou, à laquelle est attaché le bras<sup>I.1a</sup>. Du point formel, cette définition est tout à fait acceptable, conformément aux Règles 1-5 d'élaboration de la définition lexicographique. Cependant, elle ne satisfait pas au Critère 6 (« Pertinence linguistique des composantes sémantiques »), qui aide à rendre compte des liens sémantiques entre des lexies apparentées : cette définition ne prend pas en charge le lien évident avec, par exemple, 'PORTER SUR SES ÉPAULES'<sup>1</sup> (dans les deux sens : physique et métaphorique, c'est-à-dire, '≈ avoir une grosse responsabilité'), ÉPAULER (≈ appuyer) ou 'DONNER UN COUP D'ÉPAULE'<sup>1</sup> (≈ appuyer). Il faut donc enrichir la description du sémantisme de ÉPAULE par les éléments sémantiques 'porter' [physique] et 'appuyer' [métaphorique]. Il reste à décider si nous devons les inclure dans la définition ou dans les connotations. Utilisons encore une fois le Critère 11. La phrase (34a) est normale, tandis que (34b) relève du jeu de mots :

(34) a. *J'ai mal à l'épaule, et je ne peux rien porter.*

b. *\*J'ai mal à l'épaule, et je ne peux pas t'épauler* (<t'appuyer>) *pour cette décision.*

En conclusion, le sens 'porter' doit être intégré dans la définition, qui se lira 'partie supérieure horizontale du corps<sup>19</sup>II.1d de X située à côté du cou, à laquelle est attaché le brasI.1a et qui peut être utilisée par X pour porter quelque chose dessus'. Quant au sens 'appuyer' [métaphorique], il ne représente qu'une connotation.

Il convient de signaler que les deux critères de connotation introduits ci-dessus n'ont aucune prétention ni à la généralité ni à l'exhaustivité. Surtout le Critère 11 est très spécifique, n'étant applicable qu'aux noms d'organes et d'artefacts. Cependant, malgré leur caractère restreint, ces critères sont, comme on vient de le voir, assez efficaces; il serait souhaitable d'en élaborer d'autres.

En résumé, les connotations d'une lexie L sont des informations sémantiques associées à L qui sont 1) extérieures à la définition de L, mais 2) indispensables pour l'indication explicite des liens intuitivement perçus entre des lexies apparentées. Les connotations peuvent varier beaucoup d'une langue à l'autre, et ceci, de façon imprévisible. Cela veut dire que deux lexies appartenant à deux langues différentes, même si elles ont les mêmes définitions (elles sont donc des équivalents idéaux), n'auront pas forcément les mêmes connotations. En russe, par exemple, au lieu de *C'est du vent*, qui serait agrammatical et même incompréhensible, pour une mauvaise thèse, on dira *C'est de l'eau*<sup>19</sup>, ce qui est impossible en français. Bien plus, on a des exemples similaires dans la même langue. En russe (exemple d'Apresjan), les lexies OSĚL ('âne') et IŠAK ('âne, bourricot') désignent le même animal domestique et ont donc une même définition; cependant, leurs connotations sont tout à fait différentes : OSĚL connote la stupidité et le caractère têtu (comme ÂNE en français), et IŠAK, seulement la soumission à une charge de travail excessive (comme BOURRIQUE en français). Cf. :

- (35) a. *uprjamyj kak osël (\*kak išak) ~ têtu comme un âne ('comme une bourrique, \*comme un bourricot)*  
 b. *rabotat' kak išak (\*kak osël) ~ travailler comme une bourrique (\*comme un âne, \*comme un bourricot)*  
 c. *nagružennyj kak išak (\*kak osël) ~ chargé comme une bourrique (comme un âne, comme un bourricot)*

Cela montre clairement la nécessité d'indiquer, dans une description lexicographique qui se veut exhaustive, toutes les connotations d'une lexie.

L'utilisation des connotations permet de résoudre de façon naturelle le problème de lien sémantique à établir entre la lexie L, d'une part, et les phrasèmes et les dérivés qui lui sont sémantiquement et formellement apparentés, d'autre part, alors que la composante qui assure ce lien sémantique ne peut être incluse dans la définition de L. En fait, le recours aux connotations enrichit le concept de

19 De plus, on dit en russe *essorer un texte*, ce qui veut dire 'en éliminer les fioritures inutiles'. Comme on le voit, la présence d'une connotation peut entraîner l'existence de toute une famille d'expressions sémantiquement apparentées.

pont sémantique (voir Ch. IV, 1.1, p. 157) : un pont sémantique n'est plus seulement une composante sémantique (de définition) partagée par les lexies  $L_1$  et  $L_2$ , mais aussi une composante sémantique de  $L_2$  incluant une connotation de  $L_1$ .

### 3.4 Zone de combinatoire syntaxique

De façon générale, la définition d'une lexie vedette  $L$  ne permet pas de prévoir tous les cas de son utilisation syntaxique; ainsi, rien dans la définition du verbe *EMPRUNTER* n'indique qu'il régit la préposition *À* (*emprunter à mon père*), de même que celle de *S'APPROCHER* ne spécifie le régime en *DE* (*s'approcher DE la rive*). À part les usages soumis aux règles générales de la syntaxe,  $L$  peut cependant avoir quantité de propriétés syntaxiques dont la description doit être fournie dans l'article de dictionnaire de  $L$  – dans la zone de combinatoire syntaxique.

On peut répartir ces propriétés syntaxiques en deux grands groupes : celles qui caractérisent la lexie  $L$  elle-même comme dépendant syntaxique et celles qui concernent la façon dont  $L$  détermine le comportement syntaxique de ses dépendants. Les propriétés du premier groupe concernent les capacités de  $L$  d'entrer dans certaines constructions syntaxiques; elles sont plus grammaticales que lexicales, et nous les négligerons ici.

Quant aux propriétés syntaxiques du deuxième groupe, elles visent avant tout les ACTANTS SYNTAXIQUES [= ASynt] de  $L$  et de ce fait sont intimement liées à ses propriétés lexicales. Elles nous intéressent donc.

En conformité avec la théorie Sens-Texte, on distingue les ASynt DE SURFACE et les ASynt PROFONDS. Nous ne pouvons pas expliquer ici toutes les considérations sous-jacentes à la subdivision des représentations syntaxiques en ces deux niveaux (le lecteur intéressé peut consulter les ouvrages indiqués dans la note 4, Ch. I, 1, p. 17); nous nous limiterons donc au minimum d'information indispensable.

Les ASynt de surface [= ASyntS] correspondent aux compléments de type sujet et objets de la grammaire traditionnelle. Dans ce livre, nous ne faisons intervenir des ASyntS que très rarement; nous les mentionnons ici seulement pour mieux faire ressortir la différence avec les ASynt profonds, qui jouent un rôle crucial dans la lexicologie et que nous allons définir maintenant.

#### Définition III.9 : actant syntaxique profond

|| Nous appelons *actant syntaxique profond* [= ASyntP] de la lexie  $L$  un syntagme qui dépend de  $L$  syntaxiquement et en exprime un actant sémantique.

Les ASyntP d'une lexie sont numérotés par des chiffres romains : I, II, III, ...

Le principe de numérotation des ASyntP d'une lexie verbale L est *grosso modo* comme suit : le S(ujet) G(rammatical) de L correspond à I, le CO<sup>dir</sup> ou central à II, les autres CO – par ordre d'importance décroissant – à III, et ainsi de suite.

Les deux précisions ci-dessous éclairent la présentation : pour mettre en relief les propriétés fondamentales du concept en question, nous opposerons les ASyntP aux ASyntS.

D'une part, un ASyntP de L NE CORRESPOND PAS obligatoirement à un ASyntS de L. Il est vrai que, pour la plupart des lexies, une telle correspondance existe; cependant, pour certaines lexies, qui ne sont peut-être pas tellement nombreuses, mais qui sont quand même très importantes dans le lexique, un ASyntP de L correspond à un complément circonstanciel de L (ou même à un modificateur de L, voir le paragraphe suivant). Par exemple, pour le verbe TRAITER (*X traite Y Z-emment* = 'X se comporte envers Y de façon Z'), les expressions circonstancielles *bien/mal, comme un chien, royalement, en ami, etc.* sont des manifestations de l'ASyntP III, qui, lui, exprime l'ASém Z. Pour le verbe SE TROUVER – *X se trouve à Y* (quelque part), le complément circonstanciel de lieu représente l'ASyntP II.

D'autre part, le concept d'ASyntP est généralisé pour couvrir les dépendants syntaxiques des lexies de TOUTES les parties du discours (et pas uniquement du verbe) – à condition, bien entendu, qu'un tel dépendant exprime un ASém de la lexie L. Par exemple, dans l'expression *arrivée de Pierre*, PIERRE est l'ASyntP I de ARRIVÉE; de la même façon, dans *le veto français*, FRANÇAIS est aussi l'ASyntP I de VETO (*le veto de la France; La France a imposé son veto, etc.*), tout comme dans *victoire américaine* (*socialiste, étudiante*) l'adjectif désignant le vainqueur est une manifestation de l'ASyntP I (l'ASyntP I = l'ASém X; *X vainc Y*).

C'est dans les actants syntaxiques profonds que sont concentrées les particularités de cooccurrence irrégulière de L avec ses dépendants.

**NB :** Par *irrégulier*, nous entendons, ici comme ailleurs, un phénomène qu'on ne peut décrire complètement par des règles générales de la syntaxe.

Les dépendants non actanciels de L, c'est-à-dire les modificateurs libres, qui ne découlent pas du sens de L, sont normalement décrits de façon tout à fait satisfaisante par la syntaxe et n'ont nul besoin de description dans le dictionnaire. Par contre, les dépendants actanciels de L ont, de par leur nature, une forte tendance à s'exprimer en fonction de L – de façon idiomatique et capricieuse. Normalement, il faut mémoriser les formes que prennent tel ou tel ASyntP. C'est ce qu'on appelle le *régime* dans la grammaire traditionnelle.

La difficulté du choix de la forme d'un actant syntaxique est connue depuis l'Antiquité; il suffit de penser aux milliers d'élèves qui se cassent la tête pour



employer la bonne préposition ou le bon cas après tel ou tel verbe dans un thème latin. En effet, comment trouver le bon régime ? Les dictionnaires ont toujours essayé de fournir les indications nécessaires, mais aucun ne semble assez complet et systématique sur cette question. Le DEC, par contre, accorde aux problèmes de régime une importance fondamentale et propose, comme solution, l'utilisation de ce que nous appelons le Tableau de régime ou **Régime**.

Le **Régime** d'une lexie L est représenté sous forme de tableau où le nombre de colonnes correspond au nombre d'ASém de L et le nombre de lignes, au nombre maximal de possibilités d'expressions existant pour les ASyntP de L.

### Colonnes du tableau de régime

Chaque ASém de L se voit assigner une colonne particulière. La manchette du tableau identifie chaque colonne par une égalité : l'ASém de la colonne, c'est-à-dire la variable qui le représente, est mis en correspondance avec un ASyntP de L, noté par un chiffre romain. Nous écrivons donc  $X = I$ ,  $Z = III$ , ...

### Lignes du tableau de régime

Les lignes du tableau renferment les renseignements pertinents sur toutes les façons possibles d'exprimer l'ASyntP donné.

En résumé, une case particulière du tableau de régime donne le moyen linguistique (une préposition, une conjonction, une forme morphologique tel un cas grammatical, l'infinitif, le gérondif, ...) qui peut être utilisé pour exprimer l'actant syntaxique profond indiqué dans la manchette de la colonne. Par exemple, si dans la case en question on voit « 3. *de N* », cela signifie que le moyen 3 utilisable pour l'expression de l'ASyntP impliqué est un syntagme nominal introduit par la préposition DE.

En plus du Tableau de régime, la zone de combinatoire syntaxique fournit des **contraintes additionnelles**, qui se rapportent aux deux aspects suivants du régime :

- Les restrictions sémantiques portant sur les moyens d'expression des actants individuels : tel actant doit/ne doit pas être exprimé de telle façon s'il a/n'a pas telles propriétés sémantiques.
- Les restrictions syntaxiques portant sur les combinaisons d'actants ou d'expressions correspondantes, c'est-à-dire sur des combinaisons impossibles/non souhaitables et des combinaisons obligatoires des actants eux-mêmes ainsi que des moyens particuliers de leur réalisation : telle réalisation de tel actant ne peut apparaître à côté de telle réalisation de tel autre actant; ou, au contraire, telle réalisation de tel actant ne peut apparaître sans telle réalisation de tel autre, etc.

Les contraintes sont numérotées pour fins de référence dans les exemples d'énoncés impossibles et non souhaitables (voir plus loin, p. 121).

Pour faciliter la compréhension, le Régime est muni d'*exemples* illustrant l'application de toutes les contraintes (en premier, les exemples positifs, puis les exemples négatifs).

Prenons, à titre d'illustration, le Régime du verbe AIDER1 (comme dans *Il vous aida aussi par son intervention à surmonter cette épreuve*). D'abord, nous donnerons le tableau de régime, suivi des contraintes et des exemples, pour apporter ensuite les explications nécessaires.

**AIDER1**

*X aide Y à Z-er par W* = 'Y étant [en train] (d'essayer) d'effectuer une action Z souhaitable pour Y, || X emploie ses ressources W [de sorte que]<sup>20</sup> W facilite ou rend possible Z pour Y'.

**Régime**

X = I	Y = II	Z = III	W = IV
1. N	1. N	1. à N 2. à V <sub>inf</sub> 3. dans N 4. pour N 5. pour V <sub>inf</sub> 6. pour que Prop <sub>subj</sub>	1. avec N 2. de N 3. par N 4. en V <sub>ant</sub>

- 1) C<sub>III.1</sub> : N désigne une action abstraite [= non physique]
- 2) C<sub>III.3</sub> : N désigne une activité
- 3) C<sub>III.4-6</sub> : N/V/PROP désigne le résultat d'actions non spécifiées
- 4) C<sub>IV.1</sub> : N désigne plutôt un objet concret, surtout pas une action
- 5) C<sub>IV.2</sub> : N ne désigne pas un objet concret; construction extrêmement limitée [N = CONSEIL, ...]
- 6) C<sub>IV.3</sub> : N ne désigne pas un objet concret

20 On notera que la composante [de sorte que] reflète bien la factivité du verbe AIDER1 (pour la factivité, voir 3.3.1.6, p. 107, n° 13) : *Jean a aidé Marie à préparer le repas* entraîne 'le repas est préparé'; cela est aussi vrai pour 'Jean a employé ses ressources W [de sorte que] W a facilité ou rendu possible la préparation du repas'.

- |       |                             |              |
|-------|-----------------------------|--------------|
| 7) a) | $C_{III.1} + C_{II}$        | } impossible |
| b)    | $C_{III.5-6}$ sans $C_{II}$ |              |
| c)    | $C_{III.4} + C_{IV.1}$      |              |
| d)    | $C_{IV.2}$ sans $C_{II}$    |              |
| e)    | $C_{II}, C_{III} = \Lambda$ |              |

- $C_I + C_{II}$  : La Grande-Bretagne aide ses voisins.
- $C_I + C_{III.1}$  : La Grande-Bretagne a aidé à la création de l'ONU  
(= à créer l'ONU).
- $C_I + C_{III.3}$  : La Grande-Bretagne a aidé dans toutes les activités de l'ONU.
- $C_I + C_{II} + C_{III.2}$  : La Grande-Bretagne a aidé l'ONU à réussir ce projet (= dans la réussite de ce projet).
- $C_I + C_{II} + C_{III.4-6}$  : La Grande-Bretagne a aidé l'ONU pour la réussite de ce projet (= pour réussir ce projet, pour que ce projet réussisse).
- $C_I + C_{II} + C_{IV.1-4}$  : La Grande-Bretagne a aidé l'ONU avec ses avions (de ses conseils, par son intervention immédiate = en intervenant immédiatement).
- $C_I + C_{III.1-2} + C_{IV.1,3-4}$  : La Grande-Bretagne a aidé à la création de l'ONU (= à créer l'ONU) avec son argent (par son intervention immédiate = en intervenant immédiatement).
- $C_I + C_{II} + C_{III.2-3} + C_{IV.1,3-4}$  : La Grande-Bretagne a aidé les autres pays à créer l'ONU (= dans la création de l'ONU) avec son argent (par son intervention immédiate = en intervenant immédiatement); Marie l'a aidé pour la préparation du dîner avec la vaisselle.

**Impossible :** \*La Grande-Bretagne a aidé au bétonnage des pistes [Contrainte 1; la forme correcte : ... à bétonner les pistes]

\*Marie l'a aidé dans la cuisson de ce (= pour cuire le) dîner [Contrainte 2; la forme correcte : ... pour le dîner]

\*Marie l'a aidé pour la préparation du (= pour préparer le) dîner [Contrainte 3; la forme correcte : ... pour le dîner]

\*Marie l'a aidé avec sa participation [Contrainte 4; la forme correcte : ... par sa participation]

*\*Marie l'a aidé de sa participation/de son amour* [Contrainte 5; la forme correcte : ... *par sa participation/par son amour*]

*\*Marie l'a aidé par son argent* [Contrainte 6; la forme correcte : ... *avec son argent*]

*\*La Grande-Bretagne a aidé ses voisins à la création de l'ONU* [Contrainte 7a; la forme correcte : ... *à créer...*]

*\*Marie a aidé pour préparer le dîner* [Contrainte 7b; la forme correcte : ... *l'a aidé pour le dîner*]

*\*Marie l'a aidé pour le dîner avec son argent* [Contrainte 7c; la forme correcte : ... *l'a aidé pour préparer le dîner...*]

*\*Marie a aidé de ses conseils* [Contrainte 7d; la forme correcte : ... *l'a aidé de ses conseils*]

*\*Marie a aidé avec son argent* [Contrainte 7e; la forme correcte : ... *l'a aidé avec son argent/a aidé à la création de la fondation avec son argent*]

Comme on peut le voir, chacun des quatre ASém d'AIDER1 est inscrit dans une colonne où la manchette précise la correspondance entre l'ASém et l'ASyntP qui l'exprime dans une phrase.

La manchette de la première colonne signale que l'ASém X (celui qui utilise ses ressources pour faciliter la tâche de Y) est réalisé par l'ASyntP I (c'est-à-dire, dans le cas du verbe conjugué, par le sujet grammatical); la manchette de la deuxième indique la réalisation de l'ASém Y (celui qui essaie de faire Z) : c'est l'ASyntP II (*grosso modo*, le complément d'objet direct); la manchette de la troisième traite de l'ASém Z (l'action de Y) : sa réalisation est l'ASyntP III (*grosso modo*, un complément d'objet prépositionnel ou un complément infinitif) qui l'exprime; et la manchette de la quatrième colonne spécifie la réalisation de l'ASém W (les ressources de X mises à la disposition de Y) par l'ASyntP IV (*grosso modo*, un complément circonstanciel ou un gérondif).

Comme nous l'avons déjà dit, les cases de chaque colonne listent la totalité des moyens de surface admissibles pour exprimer l'ASyntP spécifié : prépositions, conjonctions, cas grammaticaux, infinitif, participe, gérondif, ... Par exemple, dans la troisième colonne (= ASém Z [action visée par Y], ASyntP III), on trouve six réalisations possibles de l'ASyntP III :

- 1) syntagme nominal introduit par la préposition *À* (*Il vous aida aussi par son intervention à l'obtention d'un congé*);
- 2) syntagme verbal infinitif introduit par la même préposition (*Il vous aida aussi par son intervention à obtenir un congé*);

- 3) syntagme nominal introduit par DANS (*Il vous aida aussi dans vos études*);
- 4)-6) constructions avec POUR suivi soit d'un nom (*Il vous aida aussi pour vos études*), soit d'un infinitif (*Il vous aida aussi pour étudier*), soit d'une proposition au subjonctif introduite par la conjonction QUE (*Il vous aida aussi pour que vous fassiez vos études*).

Si l'on examine les moyens de réalisation des ASyntP listés dans le Régime du verbe AIDERI, on constate que trois cas différents peuvent se présenter :

- Les moyens sont sémantiquement et syntaxiquement équivalents (avec N = par N dans la colonne IV : *Jean m'a aidé avec = par ses conseils*).
- Les moyens sont sémantiquement équivalents, mais syntaxiquement en distribution complémentaire – en fonction des propriétés syntaxiques de l'actant (*pour V<sub>inf</sub> = pour N* dans la colonne III : *Jean m'a aidé pour transporter mes valises = pour le transport de mes valises*).
- Les moyens d'expression d'un actant sont limités par certaines propriétés sémantiques de cet actant. Ainsi, « dans N » (colonne III) n'est utilisable que pour l'ASyntP III qui désigne une activité : *Jean m'a beaucoup aidé dans mes études* [ÉTUDES étant une activité] vs \**Hier Jean m'a beaucoup aidé dans le transport de l'équipement* [le transport désignant ici une action plutôt qu'une activité]. Cependant, on peut dire *Depuis dix ans, Jean m'a beaucoup aidé dans le transport de l'équipement*, ce qui est tout à fait régulier, parce que dans cette phrase, le transport désigne l'activité d'un transporteur. (Cet exemple montre clairement dans quelle mesure le régime est sensible aux nuances sémantiques !) Comme autre exemple, considérons la préposition AVEC (n° 1 de la colonne IV). Elle sélectionne surtout des noms qui désignent des objets concrets : *Jean m'a beaucoup aidé avec son argent* (son camion), mais pas \**avec son arrivée rapide* (?avec ses soins). Ces faits sont signalés par des contraintes sémantiques concernant l'applicabilité de certains moyens à certains ASyntP.

En plus de ces contraintes, on trouve encore des contraintes d'un autre ordre : celles qui concernent la cooccurrence des moyens différents entre eux. Par exemple, on peut *aider sa mère pour l'achat d'une voiture* ou *l'aider avec son argent*, mais pas \**aider sa mère avec son argent pour l'achat d'une voiture*. On constate donc que les deux réalisations des ASyntP III et IV, qui sont individuellement parfaitement normales, produisent une combinaison gênante ou impossible.

En résumé, le tableau de régime est accompagné, dans le DEC, de contraintes des deux types :

1. **Des contraintes (sémantiques) traitant de moyens individuels.** On les donne sous forme d'une liste qui suit le tableau de régime. Chaque contrainte est identifiée par l'expression de la forme C<sub>i,j</sub>, qui spécifie la case sur laquelle

cette contrainte porte : « C » est l'abréviation de « colonne », le premier indice (*i*) représente le numéro de la colonne, et le deuxième indice (*j*) celui de la ligne. Par conséquent,  $C_{III,3}$  se rapporte à la troisième case de la troisième colonne, c'est-à-dire que, dans le régime d'AIDER1, il s'agit de l'expression « dans N ».

On comprendra ainsi que la première contrainte dans le **Régime** du verbe AIDER1, soit :

- 1)  $C_{III,1}$  : N désigne une action abstraite [= non physique]

signifie que l'ASyntP III de AIDER1 peut être introduit par la préposition À seulement si cet actant désigne une action abstraite et non un objet matériel ou encore une action physique. En effet, *aider à la création de* (qqch) est tout à fait normal, ainsi que *aider à la fondation du parti*, *à la promotion des nouvelles idées*, *au progrès de la science*, etc., mais *\*aider au bétonnement de la chaussée*, *\*aider à la plantation d'arbres* ou *\*aider à la réparation de ce moteur* sont impossibles ou gênants.

- 2) **Des contraintes traitant de combinaisons de moyens.** Elles peuvent interdire certaines combinaisons de moyens ou, au contraire, en exiger d'autres. Ainsi, la contrainte 7c d'AIDER1 refuse la cooccurrence des syntagmes *pour l'achat* et *avec son argent* comme co-dépendants de ce verbe dans une même phrase. La contrainte 7d interdit l'utilisation du syntagme *de ses conseils* si l'ASyntP II n'est pas exprimé : *\*Il a beaucoup aidé de ses conseils pour la réussite du projet* vs *Il m'a beaucoup aidé de ses conseils pour la réussite du projet*.

Les contraintes individuelles et de combinaison doivent prévoir tous les cas d'expression non libre d'ASyntP : tous les moyens d'expression qui sont sémantiquement limités, ainsi que toutes leurs combinaisons non autorisées ou obligatoires.

Pour faciliter la compréhension du Tableau de régime, le DEC fait suivre les contraintes listées par des exemples illustrant toutes les COMBINAISONS PERMISES des ASyntP avec la lexie vedette. On présente ces exemples de façon systématique : d'abord les combinaisons de la vedette et des ASyntP individuels, puis les combinaisons avec des paires d'ASyntP, puis avec des triplets d'ASyntP, et ainsi de suite, jusqu'aux combinaisons de la vedette avec tous les ASyntP. Les exemples donnés visent avant tout à montrer les possibilités de combinatoire purement grammaticale; ils n'ont donc aucune prétention stylistique.

De la même façon, le DEC liste des exemples de toutes les COMBINAISONS INTERDITES par les contraintes, en précisant, dans chaque cas, la contrainte spécifique qui empêche telle ou telle combinaison. Ces exemples sont négatifs, et comme on le sait très bien, les dictionnaires (ainsi que les grammaires) ont toujours refusé – pour des raisons soi-disant pédagogiques – de donner des exemples négatifs. Or, nous croyons que des exemples négatifs sont utiles et même indispensables en linguistique et, plus particulièrement, en lexicologie.

D'une part, ils sont la seule preuve de la validité des contraintes posées et, en même temps, ils aident l'utilisateur à bien saisir l'essence de la contrainte. Ainsi, l'exemple négatif \**La Grande-Bretagne a aidé au bétonnage* [= ASyntP II] *des pistes* démontre la nécessité de la contrainte 2 sur le régime d'AIDER1 (voulant que l'ASyntP II soit une action abstraite) et la rend plus explicite : *bétonnage des pistes* est une action physique. D'autre part, tout comme les exemples positifs contribuent à la formation de l'intuition du « possible linguistique », les exemples négatifs doivent contribuer à la formation de la bonne intuition de l'« impossible linguistique », cette dernière étant, à notre avis, tout aussi importante que la première.

**NB** : Les exemples (des deux types) sont pour l'instant réservés à l'utilisateur humain; aujourd'hui, un ordinateur qui utilise un DEC d'une façon quelconque ne peut encore en tirer profit. Cependant, rien n'exclut l'élaboration d'un logiciel qui permette, en traitant les exemples, d'en extraire les règles généralisées de combinatoire et de les formuler adéquatement.

### 3.5 Zone de combinatoire lexicale

#### 3.5.1 Présentation des fonctions lexicales

De façon générale, les données sémantiques et les données syntaxiques, même prises ensemble, ne suffisent pas à déterminer entièrement l'utilisation d'une lexie vedette. Il faut, d'une part, spécifier l'ensemble de toutes les possibilités dans le même « paradigme » sémantique. Ainsi, supposons qu'un locuteur veuille parler, par exemple, d'une situation de 'production de marchandises'; son point d'entrée dans le DEC pourrait être la lexie PRODUCTION. Le DEC se doit de lui fournir, sous la vedette PRODUCTION, un inventaire des lexies se rapportant à la situation de production. Un tel inventaire comprend : PRODUIRE, FABRIQUER (des vêtements), CONSTRUIRE (des avions), PRODUCTION, FABRICATION, CONSTRUCTION, CONFECTION, PRODUCTEUR, FABRICANT, CONSTRUCTEUR, INDUSTRIEL, INDUSTRIE, USINE, CHANTIER, PRODUIT, MARCHANDISE, ... Cette liste permet au locuteur de faire une première sélection de la lexie vedette qui correspond le mieux à son idée : il peut retenir PRODUCTION ou préférer une autre lexie. Bien entendu, pour compléter ce choix, le locuteur devra consulter l'article de dictionnaire correspondant où il trouve une description sémantique et syntaxique exhaustive de la lexie choisie. En fait, ce qui est indiqué dans l'inventaire du type présenté sous PRODUCTION, ce sont des (quasi-)synonymes et des dérivés syntaxiques de la lexie PRODUCTION. Pour d'autres situations, un tel inventaire comprend en plus les antonymes et les quasi-antonymes (par exemple, pour TEMPS [état d'atmosphère], il faut inclure *temps* ENSOLEILLÉ vs COUVERT; ORAGE vs ACCALMIE; ...), les

converses et les quasi-converses (pour ACHAT-VENTE, on donne ACHETER vs VENDRE vs PAYER vs DÉPENSER, MAGASIN vs CLIENT, ...), etc. On pourrait parler ici d'aide aux CHOIX PARADIGMATIQUES.

Le DEC doit, d'autre part, prévoir, pour la lexie sélectionnée, les données permettant un choix des cooccurents lexicalement contraints, c'est-à-dire les cooccurents dont la combinatoire n'est déterminée ni par leur sémantisme ni par leurs propriétés syntaxiques. Ainsi, on dit couramment *fort comme un Turc*, mais *jaloux comme un tigre*; on ne peut inverser ces expressions : \**fort comme un tigre*, \**jaloux comme un Turc*. Rien dans le sémantisme ou encore dans la syntaxe ne force ce choix; les expressions correspondantes ne sont pas prévisibles et doivent être apprises. De la même façon, on dit *avoir recours à = faire usage de*, et ici non plus on ne peut inverser (\**faire recours à*, \**avoir usage de*). Comme on le voit, on a, dans un cas, des intensificateurs ('degré élevé', 'beaucoup', 'très') et dans l'autre, des verbes supports (AVOIR, FAIRE). Comme il est impossible de déterminer ces expressions par règles, un DEC doit les recenser systématiquement, en fait, de la même façon que les synonymes, les antonymes, les dérivés, etc. Ici, il s'agit des CHOIX SYNTAGMATIQUES.

Pour faire face à ces deux tâches – le choix du mot juste et la production de la combinaison adéquate, le DEC préconise la méthode des fonctions lexicales.

Du point de vue formel, une *fonction lexicale* [= FL] est une fonction au sens mathématique; elle peut être représentée par la formule traditionnelle :

$$f(x) = y,$$

où  $x$  est l'*argument* de la fonction et  $y$ , sa *valeur*. Par exemple, dans *fort comme un Turc* et *jaloux comme un tigre*,  $f$  est 'intensification',  $x = fort, jaloux$ , et  $y = comme un Turc, comme un tigre$ , de sorte qu'on a :

$$\begin{aligned} \text{'intensification'}(fort) &= \text{comme un Turc} \\ \text{'intensification'}(jaloux) &= \text{comme un tigre} \end{aligned}$$

Ces fonctions sont appelées lexicales parce qu'elles ont comme particularité de n'accepter en tant qu'argument que des LEXIES et en tant que valeur, que des ENSEMBLES DE LEXIES.

Dorénavant, nous appellerons l'argument d'une FL son *mot-clé* (évitant ainsi la confusion potentielle due à l'ambiguïté multiple du terme *argument*).

**NB :** On trouve dans la documentation lexicologique une terminologie parallèle (F. J. Hausmann), où le *mot de base* correspond à notre *mot-clé*, et le *collocatif*, à un *élément de la valeur* (de la fonction lexicale). C'est une terminologie très commode, systématique et mnémotechnique; si nous utilisons quand même *mot-clé* vs *valeur*, c'est que nous voulons mettre en relief l'aspect fonctionnel de la description des collocations que nous proposons.

Du point de vue du contenu, on peut dire qu'une fonction lexicale est *grosso modo* un sens fort spécifique qu'on caractérise par les trois propriétés suivantes :



- Le sens associé à une FL est très abstrait et très général. Une FL peut donc avoir de nombreux arguments, c'est-à-dire que son sens peut se combiner avec plusieurs lexies. Par exemple, l'intensification peut qualifier de très nombreuses lexies (en fait, toutes celles dont le sens inclut une composante de gradation : 'plus' ~ 'moins'). Ce sens assure donc pour une FL un nombre élevé de  $x$ .
- Le sens associé à une FL a un nombre extrêmement élevé d'expressions possibles. Par exemple, l'intensification se manifeste par de très nombreuses lexies : *comme un pot* [sourd]; *nourris, frénétiques, à tout casser, à tout rompre* [applaudissements]; *à plate couture* [battre]; *comme un sourd* [taper]; *largement* [recourir]; *vivement, profondément, énormément, infiniment, de tout cœur* [remercier]; *étroitement, de près* [surveiller]; *d'orfraie* [pousser des cris]; ... Le sens associé à une FL assure donc un nombre élevé de  $y$ .
- L'expression appropriée pour le sens en question est choisie en fonction de la lexie  $L$  à laquelle ce sens se rapporte ( $L$  étant le mot-clé de la fonction lexicale concernée). On peut citer l'exemple classique de Charles Bally, où l'intensification est exprimée en fonction de la lexie « intensifiée » : *grièvement blessé* vs *gravement malade*. On peut encore signaler le cas de *Je vous remercie infiniment* (énormément) vs *merci infiniment* (\*énormément). Ces collocations sont sensibles aux différences lexicales les plus subtiles. Ainsi, l'acception performative de REMERCIER peut être caractérisée par l'intensificateur *énormément*, mais pas son acception descriptive : *Je vous remercie infiniment* vs \**Je l'ai remercié infiniment*. Le sens associé à une FL est donc tel que  $y$  est toujours déterminé par  $x$ .

Soulignons que la description donnée ci-dessus ne traite en fait que d'un seul type de FL, soit des FL *standard simples*. Comme il s'agit du type de FL le plus important, autant quantitativement que qualitativement, dans la section suivante nous ne parlons que des FL standard simples. (Des explications au sujet des FL *non standard* et des FL *complexes* seront données dans la sous-section 3.5.3.)

Les explications ci-dessus nous permettent de rassembler tous les éléments importants du concept de FL standard dans une définition formelle.

### Définition III.10 : fonction lexicale standard

Nous appelons *fonction lexicale standard* [= FL] une fonction  $f$  qui associe à une lexie  $L$  un ensemble de lexies  $f(L)$  tel que les quatre conditions suivantes soient satisfaites :

1. Pour toute paire de lexies  $L_1$  et  $L_2$ , les lexies  $f(L_1)$  et  $f(L_2)$  montrent des relations sémantico-syntaxiques (presque) identiques à ces lexies :

$$\frac{f(L_1)}{L_1} \approx \frac{f(L_2)}{L_2}$$

2. En règle générale,  $f(L_1)$  et  $f(L_2)$  sont différents :  $f(L_1) \neq f(L_2)$ .
3. La fonction  $f$  a un nombre élevé d'arguments (= de mots-clés). [En d'autres mots, le sens ' $f$ ' est très abstrait et très général et s'applique à beaucoup d'autres sens.]
4. La fonction  $f$  a un nombre élevé d'éléments dans sa valeur (= d'expressions).

Pour bien faire comprendre cette définition, nous croyons utile d'illustrer chaque condition par des exemples.

Soit la FL  $f$  = 'intensificateur', mentionnée ci-dessus. Illustrons d'abord la Condition 1. Si  $L_1$  = PLEURER et  $L_2$  = PLUIE, alors on a :

$f(L_1)$  = *amèrement, à chaudes larmes, comme une Madeleine, toutes les larmes de son corps, comme un veau, comme une vache, comme un enfant;*

$f(L_2)$  = *grosse | prépos, diluvienne, torrentielle, violente, ...*

Tout élément du premier ensemble (par exemple, *comme une Madeleine*) se trouve par rapport à PLEURER dans une relation sémantico-syntaxique qui est identique à la relation qu'entretient tout élément du deuxième ensemble (par exemple, *grosse*) avec PLUIE :

$$\frac{\text{comme une Madeleine}}{\text{PLEURER}} \approx \frac{\text{grosse}}{\text{PLUIE}} \approx \dots$$

Bien entendu, nous ne disons pas que *comme une Madeleine* et *grosse* sont sémantiquement ou syntaxiquement équivalents; mais l'expression *comme une Madeleine* remplit par rapport à PLEURER (à peu près) le même rôle que l'adjectif préposé *grosse* par rapport à PLUIE : les deux sont des modificateurs intensificateurs qui signifient – dans le contexte donné – 'beaucoup', 'très', 'intense/intensément'. La proportion donnée ci-dessus peut être prolongée *ad libitum*. Pour être une FL, une dépendance lexicale doit donc donner lieu à un grand nombre de proportions de ce genre.

L'importance de la Condition 2 peut être illustrée comme suit. Si la fonction  $f$  considérée donne lieu à des proportions comme celles ci-dessus (donc, si elle satisfait la Condition 1) mais que l'on a toujours le même numérateur pour des dénominateurs différents, une telle dépendance  $f$  est triviale en  $\mathcal{L}$  : elle n'a pour nous aucun intérêt, puisque le résultat de son application n'est pas une collocation; nous ne voulons pas que  $f$  soit retenue comme FL. Ce n'est qu'un simple cas de signification lexicale. Par exemple, le sens 'cher' [= 'qui est d'un prix élevé'] ne correspond pas à une FL en français, car avec n'importe quelle lexie il peut toujours être exprimé par le même lexème CHER; cela signifie que son expression ne dépend pas de la lexie modifiée :

$$\frac{\text{cher}}{\text{VOITURE}} = \frac{\text{cher}}{\text{VOYAGE}} = \dots$$

Pourtant le sens 'très' (= 'intense') est une FL en français :

- 'très' (*malade*) = *très, gravement* (<\**grièvement*>), mais 'très' (*blessé*) = *gravement, grièvement* (<\**très*>);  
 'très' (*grippe*) = *carabinée*, mais 'très' (*prix*) = *haut, élevé, ...*;  
 'très' (*lutter*) = *sans relâche, à corps perdu*, mais 'très' (*battre*) = *à plate couture*; etc.

Comme on le voit, l'expression de 'très' [= de l'idée d'intensité] dépend de la lexie modifiée. Ce sens correspond, en fait, à la FL **Magn**, voir ci-dessous.

Quant aux Conditions 3 et 4, les exemples cités montrent à l'évidence que le sens 'très', qui correspond à une FL d'intensification, est compatible avec un nombre incalculable de mots-clés et que ces expressions sont aussi fort nombreuses.

### 3.5.2 Inventaire des fonctions lexicales standard simples

Les fonctions lexicales standard simples sont au nombre de **56**; elles couvrent, dans toutes les langues, les relations PARADIGMATIQUES entre les lexies (synonymes, antonymes, dérivés, ...), ainsi que leurs relations SYNTAGMATIQUES (intensificateurs, verbes supports, ...). La vocation des fonctions lexicales est de fournir au locuteur la totalité des moyens lexicaux nécessaires à l'expression la plus riche, la plus variée et la plus complète de la pensée et, en même temps, de garantir le choix le plus précis de la formulation appropriée. En d'autres termes, les FL (ainsi que le reste du DEC) alimentent un système puissant de PARAPHRASAGE, qui est à la fois une raison d'être des FL et un outil fondamental de leur vérification. L'inventaire des FL qui suit cherchera à mettre en relief cet aspect.

#### *Les FL paradigmaticques*

Nous commençons par présenter les FL de base (**1-3**), qui reflètent les relations sémantico-syntaxiques fondamentales : synonymie, conversion, antonymie. Les FL qui suivent (**4-7**) s'apparentent, d'une certaine façon, aux premières (par exemple, **Figur** est un cas particulier de **Syn**). Les FL **8** sont des dérivés syntaxiques (nom déverbatif, adjectif dénominal, etc.).

**1. Synonyme [Syn].** Même si cette notion est fort connue, nous tenons cependant à préciser que, dans un DEC, l'accent est mis sur les synonymes approximatifs (= para/quasi-synonymes) plutôt que sur les synonymes absolus, phénomène assez rare. On distingue trois types de relations entre les sens :

- le sens 'A' est plus spécifique/plus riche que le sens 'B' : 'A'  $\supset$  'B', c'est-à-dire que 'A' inclut tous les sèmes de 'B' et au moins un autre;
- le sens 'A' est moins spécifique/moins riche que le sens 'B' : 'A'  $\subset$  'B', c'est-à-dire que 'B' inclut tous les sèmes de 'A' et au moins un autre;

- les sens 'A' et 'B' ont une intersection non vide : 'A'  $\cap$  'B'  $\neq$   $\Lambda$ , c'est-à-dire que les sens 'A' et 'B' renferment des sèmes communs ainsi que des sèmes différents.

Par conséquent, le DEC indique quatre types de synonymes : absolus (**Syn**), plus spécifiques (**Syn**<sub>⊃</sub>), moins spécifiques (**Syn**<sub>⊂</sub>) et à intersection (**Syn**<sub>∩</sub>).

**Syn**(voiture) = auto(mobile)

**Syn**<sub>⊃</sub>(véhicule) = voiture

**Syn**<sub>⊂</sub>(voiture) = véhicule

**Syn**<sub>∩</sub>(voiture) = camionnette

Il est entendu que les mêmes types d'approximation sémantique s'appliquent aux autres FL. Par conséquent, nous utiliserons la même notation.

**NB** : Par rapport à nos pratiques précédentes, entre autres dans les volumes publiés du DEC français (DEC-1-3), nous changeons ici les sens des indices  $\supset$  et  $\subset$ . La tradition bien établie est de parler de synonymes plus larges ou plus étroits; cependant, cela veut dire qu'on considère le référent et non le signifié des lexies en question. En effet, quand on dit que ANIMAL est un synonyme plus large de MAMMIFÈRE, cela est exact dans la mesure où la classe des animaux inclut celle des mammifères et donc il y a plus d'animaux que de mammifères. Or, ce raisonnement, venant de la logique, ne s'applique pas à la sémantique au sens strict : le lexicologue doit comparer les SENS des lexies et non pas les CLASSES D'OBJETS auxquelles ces lexies réfèrent. Ainsi, le sens 'mammifère' est évidemment plus riche (= plus complexe, plus spécifique) que le sens 'animal'; 'mammifère' inclut 'animal', c'est-à-dire 'mammifère'  $\supset$  'animal'. En parlant de synonymes, il faut donc bannir les termes *plus large/plus étroit*, qui, en faisant intervenir le référent, mettent sur une mauvaise piste. Par contre, les termes *plus spécifique* et *moins spécifique* sont plus appropriés : ils ne concernent que le sens; les indices  $\supset$  et  $\subset$  présentent sans équivoque les relations entre deux sens.

2. **Conversif** [**Conv**<sub>ij</sub>]. Cette FL est liée à la correspondance entre les actants sémantiques et les actants syntaxiques profonds d'une lexie. On dit que la lexie  $L_1$  est un conversif de la lexie  $L_2$  si et seulement si leur sémantisme est identique et les actants SyntP de  $L_1$  sont inversés par rapport à ceux de  $L_2$  (les indices actanciels montrent l'ordre des actants SyntP inversés).

**NB** : 1. Rappelons que les ASyntP sont notés par des chiffres romains – pour les distinguer graphiquement des ASém. Cependant, pour noter les permutations des ASyntP, nous préférons les représenter en indice par des chiffres arabes, puisque c'est plus économique et plus lisible.

2. Le cas très connu de conversifs dans la grammaire est celui de la voix : l'actif et le passif sont des conversifs réciproques.

Illustrons la FL **Conv** par des exemples :

**Conv**<sub>21</sub>(*inclure*) = [*faire partie*]<sup>1</sup>; cela veut dire que les deux phrases suivantes sont synonymes puisque l'ASyntP I de **INCLUDE** est l'ASyntP II de [**FAIRE PARTIE**]<sup>1</sup> et *vice versa* :

(36) a. *Sa présentation* [= ASyntP I] **inclut** ces trois schémas [= ASyntP II].

b. *Ces trois schémas* [= ASyntP I] **font partie** de sa présentation [= ASyntP II].

**Conv**<sub>21</sub>(*effrayer*) = *craindre* [*La défaite m'effraie* = *Je crains la défaite*]

**Conv**<sub>3214, n</sub>(*acheter*) = *vendre*.I.I [*Jean a acheté de Marie une paire de skis pour 100 \$* = *Marie a vendu à Jean une paire de skis pour 100 \$*]

3. **Antonyme** [**Anti**]. Comme cette notion est également fort connue, nous nous limiterons à préciser que la lexie  $L_1$  est un antonyme de la lexie  $L_2$  si et seulement si leurs signifiés sont identiques sauf pour la négation se trouvant « au sein » d'un des deux signifiés. Ainsi, **CONSTRUIRE** [*une maison*] signifie (≈ causer que [la maison] commence à exister); son antonyme, **DÉTRUIRE**, a le sens de (causer que [la maison] commence à NE PAS exister). (Une négation se trouvant « en dehors » d'un signifié ne crée pas, en règle générale, d'antonymie : (ne pas construire) ≠ (détruire).)

**Anti**(*respect*) = *irrespect*                      **Anti**<sub>c</sub>(*mépris*) = *respect*

**Anti**<sub>v</sub>(*désespoir*) = *espoir*                      **Anti**<sub>n</sub>(*aider*) = *gêner*

**Anti** se combine facilement avec d'autres FL (surtout avec **Magn**, **Bon**, **Ver**, **Real**) pour former des FL complexes (dont il sera question plus loin, 3.5.3.1, p. 148) :

**AntiMagn**(*majorité*) = *courte, faible*      [**Magn**(*majorité*) = *forte* | prépos < écrasante]

**AntiBon**(*choix*) = *mauvais* | prépos      [**Bon**(*choix*) = *bon* | prépos]

**AntiVer**(*reprocher*) = *à tort*              [**Ver**(*reprocher*) = *à bon escient*]

**AntiReal**<sub>3</sub>(*ordre*) = *défier*                  [**Real**<sub>3</sub>(*ordre*) = *exécuter*]

**NB** : Illustrons la différence importante entre les conversifs et les antonymes.

**ENVOYER** [*une lettre à Jean*] :

**Conv**<sub>31</sub>(*envoyer*) = *recevoir* [*Pierre envoie une lettre à Jean* ~ *Jean reçoit une lettre de Pierre*]

**Anti**(*envoyer*) = *intercepter* [(≈ causer que la lettre n'arrive PAS)]

**ÉGAL :**

- Conv**<sub>21</sub>(*égal*) = *égal* [*Pierre est égal à Jean ~ Jean est égal à Pierre*; c'est-à-dire, ÉGAL est son propre conversif]  
**Anti**(*égal*) = *inégal*

**ÉPOUX [de Jeanne] :**

- Conv**<sub>21</sub>(*époux*) = *épouse* [*Pierre est l'époux de Jeanne ~ Jeanne est l'épouse de Pierre*]  
**Anti**(*époux*) = *amant* [≈ 'homme remplissant auprès d'une femme la fonction sexuelle d'un époux et qui n'est PAS son époux']

4. **Contrastif [Contr]**. Cette FL spécifie pour la lexie L<sub>1</sub> une expression contrastive consacrée par l'usage (à utiliser dans une figure de rhétorique) :

- Contr**(*d'acier*) = *de velours* [*un regard d'acier vs des yeux de velours*]  
**Contr**(*merl.1*) = *terre*                      **Contr**(*tête1.4*) = *cœur1.4a*  
**Contr**(*ciel*) = *terre*                      **Contr**(*glace*) = *feu*

5. **Épithète pléonastique [Epit]**. C'est un adjectif ou un adverbe ajouté à L<sub>1</sub> sans contribution sémantique dans le cadre d'un cliché :

- Epit**(*océan*) = *immense*              **Epit**(*gagnant*) = *heureux*  
**Epit**(*défier*) = *ouvertement*

6. **Générique [Gener]**. La FL **Gener** détermine pour L<sub>1</sub> un mot générique qui peut apparaître au moins dans une des deux constructions suivantes :

- 1) '**Gener**(L)  $\xrightarrow{\text{ATTR}}$  DER(L)' = 'L' [où DER est un dérivé syntaxique, voir le groupe des FL au n°8 ci-dessous, et ATTR désigne la relation SyntP de modification];

- 2) énumérations du type X<sub>1</sub>, X<sub>2</sub>, ... et autres **Gener**(X).

- Gener**(*gaz*) = *substance* [*gazeuse*]      **Gener**(*république*) = *État* [*républicain*]  
**Gener**(*pistolet*) = *arme à feu* [*fusils, pistolets et autres armes à feu*]  
**Gener**(*armoire*) = *meuble* [*armoires, chaises et autres meubles*]

7. **Figuratif [Figur]** : correspond à un nom métaphorique L<sub>2</sub> codifié par la langue dont la combinaison avec L<sub>1</sub> (**Figur**(L)  $\xrightarrow{\Pi}$  L) est un synonyme (plus riche) de L<sub>1</sub>.

- Figur**(*fumée*) = *rideau* [*de ~*]              **Figur**(*haine1*) = *feu* [*de la ~*]  
**Figur**(*jalousie*) = *démon* [*de la ~*]

**8. Dérivés syntaxiques** : nominalisation [ $S_0$ ], verbalisation [ $V_0$ ], adjectivisation [ $A_0$ ], adverbialisation [ $Adv_0$ ]. Ces FL représentent des dérivés syntaxiques qui ont nécessairement le même contenu sémique que la lexie vedette – bien que leurs caractéristiques syntaxiques soient différentes. Il s'agit du phénomène bien connu que L. Tesnière (1969 : 361 *ssq.*) a fort à propos appelé « translation » :

(37) a. *Il a présenté cette théorie pendant deux heures = Sa présentation [=  $S_0$ ] de cette théorie a duré deux heures.*

b. *Son serment de fidélité à Marie n'était qu'un faux-fuyant = Il a juré [=  $V_0$ ] d'être fidèle à Marie, mais ce n'était qu'un faux-fuyant.*

**NB** : Le verbe JURER est un dérivé syntaxique [=  $V_0$ ] supplétif de SERMENT, tout à fait comme TOMBER est le  $V_0$  supplétif de CHUTE, etc. Notons que, de la même façon, PENDANT =  $Adv_0$ (DURER).

c. *La correction absolue de son comportement facilite l'entente = Son comportement est tout à fait correct [=  $A_0$ ], ce qui facilite l'entente.*

d. *La livraison du courrier était rapide = Le courrier était livré rapidement [=  $Adv_0$ ].*

Ces quatre FL sont réversibles dans ce sens que si  $S_0(\text{Verbe}) = X$ , alors Verbe =  $V_0(X)$  ; c'est-à-dire que si PRÉSENTATION est une nominalisation de PRÉSENTER, alors PRÉSENTER est une verbalisation de PRÉSENTATION. De la même façon, si PENDANT =  $Adv_0$ (DURER), alors DURER =  $V_0$ (PENDANT); ABSOLU =  $A_0$ (<sup>1</sup>TOUT À FAIT<sup>1</sup>) et <sup>1</sup>TOUT À FAIT<sup>1</sup> =  $Adv_0$ (ABSOLU); etc.

Comme les exemples le montrent, il ne s'agit nullement de dérivation morphologique. Les FL « dérivées » ne reflètent que des relations lexicales pures, où les liens morphologiques sont fortuits.

Dans la suite de notre liste, les FL sont regroupées – strictement pour des commodités de lecture – par leur partie du discours : de **9** à **17**, les FL nominales, de **18** à **27**, les FL adjectivales, de **28** à **31**, les FL adverbiales, et de **32** à **56**, les FL verbales.

#### FL nominales

**9. Dérivés sémantiques nominaux actanciels** : nom typique de l'ASyntP I, ou *grosso modo*, « nom d'agent » [ $S_1$ ]; nom typique de l'ASyntP II, ou *grosso modo*, « nom de patient » [ $S_2$ ]; nom typique de l'ASyntP III [ $S_3$ ];... Ces FL représentent des dérivations connues depuis longtemps de la grammaire, mais qui ne sont pas consignées de façon systématique dans les dictionnaires.

$S_1$ (parler) = locuteur

$S_2$ (parler) = paroles, propos, le dit, discours, ...

$S_3$ (parler) = allocutaire, destinataire [ $S_1$  parle à  $S_3$  en lui disant  $S_2$ ]

$S_1$ (louer [un appartement à qqn]) = *propriétaire, fam proprio*

$S_3$ (louer.[un appartement à qqn]) = *locataire*

- (38) a. *Ce qu'il a dit dans notre conversation n'était pas convaincant = Ses paroles [=  $S_1$ ] dans notre conversation n'étaient pas convaincantes.*  
 b. *Jean loue un appartement dans ma maison ≈ Je suis le proprio [=  $S_1$ ] de Jean = Jean est mon locataire [=  $S_3$ ].*

**NB** : Soulignons que les indices actanciels, tels que présentés dans les exemples ci-dessus, sont aussi utilisés systématiquement avec certaines fonctions lexicales de façon similaire, c'est-à-dire pour indiquer l'orientation de la fonction lexicale du mot-clé L par rapport à un ASyntP de L plutôt que par rapport à L. Ainsi, prenons la FL **Magn** – un intensificateur, voir plus loin, et le mot-clé **ACHERETER1**. Appliquée à **ACHERETER1**, **Magn** ne peut intensifier l'action d'acheter elle-même car le sens d'acheter ne contient pas de composante intensifiable; par contre, **Magn** peut intensifier deux de ses ASyntP : la quantité de marchandises achetées (= ASyntP II), ce qui s'écrira **Magn**<sub>2</sub><sup>quant</sup>, ou bien le prix payé (= ASyntP IV), ce qui s'écrira **Magn**<sub>4</sub>. Comme résultat, on a : **Magn**<sub>2</sub><sup>quant</sup>(*acheter1*) = *massivement*, **Magn**<sub>4</sub>(*acheter1*) = *cher*.

**10. Dérivés sémantiques nominaux circonstantiels** : nom d'instrument [ $S_{instr}$ ], nom de lieu [ $S_{loc}$ ], nom de moyen [ $S_{med}$ ], nom de manière [ $S_{mod}$ ] et nom de résultat [ $S_{res}$ ]. Ces FL couvrent les noms typiques des circonstants mentionnés.

$S_{instr}$ (*parler*) = *langue*

$S_{loc}$ (*parler*) = *parloir*

$S_{med}$ (*laver*) = *savon, lessive, détergents, ...*

$S_{mod}$ (*parler*) = *façon [de ~] // parlure*

$S_{res}$ (*laver*) = *lessive [Il était en train de ranger la lessive]*

Le symbole « // » devant le lexème **PARLURE** représente un élément fusionné de la valeur de la FL. Il en sera question plus loin (3.5.4, p. 151); pour le moment, contentons-nous d'indiquer qu'un élément fusionné exprime de façon cumulative le sens de la FL et de son mot-clé, de sorte que ce dernier ne doit plus être repris : *façon de parler vs \*parlure de parler*.

**11. Singulatif [Sing]** : 'unité minimale régulière de...' (correspond en français à un nom).

**Sing**(*flotte*) = *// navire*

**Sing**(*riz*) = *grain [de ~]*

**Sing**(*pluie*) = *goutte [de ~]*

**Sing**(*ail*) = *gousse [d'~]; tête [d'~]*

**Sing**(*fumée*) = *bouffée [de ~]*

**Sing**(*neige*) = *flocon [de ~]*



**12. Collectif [Mult]** : 'ensemble régulier de...' (correspond en français aussi à un nom; représente une FL inverse par rapport à **Sing**).

**Mult**(navire) = // flotte                      **Mult**(chien) = meute [de ~s]  
**Mult**(barbare) = horde [de ~s]              **Mult**(abeille) = essaim, nuée [d'~s]  
**Mult**(oiseau) = volée [d'~x]                **Mult**(poisson) = banc [de ~s]

**13. Nom de chef [Cap]** : 'chef de...' (correspond à un nom).

**Cap**(université) = président              **Cap**(avion) = commandant (de bord)  
**Cap**(théâtre) = directeur

**14. Nom d'équipe [Equip]** : 'équipe de...' (un nom).

**Equip**<sub>3</sub>(université) = corps professoral    **Equip**(avion) = équipage  
**Equip**<sub>3</sub>(théâtre) = troupe

**15. Nom de « démarrage » [Germ]** : 'germe/origine de' (un nom).

**Germ**(colèreI) = ferment, levain [de la ~]  
**Germ**(guerre) = les premiers coups de feu [de la ~]

**16. Nom du centre [Centr]** : 'le centre de...', 'le milieu de...' (un nom).

**Centr**(problème) = cœur [du ~]    **Centr**(âme) = le plus profond [de l'/A<sub>poss</sub> ~]

Cette FL est souvent utilisée dans des FL complexes, par exemple :

**Loc<sub>m</sub> Centr**(hiver, nuit, mêlée) = au cœur [de ART<sub>def</sub> ~]

**Loc<sub>m</sub> Centr**(rue) = en pleine [~]

Pour **Loc<sub>m</sub>**, voir plus loin, n°30.

**17. Nom du point culminant [Culm]** : 'culmination de...' (un nom).

**Culm**(joieI) = combleII [de la ~]    **Culm**(colèreI) = paroxysme [de la ~]

Tout comme **Centr**, cette FL connaît, elle aussi, beaucoup d'emplois avec

**Loc<sub>m</sub>** :

**Loc<sub>m</sub> Culm**(dépression [psychol.]) = // au trente-sixième dessous

On a ici encore une valeur fusionnée de la FL, indiquée par //.

### FL adjectivales

**18. Dérivé sémantique adjectival actanciel** : modificateur adjectival typique de la lexie L<sub>2</sub> en tant qu'ASyntP I/II/III/... de la lexie L<sub>1</sub> - A<sub>1</sub>, A<sub>2</sub>, A<sub>3</sub>, ... ['tel qu'il est...'].

A<sub>1</sub>(mépris) = plein, rempli [de ~]    A<sub>2</sub>(mépris) = couvert [de ~]  
A<sub>1</sub>(chercher) = // en quête [de N]    A<sub>2</sub>(diriger) = // sous la direction [de N]  
A<sub>1</sub>(forme) = sous [~ de N]            A<sub>2</sub>(contrôler) = // sous le contrôle [de N]

**19. Dérivé sémantique adjectival potentiel** : modificateur adjectival typique de la lexie  $L_2$  en tant qu'ASyntP I/II/III/... potentiel de la lexie  $L_1$  – **Able**<sub>1</sub>, **Able**<sub>2</sub>, **Able**<sub>3</sub>, ... ['tel qu'il peut...'], ['tel qu'on peut le...'], etc.].

**Able**<sub>1</sub>(nuire) = nocif    **Able**<sub>1</sub>(peur) = peureux    **Able**<sub>2</sub>(peur) = effrayant  
**Able**<sub>1</sub>(pleurer) = pleurnicheur    **Able**<sub>2</sub>(brûler) = combustible  
**Able**<sub>2</sub>(lire) = lisible

**Able**<sub>1</sub> se combine très souvent avec **Anti** :

**AntiAble**<sub>1</sub>(peur) = courageux    **AntiAble**<sub>2</sub>(brûler) = ignifuge

**20. Dérivé sémantique adjectival virtuel** : modificateur adjectival typique de la lexie  $L_2$  qui entraîne **Able**<sub>1</sub> avec une forte probabilité – **Qual**<sub>1</sub> ['tel qu'il est fort probable qu'il peut...'], ['tel qu'il est fort probable qu'on peut le...'], etc.].

**Qual**<sub>1</sub>(tromper) = malhonnête    [**Able**<sub>1</sub>(tromper) = trompeur]  
**Qual**<sub>2</sub>(tromper) = naïf    [**Able**<sub>2</sub>(tromper) = trompable]

Les FL syntagmatiques

*FL adjectivales*

**21. Intensificateur [Magn]** : modificateur adjectival ou adverbial de la lexie-clé L qui exprime le sens de 'très', 'intense', 'à un degré élevé'.

**Magn**(amour) = ardent, fou    **Magn**(boire) = comme un trou  
**Magn**(fidèle) = comme un chien    **Magn**(souffrir) = atrocement  
**Magn**(fièvre) = de cheval    **Magn**(contrôler) = minutieusement  
**Magn**(peur) = bleue    **Magn**(pleurer) = comme une Madeleine

**22-23. Comparatifs [Plus/Minus]** : expriment les degrés de comparaison; ne s'utilisent qu'avec d'autres fonctions, produisant le plus souvent un verbe signifiant 'devenir plus/moins **Magn**'.

**IncepPredPlus**(ouragan) = se déchaîne  
**IncepPredPlus**(fièvre) = augmente  
**IncepPredMinus**(ouragan) = se calme  
**IncepPredMinus**(fièvre) = baisse, diminue

**24. « Confirmateur » [Ver]** : modificateur adjectival ou adverbial de la lexie-clé L qui exprime le sens 'tel qu'il faut', 'tel que cela sied'.

**Ver**(argument) = valable    **Ver**(succès) = (bien) mérité  
**Ver**(peur) = justifiée    **Ver**(jalousie) = à bon escient

**25. Laudatif [Bon]** : modificateur adjectival ou adverbial de la lexie clé L qui exprime le sens 'bon' en tant qu'approbation subjective du locuteur. La FL **Bon** s'utilise très souvent avec **Anti**, exprimant ainsi la désapprobation du locuteur.

<b>Bon</b> (conseil) = précieux	<b>Bon</b> (se porter) = comme un charme
<b>Bon</b> (repas) = // gueuleton	<b>Bon</b> (choix) = heureux
<b>AntiBon</b> (critique) = virulente	<b>AntiBon</b> (temps) = de chien
<b>AntiBon</b> (film) = // navet	<b>AntiBon</b> (choix) = malheureux

26. **Péjoratif [Pejor]** : exprime le sens 'MinusBon' [≈ 'pire'] et s'utilise surtout dans des FL complexes.

**IncepPredPejor**(santé) = détériore

**IncepPredPejor**(situation) = s'aggrave

**IncepPredPejor**(temps) = se gâte

**IncepPredPejor**(discipline) = se relâche

**CausPredPejor**(joie) = gâcher

**S<sub>1</sub>CausPredPejor**(joie) = rabat-joie

Signalons que **IncepPredPejor** = **Degrad**, n°49.

27. **Positif [Pos<sub>2</sub>]** : exprime le sens d'évaluation positive de l'ASyntP II de la part de l'ASyntP I ; de par sa nature, se combine facilement avec **Anti**.

**Pos<sub>2</sub>**(opinion) = grande | prépos, favorable

**Pos<sub>2</sub>**(compte rendu) = élogieux, favorable, positif

**Pos<sub>2</sub>**(avis) = favorable

**AntiPos<sub>2</sub>**(avis) = défavorable, négatif

#### FL adverbiales

28. **Dérivés sémantiques adverbiaux actanciels** : modificateur adverbial typique de la lexie L<sub>2</sub> en tant qu'ASyntP I/II/III... de la lexie L<sub>1</sub> – **Adv<sub>1</sub>**, **Adv<sub>2</sub>**, **Adv<sub>3</sub>**, ... ['de telle façon que...'].

**Adv<sub>1</sub>**(mépris) = avec [~]      **Adv<sub>2</sub>**(feu [tir]) = sous [le ~ de N]

**Adv<sub>1</sub>**(chercher) = //en quête [de N]      **Adv<sub>2</sub>**(diriger) = //sous la direction [de N]

**Adv<sub>1</sub>**(joieI) = avec [~]      **Adv<sub>2</sub>**(joieI) = à [la ~]

29. **Instrumental [Instr]** : préposition ou expression prépositive exprimant le sens 'avec, par'.

**Instr**(main) = à, avec [la ~], de [la ~]

**Instr**(arguments) = par [ART ~], 'à l'aide' [de ART ~]

**Instr**(téléphone) = par [~]

**Instr**(outil) = avec [ART ~]

30. **Locatif [Loc]** : préposition ou expression prépositive exprimant le sens 'se trouvant dans' [spatialement ou temporellement] (= **Loc<sub>m</sub>**), 'se déplaçant à partir de' (= **Loc<sub>ab</sub>**), 'se déplaçant pour se trouver dans' (= **Loc<sub>ad</sub>**).

**Loc<sub>in/ad</sub>**(gare) = à [ART ~]      **Loc<sub>m</sub>**(personnel) = 'au sein' [de ART<sub>def</sub> ~]

**Loc<sub>in/ad</sub>**(ville) = en [~]      **Loc<sub>in/ad</sub>**(campagne) = à [la ~]

$\text{Loc}_{\text{in}}^{\text{temps}}$ ( <i>dictature</i> )	=	<i>sous</i> [ <i>la</i> ~]
$\text{Loc}_{\text{in}}^{\text{temps}}$ ( <i>période</i> )	=	<i>pendant</i> [ART ~]
$\text{Loc}_{\text{in}}^{\text{temps}}$ ( <i>Antiquité</i> )	=	<i>dans</i> [ <i>l'</i> ~]
$\text{Loc}_{\text{ab}}^{\text{temps}}$ ( <i>Antiquité</i> )	=	<i>depuis</i> [ <i>l'</i> ~]

**31. Consécutif [Propt]** : préposition ou expression prépositive exprimant le sens 'à cause de'.

$\text{Propt}$ ( <i>jalousie</i> )	=	<i>par</i> [~]	$\text{Propt}$ ( <i>maladie</i> )	=	<i>pour cause</i> [ <i>de</i> ~]
$\text{Propt}$ ( <i>fatigue</i> )	=	<i>de</i> [~]	$\text{Propt}$ ( <i>peur</i> )	=	<i>de, par</i> [~]
$\text{Propt}$ ( <i>alcool</i> )	=	' <i>sous l'empire</i> ' [ <i>de l'</i> ~]			

### FL verbales

**32. Pred** : verbe ayant le sens 'être', qui n'apparaît qu'en combinaison avec d'autres FL, voir n<sup>os</sup> 22-23.

Les fonctions verbales suivantes se présentent de façon naturelle par triplets (la raison de ce phénomène sera donnée plus loin).

### 33-35. Verbes supports [Oper, Func, Labor]

Le premier triplet est formé par les FL **Oper**<sub>i</sub>, **Func**<sub>i</sub> et **Labor**<sub>ij</sub>, qui formalisent la notion de *verbe support*, notion bien connue en linguistique française (par exemple, AVOIR et ÊTRE dans *Max a un certain dynamisme* et *Max est en bagarre avec Luc*; voir, entre autres Gross 1981 : 16 *ssq.* et Gross 1989). Ces FL et leurs valeurs sont des verbes sémantiquement vides (ou vidés dans le contexte de leur mot-clé); elles servent à « verbaliser » les noms prédicatifs (c'est-à-dire les noms dont le sens est un prédicat sémantique), en exprimant le mode et le temps. La vocation de ces FL est simplement syntaxique; elles se distinguent entre elles uniquement 1) par le rôle syntaxique joué auprès d'elles par le mot-clé lui-même et 2) par le rôle des actants SyntP du mot-clé.

#### La FL support et son mot-clé

La FL **Oper**<sub>i</sub>, **Func**<sub>i</sub> ou **Labor**<sub>ij</sub> est déterminée par le rôle syntaxique que le mot-clé joue par rapport à elle :

- **Oper**<sub>i</sub> : son mot-clé est son CO<sup>dir</sup> (*faire une ERREUR, recevoir un ORDRE*);
- **Func**<sub>i</sub> : son mot-clé est son SG (*Cet ORDRE vient de..., Cet ORDRE vous concerne*);
- **Labor**<sub>ij</sub> : son mot-clé est son CO<sup>indir</sup> (*soumettre... à une ANALYSE, prendre... en LOCATION*).

Comme on le voit, la trinité des FL supports reflète le fait fondamental de la syntaxe des langues naturelles, qui ne distingue que TROIS rôles syntaxiques majeurs, soit trois types d'actants syntaxiques de surface : le Sujet Grammatical,

le Complément d'Objet direct et les Compléments d'Objet indirect (les CO<sup>indir</sup>, quel que soit leur nombre, ne représentent qu'un même rôle syntaxique majeur).

### La FL support et les ASyntP de son mot-clé

L'indice actanciel de chaque FL support est déterminé par le rôle des actants Synt(axiques) P(rofonds) du mot-clé : l'indice 1 renvoie à l'actant SyntP I du mot-clé, l'indice 2 à l'actant SyntP II, l'indice 3 à l'actant SyntP III, etc.

Plus spécifiquement, avec un **Oper**, l'indice 1 signifie que le SG de cet **Oper** est l'ASyntP I du mot-clé; l'indice 2, lui, signifie que le SG de l'**Oper** est l'ASyntP II du mot-clé ; et ainsi de suite.

Avec les **Func**, la situation est légèrement différente : les indices actanciels ne renvoient pas au SG mais au CO<sup>dir</sup>. L'indice 1 signifie donc que le CO<sup>dir</sup> du **Func** donné est l'ASyntP I du mot-clé, et l'indice 2, que le CO<sup>dir</sup> du **Func** est l'ASyntP II du mot-clé.

Enfin, avec les **Labor**, l'indice actanciel en première position (= i) renvoie toujours au SG du **Labor**, et l'indice en deuxième position (= j), toujours au CO<sup>dir</sup>. Ainsi, l'indice 3 en première position montre que le rôle de SG du **Labor** est rempli par l'ASyntP III du mot-clé, alors que l'indice 2 en deuxième position signale que c'est l'ASyntP II du mot-clé qui remplit le rôle de CO<sup>dir</sup> du **Labor**. Si nous écrivons que SOUMETTRE = **Labor**<sub>12</sub>(INTERROGATOIRE), cela signifie que l'ASyntP I (I correspond à l'indice 1), c'est-à-dire celui qui interroge, est le SG de SOUMETTRE, et l'ASyntP II (II correspond à l'indice 2), c'est-à-dire celui qui est interrogé, est le CO<sup>dir</sup> de SOUMETTRE (*Le juge [= I] a soumis Jean [= II] à une interrogation serrée [= III]*).

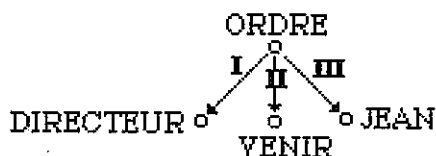
Il faut préciser ici que, pour simplifier l'exposé, nous venons de parler juste de CO<sup>dir</sup>. En fait, il s'agit d'un Complément d'Objet CENTRAL : le CO le plus important pour la lexie en cause. Bien entendu, pour un verbe transitif, le CO central est un CO<sup>dir</sup>, mais ce n'est pas toujours le cas : par exemple, pour une copule, le CO central est le complément prédicatif (DEVENIR N), et pour un verbe intransitif, le CO central peut être un CO prépositionnel (APPARTENIR à N, S'APPROCHER de N, INSISTER sur N, ...).

Comme on a pu le voir, la détermination de la bonne FL pour décrire une collocation donnée « Verbe support ~ nom L » (par exemple, *Leurs PRÉOCCUPATIONS VONT aux luttes d'influence...*) dépend, de façon cruciale, de la configuration des ASyntP en cause. Cela nous amène au problème suivant :

|| Comment déterminer les ASyntP du mot-clé ?

Souvent, la composition et la numérotation d'ASyntP d'une lexie L ne sont pas évidentes, surtout si L est un nom, puisque notre intuition linguistique privilégie avant tout l'analyse actancielle des verbes. Si le mot-clé L est un nom (un adjectif ou un adverbe), il faut chercher un verbe – ou, en cas d'inexistence d'un verbe, une expression verbale – qui décrit la même situation (= la même action, le même événement, le même état, le même comportement, ...) que L de la façon la plus

proche possible; c'est-à-dire qu'on recherche un (quasi-)synonyme verbal de L. On dégage les ASyntP de ce synonyme verbal, qui deviennent du coup ceux du mot-clé L<sup>21</sup>. Ainsi, pour L = ERREUR, le verbe synonyme est SE TROMPER : chaque fois que quelqu'un *se trompe*, il y a une *erreur*, et chaque fois qu'il y a une *erreur*, quelqu'un *se trompe*. Comme on dit, par exemple, *Jean s'est trompé dans ses calculs*, JEAN est l'ASyntP I de SE TROMPER et CALCULS – son ASyntP II. Par analogie, JEAN est l'ASyntP I de ERREUR (*l'erreur de Jean*) et CALCULS – son ASyntP II (*l'erreur dans les calculs*). Par conséquent, dans *FAIRE une erreur*, le verbe FAIRE est **Oper<sub>1</sub>**(*erreur*) puisque celui qui fait l'erreur, c'est-à-dire le SG du verbe FAIRE, est l'ASyntP I de ERREUR. Par contre, dans *RECEVOIR un ordre*, recevoir = **Oper<sub>3</sub>**(*ordre*), puisque celui à qui on ordonne de faire quelque chose (*Le directeur a ordonné à Jean de venir à l'assemblée; Je ferai tout ce que vous m'avez ordonné*) et qui est le SG de RECEVOIR (*Jean a reçu l'ordre de venir...*) est l'ASyntP III de ORDONNER et, par ricochet, de ORDRE :



(*l'ordre du directeur à Jean de venir à l'assemblée*)

**Oper<sub>1</sub>** relie donc l'ASyntP I du mot-clé au mot-clé, de sorte que cet actant devient le SG de **Oper<sub>1</sub>** et le mot-clé lui-même devient son CO<sup>dir</sup> (*Le directeur DONNE à Jean l'ordre de venir...*); **Oper<sub>2</sub>** relie l'ASyntP II du mot-clé, cet actant étant son SG, au mot-clé, qui est son CO<sup>dir</sup> ou, plus précisément, son CO central (*Venir à l'assemblée ÉTAIT l'ordre du directeur*, où ORDRE est le CO central du verbe ÊTRE); et ainsi de suite. **Func<sub>1</sub>** relie le mot-clé, qui est son SG, à l'ASyntP I du mot-clé, qui, lui, est son CO<sup>dir</sup> ou son CO central (*L'ordre de venir à l'assemblée VIENT du directeur*); et ainsi de suite. **Func<sub>0</sub>** n'a pas de complément : cette FL représente les verbes intransitifs absolus dont la signification est approximativement 'avoir lieu' (*La réunion EST EN COURS, Le vent SOUFFLE, La pluie TOMBE, ...*).

<b>Oper<sub>1</sub></b> ( <i>suprématie</i> )	= <i>détenir</i> [ART ~]
<b>Oper<sub>1</sub></b> ( <i>remarque</i> )	= <i>faire</i> [ART ~]
<b>Oper<sub>1</sub></b> ( <i>méfait</i> )	= <i>perpétrer</i> [ART ~]
<b>Oper<sub>2</sub></b> ( <i>danger</i> )	= <i>courir</i> [ART ~]
<b>Oper<sub>2</sub></b> ( <i>applaudissements</i> )	= <i>recueillir</i> [ART ~]

21 Dans le cas où le verbe correspondant n'est pas évident, on doit agir selon la règle empirique suivante : considérer l'expression verbale avec ÊTRE. Ainsi, pour le nom DANGER, on choisira l'expression *X est un danger pour Y*.

**Oper<sub>3</sub>**(conseil) = recevoir [ART ~] [Celui qui reçoit un conseil est l'ASyntP III de CONSEIL parce que l'expression sous-jacente est *X conseille Y à Z*; cf. ORDRE ci-dessus.]

**Func<sub>0</sub>**(réunion) = est en cours

**Func<sub>1</sub>**(aide) = vient [de N]

**Func<sub>1</sub>**(responsabilité) = incombe [à N]

**Func<sub>1</sub>**(problème) = réside [dans N]

**Func<sub>2</sub>**(danger) = menace [N]

**Func<sub>2</sub>**(liste) = comprend, contient [N]

**Func<sub>2</sub>**(interdiction) = frappe [N]

**Labor<sub>12</sub>**(liste) = mettre [N sur ART ~]

**Labor<sub>12</sub>**(estime) = tenir [N en ~]

**Labor<sub>12</sub>**(location) = donner [N en ~] [La situation de LOCATION (de voitures) se présente ainsi : Budget [= I] LOUE des voitures [= II] luxueuses à des millionnaires [= III] au mois [= IV] pour des prix salés [= V].]

**Labor<sub>32</sub>**(location) = prendre [N en ~] [Des millionnaires prennent des voitures luxueuses en location.]

Pour mieux illustrer encore ces FL, nous citerons le cas d'une seule lexie, en l'occurrence ANALYSE (comme dans *analyse du problème par Jean*), accompagnée, en tant que mot-clé, de quatre verbes supports qui sont possibles avec elle :

**Oper<sub>1</sub>**(analyse) = faire [ART ~]

**Func<sub>1</sub>**(analyse) = provient [de N]

**Func<sub>2</sub>**(analyse) = concerne [N], traite [de N], porte [sur N]

**Labor<sub>12</sub>**(analyse) = soumettre [N à ART ~]

### 36-38. Verbes de réalisation [Real, Fact, Labreal]

Le deuxième triplet de FL verbales est formé par les FL **Real<sub>i</sub>**, **Fact<sub>i</sub>** et **Labreal<sub>ij</sub>**, qui expriment *grosso modo* le sens (réaliser les « objectifs » inhérents de la chose [désignée par le mot-clé]). Ces FL sont donc des verbes sémantiquement pleins. Leur syntaxe est cependant identique à celle des FL **Oper<sub>i</sub>**, **Func<sub>i</sub>** et **Labor<sub>ij</sub>**, de sorte que **Real<sub>i</sub>** correspond à **Oper<sub>i</sub>**, **Fact<sub>i</sub>**, à **Func<sub>i</sub>** et **Labreal<sub>ij</sub>**, à **Labor<sub>ij</sub>**. Ainsi, **Real<sub>i</sub>** prend le mot-clé comme son CO<sup>dir</sup>, **Fact<sub>i</sub>**, comme son SG, et **Labreal<sub>ij</sub>**, comme son CO<sup>indir</sup>, les indices actanciels sont déterminés de la même façon que ci-dessus : **Real<sub>i</sub>** a en tant que SG l'actant SyntP I du mot-clé, le SG de **Real<sub>2</sub>** est l'actant SyntP II du mot-clé, etc.

**Real<sub>1</sub>**(film) = donner [ART ~] [Ce soir, à l'« Odéon », on donne « Danton ».]

**Real<sub>1</sub>**(peine) = imposer, infliger [ART ~]

**Real<sub>2</sub>**(peine) = purger [ART ~]

**Real<sub>2</sub>**(examen) = réussir [à ART ~]

<b>Real</b> <sub>2</sub> (piège [littéral])	= <i>tomber</i> [dans ART ~]
<b>Real</b> <sub>2</sub> (piège [figuré])	= <i>donner, tomber</i> [dans ART ~]
<b>Real</b> <sup>I</sup> <sub>2</sub> (invitation)	= <i>accepter</i> [ART ~]
<b>Real</b> <sup>II</sup> <sub>2</sub> (invitation)	= <i>donner suite</i> [à ART ~]
<b>Real</b> <sub>3</sub> (ordre)	= <i>exécuter</i> [ART ~]
<b>Fact</b> <sub>0</sub> (film)	= <i>est à l'affiche</i> [Ce soir, à l'« Odéon », « Danton » est à l'affiche.]
<b>Fact</b> <sub>0</sub> (rêve)	= <i>se réalise</i>
<b>Fact</b> <sup>I</sup> <sub>2</sub> (médecin)	= <i>reçoit, voit</i> [N]
<b>Fact</b> <sup>II</sup> <sub>2</sub> (médecin)	= <i>traite</i> [N]
<b>Fact</b> <sup>III</sup> <sub>2</sub> (médecin)	= <i>guérit</i> [N]

**NB** : On notera que, pour les verbes de réalisation, comme **Real**<sub>2</sub>(invitation) ou, **Fact**<sub>2</sub>(médecin), il peut être indispensable de fournir l'indication du DEGRÉ de réalisation, montrée par des chiffres romains en exposant. Ainsi, une invitation est censée d'abord être acceptée par l'invité, qui ensuite doit faire ce à quoi on l'a invité (= *donner suite à l'invitation*).

<b>Labreal</b> <sub>12</sub> (piège [littéral])	= <i>prendre</i> [N dans ART ~]
<b>Labreal</b> <sub>12</sub> (balle)	= <i>atteindre</i> [N avec ART ~]

### 39-41. Verbes phasiques [Incep, Fin, Cont]

Le troisième triplet de FL verbales est formé par les FL **Incep**, **Fin** et **Cont**, qui expriment les trois phases différentes d'un état ou d'un événement. Ces FL sont donc des verbes sémantiquement pleins qui ont les significations suivantes :

**Incep**(P) = 'commencer à P-er [faire l'action P]'

**Fin**(P) = **Incep**(nonP) = 'cesser de P-er' = 'commencer à ne pas P-er'

**Cont**(P) = **nonFin**(P) = **nonIncep**(nonP) = 'continuer de P-er' = 'ne pas cesser de P-er' = 'ne pas commencer à ne pas P-er'

Le fait que les FL phasiques sont au nombre de trois s'explique par des raisons logico-sémantiques : c'est tout simplement l'existence de seulement trois phases dans chaque événement, à savoir le début, la continuation et la fin. Cela découle du sémantisme de la négation et des possibilités logiques de la placer, comme on le voit dans les égalités ci-dessus.

Les **Incep**, **Fin** et **Cont** ont un caractère aspectuel et, de ce fait, doivent se combiner avec des verbes. Cependant, leur combinatoire avec les verbes français n'a rien d'intéressant du point de vue lexicographique puisque, auprès d'un verbe, ces fonctions sont presque toujours exprimées de façon régulière : **Incep**(chanter) = *commencer* [à ~], **Fin**(lire) = *cesser* [de ~], etc. Il existe quand même quelques cas particuliers : **Incep**(dormir) = // *s'endormir*; **Fin**(dormir) = // *se réveiller*; **Incep**(exister) = // *naître*, **Fin**(exister) = // *mourir*, etc.; mais comme de tels cas sont peu nombreux, nous n'en tiendrons pas compte ici (dans le DEC, ils devront être couverts). Par contre, la combinatoire des FL phasiques avec les noms prédicatifs est très riche et variée; notons cependant qu'elle doit s'exprimer inévitablement en combinaison avec d'autres FL.



Les trois FL phasiques n'ont pas de structure actancielle propre et ont donc besoin de s'appuyer sur les FL actanciennes, telles que **Oper<sub>i</sub>**, **Func<sub>i</sub>** et **Labor<sub>ij</sub>**, ou **Real<sub>i</sub>**, **Fact<sub>i</sub>** et **Labreal<sub>ij</sub>**.

Nous nous voyons obligés d'utiliser encore une fois les FL **complexes**, sans en avoir introduit la notion, ce qui sera fait en 3.5.3.1, p. 148 *ssq.*; pour le moment, il suffit de dire qu'une FL complexe est une combinaison de FL qui s'exprime lexicalement comme un tout.

<b>IncepOper<sub>1</sub></b> (suprématie)	=	<i>arriver [à ART ~], obtenir [ART ~]</i>
<b>FinOper<sub>1</sub></b> (suprématie)	=	<i>perdre [ART ~]</i>
<b>ContOper<sub>1</sub></b> (suprématie)	=	<i>retenir [ART ~]</i>
<b>IncepOper<sub>1</sub></b> (caractère)	=	<i>prendre [ART ~]</i>
<b>IncepOper<sub>1</sub></b> (feu [tir])	=	<i>ouvrir [ART ~ sur N]</i>
<b>ContOper<sub>2</sub></b> (feu [tir])	=	<i>rester [sous ART ~]</i>
<b>IncepFact<sub>0</sub></b> (film)	=	<i>sort [Ce film sort demain à Paris.]</i>

#### 42-44. Verbes causatifs [Caus, Liqu, Perm]

Le quatrième triplet est formé par les FL **Caus**, **Liqu** et **Perm**, qui expriment les trois types de causation d'un état ou d'un événement. Ces FL, qu'on pourrait qualifier de « causatives », sont donc des verbes sémantiquement pleins qui ont les significations suivantes :

**Caus(P)** = 'causer que P [faire en sorte que P a lieu]

**Liqu(P)** = **Caus(nonP)** = 'liquider P' = 'causer que non P'

**Perm(P)** = **nonLiqu(P)** = **nonCaus(nonP)** = 'permettre P' = 'ne pas liquider P' = 'ne pas causer que non P'

Là encore, on retrouve une trinité de FL pour la même raison logique que dans le cas des FL phasiques : le comportement de la négation, c'est-à-dire les deux places possibles pour l'exprimer.

Pour bien comprendre l'usage des FL causatives, il nous faut toucher à deux aspects de leur comportement : d'une part, à leur structure actancielle, et d'autre part, à leur lien avec les sens phasiques.

**FL causatives et structure actancielle.** À la différence des autres FL qui ne changent jamais la structure actancielle (sémantique, ainsi que syntaxique profonde) de la situation décrite par la lexie vedette, une FL causative introduit toujours un nouvel actant sémantique : le causateur. Ce dernier est nécessairement exprimé comme l'actant SyntP I de la lexie vedette L, et par conséquent, l'ASyntP I de départ de L change de numéro. On le voit très bien dans l'exemple banal de la construction causative française, où l'actant SyntP I de départ devient l'ASyntP III :

(39)	<i>Jean</i>	[= I]	<i>écrit</i>	<i>une lettre</i>	[= II]
	vs				
	<i>La mère</i>	[= I]	<i>fait écrire</i>	<i>une lettre</i>	[= II] à <i>Jean</i> [= III].

Quant aux autres ASyntP de départ de L, leur sort dépend de la langue particulière : ils peuvent demeurer intacts ou être soumis à un changement de numéro (nous n'entrerons pas dans ce problème ici). De façon générale, on doit parler du DÉCALAGE des ASyntP provoqué par une FL causative.

Ce décalage est exprimé au moyen des FL verbales vides – **Oper<sub>i</sub>**, **Func<sub>i</sub>** et **Labor<sub>ij</sub>**, c'est-à-dire que nous faisons intervenir les combinaisons (= FL complexes) du type **CausOper<sub>i</sub>**, **LiquFunc<sub>i</sub>**, etc. Pour bien présenter la systématique de cette description, nous allons analyser de façon détaillée un exemple. Soit une lexie vedette : ENVIE, au sens de 'désir causé par un besoin'. Elle a deux actants SyntP : I – celui qui ressent l'envie, et II – l'objet de l'envie; voir la phrase (40) :

(40) *Pierre [= I] a [= Oper<sub>i</sub>] ENVIE d'y aller [= II].*

On remarque que cette phrase peut être « enchâssée » dans une expression causative, par exemple :

- (41) a. *Cette aventure a privé Pierre de l'ENVIE d'y aller.*  
 b. *Cette aventure a ôté à Pierre l'ENVIE d'y aller.*

Ces deux phrases sont équivalentes quant à leur contenu propositionnel, et les verbes PRIVER et ÔTER sont clairement des FL d'ENVIE. Mais comment les décrire par le symbolisme des FL ?

Sémantiquement, les deux verbes veulent dire la même chose : 'CAUSER LA NON EXISTENCE [de l'envie de Pierre d'y aller]'; ce sens doit être exprimé par la FL **Liqu**.

Syntaxiquement, cependant, les deux verbes diffèrent par leur régime, et pour exprimer cette différence, nous avons besoin des FL du type de **Oper<sub>i</sub>**, **Func<sub>i</sub>** et **Labor<sub>ij</sub>**. Plus précisément, *priver Pierre de l'envie d'y aller* s'interprète comme 'causer que Pierre n'a plus d'envie'; comme *a* [= AVOIR] est **Oper<sub>i</sub>** d'ENVIE, *priver* s'écrit **LiquOper<sub>i</sub>**(envie). A son tour, *ôter à Pierre l'envie d'y aller* s'interprète comme 'causer que l'envie n'est plus à Pierre'; *est à* [= ÊTRE à] est **Func<sub>i</sub>** de ENVIE (non admissible comme tel en français), de sorte que *ôter* s'écrit **LiquFunc<sub>i</sub>**(envie).

Un autre exemple permettra de mieux comprendre le fonctionnement de notre système de description. Soit la lexie vedette FORME et la situation de départ décrite par la phrase :

(42) *Cette expression a [= Oper<sub>i</sub>] la forme suivante.*

La phrase

(43) *Alain mettra cette expression sous la forme suivante.*

s'analyse alors comme 'Alain causera que [cette expression a la forme suivante]', et nous écrivons

**CausOper<sub>i</sub>**(forme) = *mettre* [N sous ART ~].

Par contre, la phrase

(44) *Alain donnera à cette expression la forme suivante.*

s'analyse comme 'Alain causera que [la forme suivante est à cette expression]', et nous écrivons donc

**CausFunc<sub>1</sub>**(forme) = donner [à N ART ~]

**FL causatives et FL phasiques.** Comme la causation est intimement liée à la phase du fait causé (on cause soit le commencement, soit la continuation, soit la cessation d'un procès, d'un événement, etc.), nous devrions, pour être rigoureux, toujours indiquer la FL phasique correspondante après une FL causative. Cependant, pour alléger l'écriture, nous adoptons la convention suivante :

|| Étant donné que le cas le plus courant est la causation du commencement, au lieu de **CausIncep**, nous écrivons simplement **Caus**.

Par contre, les autres combinaisons « causation-phase » devront toujours être explicitement marquées; ainsi nous écrivons **CausCont** pour bien indiquer la continuation, et **Liqu**, qui, par définition, est l'équivalent de **CausFin**. Ainsi, pour mettre N sous la forme de... nous aurions dû écrire **CausIncepOper<sub>1</sub>**(forme); de même, *établir la paix* est, strictement parlant, **CausIncepFunc<sub>0</sub>**(paix), que nous écrivons, d'après la convention retenue, comme **CausFunc<sub>0</sub>**(paix). Mais, par exemple, *maintenir la paix* devra obligatoirement s'écrire au complet : **CausContFunc<sub>0</sub>**(paix).

**45. Verbe d'implication [Involv]** : verbe ayant le sens 'impliquer, affecter' qui a le mot-clé L comme sujet grammatical [= SG] et dont le complément d'objet central est le nom de l'entité impliquée dans la situation 'L' sans en être un actant « légitime »; autrement dit, c'est une action de L qui n'est pas prévue par sa définition mais qui est pourtant assez typique de L pour avoir une expression particulière. Très souvent, l'impact décrit par la FL **Involv** est perçu comme négatif, d'où l'emploi d'une *configuration* de FL (voir 3.5.3.2, p. 149) qui comporte **AntiBon** et ainsi exprime l'évaluation.

**Involv**(odeur) = remplit [la pièce]

**Involv**(soleil) = inonde [la pièce]

**AntiBon<sub>2</sub>** + **Involv**(vague) = balaie [la côte]

**AntiBon<sub>2</sub>** + **Involv**(soleil) = brûle [N]

**AntiBon<sub>2</sub>** + **Involv**(voiture) = écrase, fauche, renverse [N humain], accroche, heurte [N inanimé]

**AntiBon<sub>1</sub>** + **Involv**(voiture) = heurte [N], rentre [dans N]

**46. Verbe de manifestation [Manif]** : verbe ayant le sens 'se manifester dans' qui a le mot-clé L comme SG et dont le complément d'objet central est le nom de l'entité dans laquelle 'L' se manifeste. Cette FL se combine souvent avec **Conv<sub>21</sub>**, **Magn/AntiMagn** et **Caus<sub>1</sub>**.

**Manif**(cynisme) = se lit [dans ART<sub>def</sub> yeux/sur ART<sub>def</sub> visage]

**Manif**(sourire) = se dessine, s'esquisse [sur ART<sub>def</sub> lèvres]

**Manif**(joie) = inonde [ART<sub>def</sub> visage]

**Caus<sub>1</sub>Manif**(colère) = montrer [A<sub>poss</sub> ~]

**AntiMagn** + **Manif**(ironie) = suinte [dans ART<sub>def</sub> paroles]

**47. Verbe de préparation [Prepar]** : verbe ayant le sens 'préparer pour' qui a le mot-clé L comme complément d'objet central; se combine souvent avec les verbes supports et avec le verbe de réalisation **Fact<sub>0</sub>**.

- PreparOper<sub>1</sub>**(voyage) = // faire les valises  
**PreparFact<sub>0</sub>**(arme à feu) = charger [ART<sub>def</sub> ~]  
**PreparFact<sub>0</sub>**(voiture) = // mettre de l'essence < faire le plein  
**PreparFact<sub>0</sub>**(programme [inform.]) = charger [ART<sub>def</sub> ~ (dans la mémoire)]

**48. Verbe d'état proche [Prox]** : verbe ayant le sens 'être sur le point de' qui a le mot-clé L comme SG; se combine souvent avec les verbes supports.

- ProxOper<sub>1</sub>**(désespoir) = est au bord <à la limite> [du ~]  
**ProxOper<sub>1</sub>**(mort) = est à l'aube <au seuil> [de la ~]  
**ProxOper<sub>2</sub>**(catastrophe) = est au bord [de la ~]  
**ProxFunc<sub>0</sub>**(tempête) = s'annonce, se prépare

**49. Verbe de dégradation [Degrad]** : verbe ayant le sens 'se dégrader, devenir pire' qui a le mot-clé L comme SG.

- Degrad**(lait) = tourne  
**Degrad**(vin) = s'aigrit; bouchonne; se madérise  
**Degrad**(cœur/I.I) = faiblit  
**Degrad**(discipline) = se perd

**50. Verbe de son typique [Son]** : verbe ayant le sens 'produire le son typique' qui a le mot-clé L comme sujet grammatical.

- Son**(éléphant) = barrit  
**Son**(cochon) = grogne  
**S<sub>0</sub>Son**(cochon) = ouin-ouin  
**Son**(canon) = gronde  
**S<sub>0</sub>Son**(sabre) = cliquetis  
**Caus<sub>1</sub>Manif**(dégoût)  $\xleftarrow{\text{ATTR}}$  **S<sub>0</sub>Son**(dégoût) = pouah !, beurk !  
**Caus<sub>1</sub>Manif**(goût)  $\xleftarrow{\text{ATTR}}$  **S<sub>0</sub>Son**(goût) = mmm ! [Mmm, c'est vachement bon !]

**51. Expression impérative [Imper]** : formule exclamative qui exprime l'ordre ou la prière de faire L tout en étant différente de la forme impérative du verbe V(L).

- Imper**(aller se coucher) = Au lit !      **Imper**(descendre du cheval) = Pied à terre !  
**Imper**(viser) = En joue !      **Imper**(tirer) = Feu !  
**Imper**(secourir) = Au secours !      **Imper**(poursuivre le voleur) = Au voleur !

**52. Verbe résultatif [Result]** : verbe ayant le sens 'être dans l'état qui résulte d'avoir fait L' ; l'indice actantiel spécifie l'ASyntP du mot-clé L qui est son sujet grammatical.

<b>Result</b> <sub>1</sub> (avoir promis)	=	// être lié par A <sub>poss</sub> promesse
<b>Result</b> <sub>2</sub> (avoir promis)	=	// avoir la promesse [de N]
<b>Result</b> <sub>1</sub> (avoir acheté)	=	// posséder [N]
<b>Result</b> <sub>2</sub> (avoir acheté)	=	// être en possession [de N]

Les trois FL suivantes sont des verbes qui ont, le plus souvent, le mot-clé L comme sujet grammatical; ce cas ne sera pas explicitement marqué. Si le sujet grammatical est le nom de la personne affectée, nous utilisons l'indice 2.

**53. Expression de fonctionnement difficile [Obstr]** : verbe ayant le sens 'fonctionner difficilement'.

<b>Obstr</b> (vue)	=	<i>se brouille</i>	<b>Obstr</b> (moteur)	=	<i>hoquète, a des ratés</i>
<b>Obstr</b> (cœurI.I)	=	<i>a des ratés</i>	<b>Obstr</b> (mémoire)	=	<i>a des blancs</i>

**54. Expression de fonctionnement excessif [Excess]** : verbe ayant le sens 'fonctionner d'une façon excessive'.

<b>Excess</b> (sang)	=	<i>bout</i>	<b>Excess</b> (moteur)	=	<i>s'emballe</i>
<b>Excess</b> (cœurI.I)	=	<i>bat la chamade</i>	<b>Excess</b> (yeux)	=	<i>sortent des orbites</i>

Cette FL peut être précisée par les exposants sémantiques suivants :

<b>color</b>	: 'par rapport à la couleur'	<b>motor</b>	: 'par rapport aux mouvements'
<b>dim</b>	: 'par rapport à la dimension'	<b>stat</b>	: 'par rapport à la position verticale'
<b>fulg</b>	: 'par rapport à l'éclat'	<b>trem</b>	: 'par rapport au tremblement'
<b>t°</b>	: 'par rapport à la température'		

**55. Expression d'arrêt de fonctionnement [Stop]** : verbe ayant le sens 'arrêter de fonctionner'.

<b>Stop</b> <sub>2</sub> (vue)	=	<i>perdre [la ~]</i>	<b>Stop</b> (moteur)	=	<i>cale</i>
<b>Stop</b> (cœurI.I)	=	<i>s'arrête, flanche</i>	<b>Stop</b> (cœurI.4a)	=	<i>se brise [Son cœur s'est brisé lors du départ de Pierre]</i>

**56. Expression de symptôme d'un état [Sympt]** : expression verbale ayant le sens 'montrer un symptôme physique d'un état psychologique, ce symptôme étant un état d'une partie du corps ou d'un organe'. Les indices actanciels de la FL **Sympt** sont assignés de la façon suivante :

- l'indice 1 correspond à la partie du corps/organe;
- l'indice 2 correspond à la personne impliquée;
- l'indice 3 correspond à l'état psychologique.

L'ordonnement des indices indique le rôle syntaxique de surface rempli par l'ASyntP correspondant.

Signalons que cette FL, étant donné son sémantisme, ne s'emploie qu'avec les FL **Obstr**, **Excess** et **Stop**.

<b>Excess</b> <sup>color</sup> (visage) – <b>Sympt</b> <sub>13</sub> (colère)	=	<i>être rouge de colère</i> [Le visage de Pierre était rouge de colère]
---	---	--

<b>Excess<sup>color</sup>(visage) – Sympt<sub>23</sub>(colère)</b>	= être rouge de colère [Pierre était rouge de colère]
<b>Excess<sup>color</sup>(visage) – Sympt<sub>213</sub>(colère)</b>	= avoir le visage rouge de colère
<b>Excess<sup>motor</sup>(dents) – Sympt<sub>213</sub>(rage)</b>	= grincer des dents de rage
<b>Excess<sup>motor</sup>(peau) – Sympt<sub>2</sub>(avoir froid)</b>	= // avoir la chair de poule [Pierre avait la chair de poule]

(Sur la FL **Sympt**, voir Iordanskaja 1986.)

### 3.5.3 Les fonctions lexicales « irrégulières »

La description des fonctions lexicales présentée jusqu'ici ne concerne en fait que des fonctions « régulières », plus précisément les FL STANDARD SIMPLES. Cependant, la présentation de la cooccurrence lexicale restreinte d'une langue naturelle ne peut se faire de façon complète si le lexicologue n'a à sa disposition que l'arsenal énuméré dans la section précédente. L'appareil de FL standard simples semble bien approprié à la tâche, mais il s'avère insuffisant. En effet, le lexicologue est contraint d'avoir recours à trois autres variétés de FL, que nous expliciterons ci-dessous : FL **complexes**, **configurations** de FL et FL **non standard**.

#### 3.5.3.1 Fonctions lexicales complexes

Nous avons déjà à plusieurs reprises introduit, dans nos exemples, de nombreuses FL complexes. Cela nous permet de passer immédiatement à la définition, pour ensuite la commenter et l'illustrer.

#### Définition III.11 : Fonction lexicale complexe

|| Nous appelons *fonction lexicale complexe* un enchaînement de FL simples syntaxiquement liées, cet enchaînement ayant une valeur globale cumulative, qui exprime, de façon indécomposable, le sens de l'enchaînement entier.

Une FL complexe est écrite, comme on l'a déjà vu, sous forme d'une chaîne de symboles de FL simples constituantes : **IncepOper<sub>2</sub>**, **AntiReal<sub>3</sub>**, **IncepPred Plus**, etc., c'est-à-dire comme **fg(L)**.

Il est important de noter qu'une FL complexe n'est aucunement une COMPOSITION de FL :

$$fg(L) \neq f(g(L))$$

Ainsi, **IncepOper<sub>1</sub>(amoureux)** = tomber [~]; cependant, si **Oper<sub>1</sub>(amoureux)** = être [~], **Incep(être)** ≠ tomber (plutôt, **Incep(être)** = devenir). Par conséquent,

$$\mathbf{IncepOper}_1(\text{amoureux}) \neq \mathbf{Incep}(\mathbf{Oper}_1(\text{amoureux}))$$

Ou encore,  $\text{Adv}_1 \text{Real}_2^{\text{II}}$  (*invitation*) = à [ART<sub>def</sub> ~] [*Jean est venu à l'invitation de Colette*];  $\text{Real}_1 \text{Real}_2^{\text{II}}$  (*invitation*) = donner suite [à ART<sub>def</sub> ~], mais  $\text{Adv}_1$  (*donner suite*) n'est pas, bien entendu, à.

Une valeur spécifique correspond, de façon idiomatique, à une FL complexe comme un tout; cette valeur ne peut pas être calculée par étape (contrairement à la composition véritable de fonctions, au sens mathématique).

La règle d'utilisation des indices actanciels dans les FL complexes est très simple :

- Avec toutes les FL, sauf les FL dérivationnelles (telles que  $\text{S}_1$ ,  $\text{S}_2$ ,  $\text{Adv}_1$ ,  $\text{Adv}_2$ , etc.), les indices renvoient aux ASyntP du mot-clé (= du mot de base).
- Avec les FL dérivationnelles, les indices renvoient aux ASyntP de la FL précédente. Ainsi, dans  $\text{Adv}_1 \text{Real}_2^{\text{II}}$  (*invitation*), l'indice 2 de **Real** réfère à l'ASyntP II de la lexie INVITATION, c'est-à-dire à l'invité; cependant, l'indice 1 de **Adv** réfère à l'ASyntP I de la FL **Real**, et non pas à l'ASyntP I d'INVITATION – c'est-à-dire, toujours à l'invité.

### 3.5.3.2 Configurations de fonctions lexicales

#### Définition III.12 : configuration de fonctions lexicales

|| Nous appelons *configuration de fonctions lexicales* une suite de FL simples qui ne sont pas syntaxiquement liées entre elles, mais qui ont le même mot-clé, cette suite ayant une valeur globale cumulative qui exprime de façon indécomposable le sens de la suite entière.

Nous ne voulons pas entrer dans une discussion trop technique de ce phénomène compliqué que sont les configurations de FL, et nous nous limiterons à quelques exemples qui, espérons-le, nous permettront de clarifier le concept.

**Bon**(*joie*) + **Magn**(*joie*) = *paradisique* : « une grande [**Magn**] joie qui est très bonne [**Bon**] »

**Magn**(*maladie*) + **A<sub>1</sub>**(*maladie*) = *terrassé* [par ART ~] : « qui a [**A<sub>1</sub>**] une maladie grave [**Magn**] »

**Magn**(*orageII*) + **Caus<sub>2</sub>Func<sub>0</sub>**(*orageII*) = *déchaîner* [ART ~] : « X cause exister [**Caus<sub>2</sub>Func<sub>0</sub>**] un grand [**Magn**] orage au sujet de X »

**AntiBon**(*écoleII.1*) + **Magn**(*écoleII.1*) + **Oper<sub>3</sub>**(*écoleII.1*) = *être* [à rude ~] : « subir [**Oper<sub>3</sub>**] des expériences qui sont intenses [**Magn**] et désagréables [**AntiBon**] »

3.5.3.3 *Fonctions lexicales non standard*

Pour être tout à fait honnêtes, nous devons avouer que les FL standard telles que présentées, même en y ajoutant les FL complexes et les configurations de FL, ne couvrent pas totalement l'immense ensemble des cooccurrences lexicales restreintes. Il existe en effet un nombre imprévisible de locutions semi-figées (= collocations) qui, d'une part, sont strictement du même type que les locutions « lexico-fonctionnelles », mais qui, d'autre part, ne peuvent pas être décrites par les FL standard – puisque leur sens est trop spécifique et donc non généralisable.

Un exemple banal permettra de bien saisir le problème. Ainsi, pour dire 'sans ajout de produit laitier' en se référant à un café, un Français dit *noir*; cependant, pour le thé, le même sens ne peut pas être exprimé de cette façon : 'thé sans lait' ne s'exprime pas par \**un thé noir*. On ne dit pas non plus \**chocolat noir* pour dire 'chocolat chaud sans lait', ni \**café seul*, comme on le fait en espagnol (*café sólo*). Et la question n'est pas encore réglée pour CAFÉ : *café* 'avec du lait' se dit *café au lait*, mais *café* 'avec de la crème', c'est *café crème* (\**café lait*, \**café à la crème*) (cf. les constructions similaires en italien : *caffelatte* ou *caffè latte* vs *caffè macchiato*); *café* 'avec de l'alcool' est appelé *café arrosé* et 'avec du whisky', *café irlandais*. On voit que les sens en cause sont exprimés de façon idiomatique et en fonction du lexème CAFÉ : un thé avec de la crème n'est pas un \**thé crème*, ni un thé avec de l'alcool, un \**thé arrosé*. Tous ces sens sont donc des FL. Cependant, ils sont très précis (= riches), de sorte qu'ils s'appliquent à très peu de lexies (seulement à des noms de boissons), et ne possèdent qu'un nombre infime d'expressions différentes. On voit immédiatement que les Conditions 3 et 4 dans la définition de FL standard (= Définition III.10, p. 127) sont violées. Ces sens sont des FL *non standard*.

Pour le lexicologue, les FL non standard représentent une gêne très sérieuse : étant, de par leur nature même, extrêmement spécifiques et non systématiques, les FL non standard ne sont pas prévisibles et ne peuvent donc pas être dégagées et recensées de façon méthodique. La seule manière de les inventorier, c'est de les retrouver dans les textes. En plus, pour les mêmes raisons, on ne peut les décrire à l'aide d'un formalisme prédéterminé; au contraire, le sens de chacune doit être formulé dans le DEC au coup par coup, en français « traité » – de la même façon que pour les définitions lexicographiques. En effet, chaque description d'une FL non standard est une minidéfinition. L'exemple suivant le montre aisément. Soit l'article de dictionnaire pour le nom STEAK ; on y trouve, entre autres, les FL non standard suivantes :

STEAK, nom, masc.	
(...)	
à peine cuit	: saignant
peu cuit	: bleu
cuit	: à point
S. garni de frites	: ~ frites
S. garni de salade	: ~ salade



**NB** : Signalons que «STEAK TARTARE» est un cas tout à fait différent. Comme un *steak tartare* n'est pas un *steak* ('tranche de bœuf grillé'), cette expression forme un phrasème complet, qui a un article de dictionnaire particulier. Cependant, «STEAK TARTARE» est aussi inscrit sous l'entrée STEAK – dans la zone phraséologique (voir 3.7, p. 153), pour fins de référence.

Les FL non standard sont surtout typiques pour des mots concrets, qui sont normalement très marqués culturellement ou techniquement : noms de nourriture (fromages, vins, ...), de vêtements, de transactions financières, de procédures médicales, d'activités culturelles, politiques ou religieuses, d'armes, de parties du corps, ... Ces fonctions sont extrêmement nombreuses dans le secteur des langues spécialisées (technologie, enseignement, droit, ...). Leur recensement et leur description lexicographique représentent une tâche fort ardue.

Cependant, justement à cause de leur nature spécialisée et non systématique, les FL non standard n'ont presque aucun impact sur la structure de la langue. Ce sont, d'une certaine façon, des nomenclatures techniques. Heureusement donc pour le lexicologue, l'omission ou l'ajout d'une FL non standard représente une opération parfaitement LOCALE, sans répercussion sur l'ensemble des lexies décrites ni surtout sur l'appareil descriptif.

### 3.5.4 Éléments fusionnés des valeurs des fonctions lexicales

Comme le lecteur a pu le découvrir dans les divers exemples, la valeur d'une FL *f* pour le mot-clé *L* renferme souvent, parmi ses éléments, une expression lexicale  $f(L) = L'$  qui ne comprend pas formellement *L* mais dont le sens, lui, comprend à la fois le sens de *L* et celui de *f* de telle sorte que :

$$\langle L' \rangle = \langle L + f \rangle.$$

Un tel élément *L'* de la valeur de la FL en cause est appelé élément *fusionné*; les éléments fusionnés dans la valeur  $f(L)$  sont indiqués par une double barre inclinée //, qui sépare ainsi tous les éléments non fusionnés à sa gauche des éléments fusionnés à sa droite.

Par exemple :

**Magn**(*pluie*) = *forte* | prépos < *violente, torrentielle, ...* // *déluge* [≈ 'pluie violente']

**Magn**(*rire*<sub>v</sub>) = *beaucoup; fort; à gorge déployée* [bruyamment] // **fam** *se bidonner* [≈ 'rire beaucoup']

**AntiBon**(*cheval de course*) = // *tocard* [≈ 'mauvais cheval de course']

**AntiBon**(*film*) = // *navet* [≈ 'mauvais film']

**IncepPred**(*bleu*) = // *bleuir* [≈ 'commencer à être bleu']

**Fin**(*dormir*) = // *se réveiller* [≈ 'cesser de dormir']

\*

\*

\*

L'ORDRE de présentation des FL dans un article de dictionnaire suit celui qui a été suivi dans notre exposé en 3.5 : les FL standard d'abord, les FL non standard ensuite; les FL standard sont ordonnées d'après l'inventaire donné ci-dessus; etc. (Pour les règles détaillées d'ordonnement des FL, voir DEC-2 : 83-84.)

Nous avons ainsi terminé notre présentation de la zone de la combinatoire lexicale restreinte, et nous allons aborder maintenant l'étude des deux dernières zones de l'article de dictionnaire : la zone d'exemples et la zone phraséologique.

### 3.6 Zone d'exemples

Un article de DEC pour la lexie L idéalement rédigé fournit TOUTES les informations lexicographiques pour TOUTES les utilisations de L (cette affirmation ne fait que reprendre les principes descriptifs du DEC). On pourrait donc croire que, dans un tel cas, une zone d'exemples serait superfétatoire. Pour un ordinateur, en fait, ceci est tout à fait vrai. Pourtant, qui peut garantir que les articles de dictionnaire proposés sont vraiment idéaux ? Des oublis et des omissions, des incohérences et même des erreurs sont toujours possibles et risquent de faire trébucher le lexicologue. Des exemples peuvent donc être fort utiles pour pallier ces manquements; ils peuvent aider à redresser les fautes et fournir implicitement des informations complémentaires. Mais même si toutes les données sont parfaites, les illustrations linguistiques apportent un plus : elles permettent l'utilisation maximale de l'intuition linguistique autant du lexicologue que de l'utilisateur. Plus haut, 3.3.1.1, p. 73, nous avons déjà eu l'occasion de souligner l'importance des formulations plus linguistiques des définitions lexicographiques pour une meilleure interprétation par l'intuition linguistique; ici il s'agit du même phénomène – à l'échelle de l'article de dictionnaire en entier. Les illustrations sont donc après tout indispensables.

Dans cette zone, on n'enregistre que des phrases entières, démontrant les emplois standard de la lexie vedette L et appuyant ainsi la description lexicographique proposée de L. Bien entendu, les phrases illustratives doivent être absolument idiomatiques, courantes et naturelles, élégantes et claires, ni trop simples, ni trop compliquées, etc. Elles doivent présenter de façon aussi précise que possible tous les aspects de la syntaxe (= schéma de régime) et de la cooccurrence lexicale restreinte (= fonctions lexicales) de L. De telles phrases sont tirées des bases de données textuelles dont on dispose; mais très souvent il n'est pas possible de les transcrire telles quelles. Tout d'abord, tous les auteurs (et surtout, les « bons » auteurs) se permettent souvent des licences qui sont très belles du point de vue littéraire mais conviennent mal pour illustrer des usages purement linguistiques. Ensuite, on trouve des déviations dialectales et individuelles. Et enfin, même les phrases les plus parfaites n'ont pas été écrites pour servir d'exemple dans un dictionnaire. Le lexicologue se doit donc d'intervenir, de modifier et de transformer les phrases recensées pour les rendre conformes aux nécessités des articles illustrés.

### 3.7 Zone phraséologique

Comme le DEC met un accent particulier sur la cooccurrence et la combinatoire des lexies vedettes, il se doit de répertorier TOUTES les expressions multilexicales non libres où figure la lexie vedette L, c'est-à-dire tous les phrasèmes dont L fait partie. Cette particularité du DEC a déjà été expliquée au Ch. II, 2.2, p. 48; rappelons simplement que le DEC répertorie les phrasèmes des deux façons suivantes :

- Les phrasèmes d'un type particulier – semi-phrasèmes, ou collocations de L, où L préserve au complet son identité lexicale – sont décrits par les fonctions lexicales et consignés dans l'article de dictionnaire de L (voir 3.5 ci-dessus).
- Les autres phrasèmes – les phrasèmes complets (= locutions idiomatiques de toute sorte), où L ne préserve pas son identité, et les quasi-phrasèmes, où l'on trouve un ajout sémantique imprévisible, – sont décrits chacun par un article de dictionnaire autonome. Ainsi, comme nous l'avons déjà signalé, les phrasèmes «RAT DE BIBLIOTHÈQUE», «RAT DE CAVE» et «RAT D'HÔTEL» (ainsi que «STEAK TARTARE») seront des vedettes à part entière et seront décrits par des articles de dictionnaire complets.

Pour des raisons strictement pédagogiques – facilité de consultation du DEC – nous incluons dans l'article de dictionnaire de L la zone phraséologique, qui énumère les phrasèmes complets et les quasi-phrasèmes où L apparaît. Cette liste ne contient que des renvois aux articles de dictionnaire spécifiques et remplit ainsi un rôle mnémotechnique.

La zone phraséologique est délimitée par un losange. Le DEC utilise deux types de losange : un losange blanc ◊ signale que les phrasèmes listés ont un pont sémantique avec la lexie vedette L; un losange noir ♦ signale l'absence d'un tel pont.

Ainsi, dans l'article de dictionnaire BIBLIOTHÈQUE, le phrasème «RAT DE BIBLIOTHÈQUE» sera listé dans la zone phraséologique sous un losange blanc ◊; «RAT D'HÔTEL» sera traité de la même façon sous HÔTEL. Par contre, le renvoi au phrasème «RAT DE CAVE» dans la zone phraséologique de RAT sera précédé d'un losange noir ♦. «STEAK TARTARE» sera répertorié sous un losange blanc dans l'article STEAK, mais sous un losange noir dans l'article TARTARE.



# CHAPITRE IV

## MACROSTRUCTURE

### DU DICTIONNAIRE EXPLICATIF ET COMBINATOIRE

Après avoir présenté la MICROSTRUCTURE du DEC, c'est-à-dire l'organisation interne d'un article de dictionnaire, nous allons étudier sa MACROSTRUCTURE, c'est-à-dire l'organisation de l'ensemble du dictionnaire à partir de ses articles.

Dans le présent chapitre, nous examinons donc deux questions centrales :

- la structure du VOCABLE, qui est décrit par un superarticle du DEC (= le regroupement des lexies ayant le même signifiant et des signifiés apparentés) : section 1;
- la structure du CHAMP LEXICAL (= le regroupement des vocables dont certaines lexies appartiennent au même champ sémantique) : section 2.

#### 1 Structure d'un superarticle de dictionnaire : le vocable

L'unité de base du DEC, comme nous l'avons déjà dit, est la lexie; chaque lexie est décrite dans un article de dictionnaire spécifique. Une lexie ne correspond qu'à un seul article de dictionnaire, et *vice versa*, un article de dictionnaire correspond toujours à une seule lexie. On constate, cependant, que, dans de nombreux cas, des lexies ont un même signifiant et, en plus, manifestent entre elles des liens sémantiques assez évidents. Par exemple, on a en français au moins trois lexies ayant le signifiant /lapɛ̃/ qui sont sémantiquement liées : LAPIN 'animal...' (*Le lapin mange des carottes*), LAPIN 'viande de cet animal' (*une excellente terrine de lapin*) et LAPIN 'fourrure de cet animal' (*un manteau en lapin*)<sup>1</sup>. C'est le phénomène bien connu de la *polysémie*. La polysémie étant un fait linguistique

1 Dans cet exemple, nous ne prétendons nullement régler toutes les questions qu'on peut poser à propos de LAPIN. Mentionnons-en simplement trois :

- Usage hypochoristique : *Comment ça va, mon (petit) lapin ?* Est-ce que cette lexie LAPIN est sémantiquement liée – en français moderne – aux lexies mentionnées ci-dessus ? Il n'est pas facile de donner une réponse convaincante.

- Usage phraséologique : *un chaud lapin, poser un lapin*, etc. Dans des cas comme ceux-ci, ces phrasèmes sont des entrées à part entière, notamment :

「CHAUD LAPIN」'homme trop porté sur le sexe – comme s'il était un lapin mâle'.

[Nous voyons ici le lien sémantique avec la lexie LAPIN-animal, qui doit être étayé par une connotation particulière dans la zone sémantique de cette lexie : 'sexualité excessive'. Remarquons que, par exemple, en anglais, RABBIT, équivalent direct de LAPIN, n'a pas de connotation sexuelle. Pourtant, son synonyme enfantin, BUNNY, en a, mais inversement par rapport au français, la connotation se rapporte aux femmes.]

「POSER UN LAPIN」[à Y] 'ne pas venir au lieu d'un rendez-vous avec Y, où Y est venu conformément à l'entente préalable'.

de base – un trait fort répandu et important de toutes les langues naturelles, le DEC, comme d'ailleurs toute étude lexicologique, se doit d'en tenir compte. Notamment, les lexies montrant la relation de polysémie entre elles seront regroupées en des ensembles appelés *vocables*. Ce regroupement permet de mettre en relief les éléments communs des lexies apparentées et de cette façon de rendre explicite leur parenté.

Un vocable du DEC correspond donc à un article de dictionnaire d'un mot polysème des dictionnaires courants. Ainsi, un dictionnaire traditionnel a un seul article LAPIN, qui regroupe les diverses acceptions de LAPIN, et celles de LAPINE et de toutes les locutions comprenant *lapin*. Le DEC, par contre, aura au moins trois articles différents pour les trois lexies LAPIN illustrées ci-dessus – trois articles différents réunis cependant sous le même vocable LAPIN. Dans le DEC, donc, un vocable n'est pas un article de dictionnaire mais un ensemble d'articles apparentés par leur sémantisme et leur forme – un *superarticle*. (Notons que LAPINE et tous les phrasèmes mentionnés appartiendront à des vocables différents; les lexies correspondantes seront, bien sûr, liées entre elles par les renvois nécessaires.)

La caractérisation d'un superarticle de dictionnaire comporte au moins deux aspects : d'une part, la CONSTITUTION du superarticle, c'est-à-dire les conditions de regroupement des lexies dans un même superarticle – ce qui correspond en fait à la définition du concept de vocable; d'autre part, la STRUCTURE INTERNE du superarticle, c'est-à-dire la détermination de l'ordonnement des lexies formant le superarticle et des « distances » entre ces lexies. Pour pouvoir présenter de façon claire ces deux aspects, nous devons d'abord définir rigoureusement quelques concepts qui y sont sous-jacents – concepts que nous avons déjà évoqués subrepticement à plusieurs reprises auparavant. De plus, il faut également dire quelques mots sur le procédé de présentation que le DEC utilise pour résumer de façon succincte le contenu sémantique d'un vocable.

La présente section est donc subdivisée en cinq sous-sections :

- concepts pertinents (1.1);
- regroupements des lexies en vocables (1.2);
- ordonnancement des lexies d'un même vocable (1.3);
- distances sémantiques entre les lexies d'un même vocable (1.4);
- tableau synoptique (1.5).

---

(suite note 1)

[Ce phrasème n'a aucun lien sémantique avec les lexies LAPIN.]

Quoi qu'il en soit, un phrasème contenant le signifiant / lapē/ ne contient pas la lexie LAPIN, même s'il a un lien sémantique avec elle.

• Forme féminine : dans le DEC, l'expression LAPINE est une lexie à part, parce que, d'abord, la formation du féminin n'est pas tout à fait régulière dans la langue même pour des animaux domestiques (CHIEN ~ CHIENNE, mais TAUREAU ~ VACHE ; CHAT ~ CHATTE, mais BELIER ~ BREBIS ; etc.), et parce que les connotations de LAPINE sont bien différentes (caractère prolifique).

## 1.1 Concepts pertinents

Les discussions antérieures ainsi que les nombreux exemples analysés nous permettent de présenter ici les définitions formelles des quatre concepts de base qui sont indispensables pour la suite de l'exposé.

Le point clé est, bien entendu, la constitution d'un superarticle de dictionnaire. Celle-ci est en fait donnée par la définition du vocable. Pour définir le vocable, il faut introduire d'abord deux concepts sous-jacents : celui de 'pont sémantique' et celui de 'lien direct / indirect entre deux signifiés'.

### Définition IV.1 : pont sémantique

|| Nous appelons *pont sémantique* entre les lexies  $L_1$  et  $L_2$  une composante commune caractéristique des signifiés de ces deux lexies.

### Commentaires

1. Une composante commune est *caractéristique* si et seulement si elle satisfait simultanément les trois conditions suivantes :

- elle n'est pas un outil grammatical ou un élément assimilable (comme les composantes très générales du type 'et', 'ou', 'objet', 'personne', ...);
- elle constitue une partie proportionnellement importante au sein de chacun des signifiés ( $\mathcal{L}_1$ ) et ( $\mathcal{L}_2$ );
- elle a une fonction hiérarchique similaire au sein de chacun des signifiés ( $\mathcal{L}_1$ ) et ( $\mathcal{L}_2$ ).

NB : Nous proposons l'adjectif *caractéristique* comme équivalent de l'américanisme linguistique *non-trivial* (suggestion de J.-L. Descamps).

2. Une composante commune à deux signifiés ( $\mathcal{L}_1$ ) et ( $\mathcal{L}_2$ ) peut être explicite ou bien implicite. Ainsi, dans deux des lexèmes CHANGER, considérés plus haut (p. 58), la composante commune est EXPLICITE : 'commencer à se trouver dans une localisation **différente**', et 'commencer à avoir une caractéristique **différente**'. Cependant, dans de nombreux cas, la composante commune est IMPLICITE : elle est « cachée » dans un sens utilisé dans la définition. Par exemple, le lexème PIED<sup>1</sup>I.1a (pied humain) et le phrasème 「SUR PIED<sup>1</sup>I (Dès l'aube toute la ville était sur pied) s'étant levé de bonne heure... et étant actif debout」 (DEC-2) ont une composante sémantique commune implicite : 'pied<sup>1</sup>I.1a'; en effet, la composante 'pied<sup>1</sup>I.1a' est incluse explicitement dans la définition de 'debout' et est donc implicitement incluse dans celle de 「SUR PIED<sup>1</sup>I. Le phrasème 「SUR PIED<sup>1</sup>I et le lexème PIED<sup>1</sup>I.1a possèdent ainsi un pont sémantique, même si celui-ci n'est pas directement visible.

### Définition IV.2/3 : Lien direct/indirect

|| Nous disons que deux signifiés  $\langle L_1 \rangle$  et  $\langle L_2 \rangle$  sont *liés directement* s'ils ont un pont sémantique.  
 || Au contraire, ils sont *liés indirectement* s'ils n'ont pas de pont sémantique, mais on peut construire la chaîne  $\langle L_1 \rangle, \langle L' \rangle, \langle L'' \rangle, \langle L''' \rangle, \dots, \langle L_2 \rangle$  telle que, dans cette chaîne, n'importe quels deux éléments contigus sont liés directement.

Pour illustrer les définitions IV.2/3, considérons les trois lexies suivantes :

- MINEI            'cavité creusée sous la terre destinée à extraire un minéral'  
 MINEII.1        'cavité creusée sous la terre, située sous une fortification, contenant une charge explosive et destinée à faire sauter cette fortification'  
 MINEII.2        'engin contenant une charge explosive et destiné à faire sauter...'

Les lexies MINEI et MINEII.2 n'ont pas de pont sémantique mais elles sont liées indirectement par l'intermédiaire de MINEII.1 (le pont sémantique entre MINEI et MINEII.1 est 'cavité creusée sous la terre pour...'; le pont sémantique entre MINEII.1 et MINEII.2 est 'contenant une charge explosive et destiné à...') :



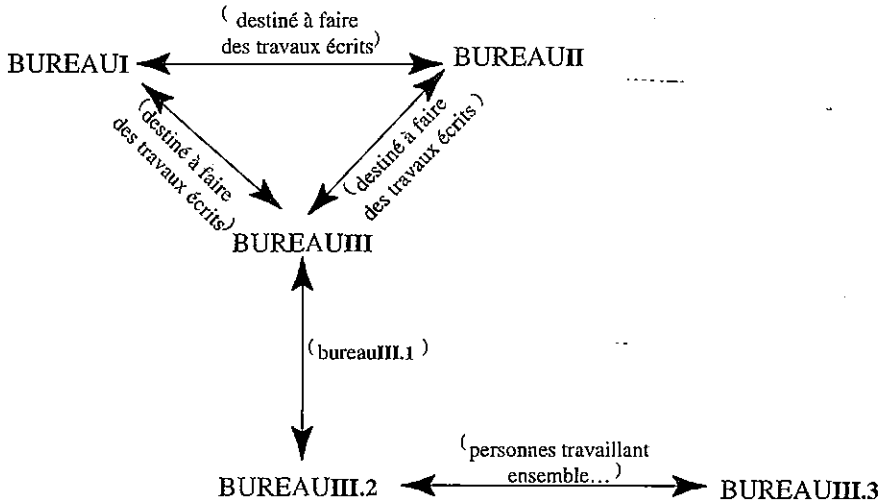
Comme deuxième exemple, prenons les lexies ci-dessous :

- BUREAU I        'table destinée à faire des travaux écrits dessus'  
                    [*bureau de style Empire*]  
 BUREAU II       'pièce destinée à faire des travaux écrits dedans'  
                    [*Je suis entré dans le bureau de l'écrivain*]  
 BUREAU III.1    'organisme destiné à faire des travaux écrits'  
                    [*le 3<sup>e</sup> bureau, le bureau de planification de la mairie*]  
 BUREAU III.2    'personnes travaillant ensemble dans un bureau III.1...'  
                    [*Son bureau lui a offert des fleurs*]  
 BUREAU III.3    'groupe de personnes travaillant ensemble pour assurer la direction d'une association' [*Bureau du parti radical*]

**NB :** Les expressions du type 'bureau de tabac' et 'bureau de poste' ne désignent pas des bureaux; ce sont des phrasèmes complets.

On constate un pont sémantique évident entre BUREAU I, BUREAU II et BUREAU III.1 : 'destiné à faire des travaux écrits'; le pont sémantique entre BUREAU III.1 et BUREAU III.2 est 'bureau III.1'; enfin, le pont sémantique entre BUREAU III.2 et BUREAU III.3 est 'personnes travaillant ensemble'.





Par conséquent, les trois lexies BUREAUI, BUREAUII et BUREAUIII.1 sont directement liées entre elles; BUREAUIII.1 est directement liée à BUREAUIII.2 et à BUREAUIII.3. Mais il n'y a pas de lien sémantique direct entre BUREAUI ('table pour écrire dessus') et BUREAUIII.3 (d'un parti) : ces lexies ne sont liées qu'indirectement – par l'intermédiaire de BUREAUIII.1 et BUREAUIII.2.

Maintenant nous pouvons définir le concept de vocable.

#### Définition IV.4 : Vocable

Un *vocable* est l'ensemble de toutes les lexies  $L_1, L_2, \dots, L_n$  qui satisfont simultanément les deux conditions suivantes :

1. les signifiants de  $L_1, L_2, \dots, L_n$  sont identiques;
2. les signifiés de deux lexies quelconques parmi  $L_1, L_2, \dots, L_n$  sont liés (directement ou indirectement).

En règle générale, il est possible d'identifier parmi les lexies d'un même vocable la *lexie de base*, élément qui est, en quelque sorte, le point de départ pour la description des autres lexies.

#### Définition IV.5 : Lexie de base d'un vocable

La *lexie de base* d'un vocable est une lexie  $L$  telle que les autres lexies du vocable font directement ou indirectement référence à  $L$  alors que  $L$  ne fait aucune référence aux autres lexies du vocable.

Ainsi, dans le vocable MINE, la lexie (en fait, un lexème) MINEI 'cavité creusée sous la terre destinée à extraire un minéral' est la lexie de base : les autres lexies de ce vocable y réfèrent. De même dans le vocable BUREAU, c'est BUREAUI 'table destinée à faire des travaux écrits dessus' qui est la lexie de base.

La lexie de base d'un vocable est donc déterminée dans le DEC par des facteurs strictement sémantiques : nous ne tenons pas compte des considérations historiques (les datations), de la filiation étymologique, de la fréquence ou même de la disponibilité des lexies considérées.

En résumé, nous pouvons constater que le vocable n'est pas autre chose que le regroupement des lexies ayant le même signifiant et des signifiés apparentés; examinons de plus près ce regroupement.

## 1.2 *Regroupement des lexies en vocables*

Tout en maintenant l'exigence d'élémentarité des articles de dictionnaire, nous rejetons, cependant, la notion de dictionnaire comme suite linéaire d'articles de dictionnaire : nous croyons à la hiérarchie sémantique dans le lexique, et nous voulons absolument la refléter dans le DEC. Pour ce faire, nous regroupons certaines lexies sous des *vocables*. Un DEC est donc un ensemble de vocables (dont chacun est évidemment un ensemble de lexies).

Le REGROUPEMENT des lexies en vocable n'est pas du tout automatique mais, comme nous venons de le dire, requiert un lien sémantique assez important entre les lexies concernées – ce que nous avons appelé le pont sémantique : une composante commune caractéristique dans leur définition. Si des lexies à signifiant identique ne possèdent pas de pont sémantique, elles sont considérées comme homographes et consignées dans des vocables différents.

Reprenons notre exemple des trois CHANGER (pp. 58-59). Le premier, dont le sens peut être représenté approximativement par 'commencer à se trouver dans une localisation différente', et le deuxième, dont le sens est 'commencer à avoir une caractéristique différente', ont un pont sémantique évident : 'commencer... différent'; le sens du troisième inclut le sens du premier (*changer Jean de fauteuil* = 'causer que Jean change de fauteuil'), qui constitue de ce fait un pont sémantique. Par conséquent, les trois CHANGER doivent être regroupés dans un même vocable.

Pour illustrer un autre type de regroupement de lexies en vocables, consultons l'entrée ÂNE dans un dictionnaire traditionnel. On y trouve les deux acceptions suivantes (dont nous avons modifié les formulations pour les rapprocher de celles du DEC) :

1. 'Mammifère domestique – bête de somme, plus petit que le cheval, ayant une grosse tête, de longues oreilles et une robe grise'.

2. 'Personne incapable de comprendre quelque chose (dans le domaine Y) alors qu'elle est censée (y) comprendre quelque chose, ce qui est causé par un manque d'intelligence' [*Jean est un âne (en math)*].

D'après l'intuition des locuteurs, ces deux acceptions – ÂNE1 et ÂNE2 – sont apparentées; le DEC se doit donc de les traiter comme deux lexèmes appartenant au même vocable, en conformité avec la démarche habituelle. Mais au lieu d'affirmer simplement ce fait, nous devons dégager FORMELLEMENT le lien sémantique entre les deux lexèmes et ainsi introduire un pont sémantique entre les acceptions 1 et 2 ci-dessus. Dans ce cas, il est très facile de le faire puisque ce lien est établi par comparaison. Nous réécrivons donc 2 comme 2' :

- 2'. 'Personne incapable de comprendre quelque chose (dans le domaine Y) alors qu'elle est censée (y) comprendre quelque chose, ce qui est causé par un manque d'intelligence, – comme si cette personne était UN ÂNE1'.

Le pont sémantique proposé – mis en relief en petites majuscules – s'appuie sur une des CONNOTATIONS du lexème ÂNE1 : 'stupidité'; voir Ch. III, 3.3.2, p. 111.

Nous croyons que cette démarche suit fidèlement l'intuition naïve d'un locuteur francophone; interrogé, pourquoi a-t-il appelé Pierre un âne, un lion ou un chien, il répondra, en toute probabilité, que c'est parce que Pierre ressemble, d'une certaine façon, à un âne, à un lion ou à un chien. Il est évident qu'il s'agit d'une comparaison, même si elle est subconsciente. La légitimité de telles comparaisons, leurs bases, la nécessité d'introduire, dans les définitions de départ, des composantes pertinentes, etc. sont des questions importantes – mais hors sujet ici. Nous ne voulons discuter ici que de la PERTINENCE des ponts sémantiques; leur nature précise relève de la discussion des composantes sémantiques en général, qui a été faite au Ch. III, 3.3, p. 72 *ssq.*

Voyons maintenant un cas opposé : deux lexies qui ont un signifiant identique et que la plupart des dictionnaires regroupent sous la même entrée; ce sont PIED ('partie du corps...') et PIED ('unité minimale rythmique d'un vers...'). Dans le DEC, il est absolument impossible de les réunir sous un vocable puisque nous n'avons pas pu dégager de pont sémantique entre ces deux lexèmes. Par conséquent, on a des homonymes PIED<sup>1</sup> et PIED<sup>2</sup>, qui forment deux superarticles. Le vocable PIED<sup>1</sup> comprend 14 lexèmes réunis en trois groupes : PIED<sup>1</sup>I : ('partie du corps...'); PIED<sup>1</sup>II : toutes sortes de support, etc. rappelant le pied; et PIED<sup>1</sup>III : ('unité de longueur – environ la longueur d'un pied<sup>1</sup>I...') (DEC-2). Comme on le voit, tous ces lexèmes sont reliés par le pont sémantique manifesté par PIED<sup>1</sup> ('partie du corps...'). Le vocable PIED<sup>2</sup> ne comprend qu'un seul lexème : PIED<sup>2</sup> 2.

2 Si le lexicologue veut absolument réunir PIED<sup>1</sup> et PIED<sup>2</sup> sous un même vocable, il se doit d'indiquer un pont sémantique entre les deux. Il peut le faire, par exemple, en ajoutant à la définition de PIED<sup>2</sup> la composante '... comme si c'était un pied<sup>1</sup>III – unité de longueur'. Avec un tel ajout, il serait formellement en règle ; mais sur le plan des faits linguistiques, les arguments pour appuyer une telle solution lui feront défaut. Dans des cas semblables, le Critère 6 (pertinence linguistique) est de rigueur.

### 1.3 *Ordonnement des lexies d'un même vocable*

Les deux opérations lexicographiques fondamentales décrites dans le chapitre précédent – la délimitation des lexies et leur description – reflètent sans doute une réalité langagière : nous pensons que les lexies sont bien distinctes dans la conscience linguistique et que chacune est associée à une structure mentale que notre description lexicographique tâche d'extérioriser (le sens d'une lexie est représenté dans le cerveau par des sens plus petits, le tout étant muni d'indications sur la cooccurrence syntaxique, lexicale, etc.). Il en est de même avec le regroupement de lexies en vocable : comme on vient de le voir, ce regroupement suit les intuitions des locuteurs quant aux liens sémantiques entre les lexies. Par contre, quant à l'ordonnement des lexies à l'intérieur d'un vocable, la question devient plus délicate. Dans un dictionnaire à l'usage des humains, en particulier dans un dictionnaire imprimé, y compris un DEC, un ordonnement LINÉAIRE des lexies est inévitable : c'est en fait l'unique façon de visualiser l'information linguistique sur papier. Or, on sait pertinemment que dans la langue, c'est-à-dire dans le cerveau des locuteurs, il n'existe pas d'ordonnement linéaire des lexies. Et ainsi, l'ordonnement des lexies dans un vocable semble ne pas correspondre à une réalité de la langue. Pourtant, même si les lexies ne sont pas linéarisées dans le cerveau, il n'en reste pas moins qu'elles sont stockées HIÉRARCHIQUEMENT, tel que le prouve le phénomène de la disponibilité : les différentes acceptions ne surgissent pas toutes en même temps, il y en a de dominantes qui apparaissent en premier. Malheureusement, ce domaine de la psycholinguistique est insuffisamment exploré, de sorte que nous ne pouvons pas nous baser sur des faits certains. Cependant, l'ordonnement des lexies dans le DEC ne peut être arbitraire ; comme tout dans le DEC, cet ordonnement doit être systématique et cohérent. Pour atteindre cet objectif, il nous faut donc énoncer des règles d'ordonnement des lexies à l'intérieur d'un vocable. Nous voulons que ces règles reflètent au maximum la hiérarchie présumée des lexies appartenant au même vocable dans la conscience linguistique.

La pratique courante des dictionnaires traditionnels est d'ordonner les acceptions à l'intérieur d'un article (ce qui correspond, dans le DEC, à l'ordonnement des lexies à l'intérieur d'un vocable) soit en fonction de considérations historiques (chronologie des attestations), soit en fonction de considérations grammaticales (d'abord les verbes, après les noms, ou *vice versa*), soit en fonction de considérations psychologiques (la fréquence et/ou la disponibilité). Nous ne suivons pas cette façon de faire : nous nous appuyons exclusivement sur des CONSIDÉRATIONS SÉMANTIQUES, ce qui constitue d'ailleurs le leitmotiv de toute notre approche. Nous espérons que le cheminement sémantique, tout en allant à l'encontre des méthodes historiques et grammaticales, reflète cependant le mieux les facteurs psychologiques.

Pour nous, le respect de l'ordonnement logique des lexies se manifeste sous la forme des cinq règles suivantes.

Soit les lexies  $L_1$  appartenant toutes au même vocable. Leur ordonnancement est dicté par les Règles 7-11; les Règles 7 et 8 sont générales, c'est-à-dire qu'elles concernent toutes les lexies, alors que les Règles 9-11 ne s'appliquent qu'à des lexies à sens figuré.

### Règle 7 : Règle d'inclusion sémantique

|| La lexie  $L_1$  dont la définition est incluse dans celle de  $L_2$  doit précéder  $L_2$ .

En d'autres termes, si  $L_2$  est définie par référence à  $L_1$ , c'est-à-dire si  $\langle L_2 \rangle \supset \langle L_1 \rangle$ , alors  $L_1$  précède  $L_2$  dans le vocable. Par exemple, la définition de DOSI.a 'partie du corps...' [=  $L_1$ ] est incluse dans la définition de DOSII.1 'partie du siège destinée à ce qu'on y appuie le dosI.a'; par conséquent, DOSI.a 'partie du corps...' précède DOSII.1 'partie du siège ...'.

Nous accordons ainsi la priorité absolue à l'inclusion sémantique; la lexie incluse précède la lexie incluante même si la première est moins disponible, ou chronologiquement plus récente, ou grammaticalement plus contrainte. Ainsi, dans le DEC, malgré la tradition lexicographique, CHANGER intransitif (*André a beaucoup changé*) précède CHANGER transitif (*Le bonheur a beaucoup changé André*), puisque CHANGER1 intransitif est défini comme 'devenir différent' et CHANGER2 transitif, comme 'causer devenir différent' = 'causer changer1' (la dernière substitution – 'devenir différent'  $\Rightarrow$  'changer1' – étant obligatoire d'après le Principe de bloc maximal).

Posons L comme lexie de base d'un vocable – lexie qui est mise en première position et par rapport à laquelle toutes les autres lexies du même vocable sont définies.

### Règle 8 : Règle de proximité sémantique

|| La lexie  $L_1$  qui est sémantiquement plus proche de L que la lexie  $L_2$  doit précéder  $L_2$ , de façon à obtenir l'ordre  $L + L_1 + L_2$ .

Comparons, par exemple, BRÛLER dans *Le bois sec* (*Le papier*) brûlait et BRÛLER dans *La ville* (*Le bateau*) brûlait; le premier – 'se consumer en produisant du feu' – est plus proche du sens de base de BRÛLER (= de celui exprimé dans *Le feu brûle*) que le deuxième – 'être en train d'être détruit par le feu'. En effet, le premier sens est presque identique au sens de base : en disant *Le bois sec brûlait*, on communique l'existence du feu; par contre, le deuxième véhicule une information supplémentaire importante : celle d'un sinistre. Par conséquent, le lexème BRÛLER dans *Le bois sec brûlait* correspond à  $L_1$  de la règle 8 et précèdera, dans le vocable BRÛLER, le lexème BRÛLER dans *La ville brûlait*, qui, lui, est  $L_2$ .

La règle suivante, qui vise uniquement les lexies à sens figuré, est fondée sur une distinction universellement reconnue entre *métonymie* et *métaphore*.

Soit deux sens,  $\langle \sigma_1 \rangle$  et  $\langle \sigma_2 \rangle$  tels que  $\langle \sigma_1 \rangle$  est inclus dans  $\langle \sigma_2 \rangle$  :  $\langle \sigma_1 \rangle \subset \langle \sigma_2 \rangle$ .

- |  |
|--|
| Entre $\langle \sigma_1 \rangle$ et $\langle \sigma_2 \rangle$ , il y a <i>relation métonymique</i> si leurs référents se trouvent dans une CONTIGUÏTÉ SPATIALE quelconque.<br>Par contre, entre $\langle \sigma_1 \rangle$ et $\langle \sigma_2 \rangle$ , il y a <i>relation métaphorique</i> si leurs référents se trouvent dans une CONTIGUÏTÉ MENTALE quelconque. |
|--|

La contiguïté spatiale se manifeste en tant que localisation (*Toute l'université riait* = 'TOUS CEUX LOCALISÉS à l'université riaient'), provenance (*Verse-moi un bordeaux* != '... VIN PROVENANT DE Bordeaux'), destination (*un doigt de gant*), etc. La contiguïté mentale, elle, se manifeste surtout en tant que ressemblance – d'après la fonction, la couleur, la forme, etc. (*l'université de la vie, une robe bordeaux, un doigt de contact* <de came, de transfert> – dans une machine), etc.

D'un point de vue formel, la métonymie est l'inclusion d'un sens introduit par une composante sémantique comme 'localisé à...' (*La France y était opposée*), 'provenant de...' (*beaujolais* – vin du Beaujolais), 'destiné à...' (*jambe de pantalon*), 'contenu dans...' (*tasse de farine*), 'fabriqué de...' (*un verre* – fabriqué de verre), '... qui a...' (*un cœur généreux*), etc.

La métaphore, elle, est l'inclusion d'un sens introduit par une composante sémantique 'rappeler...' = 'comme si...' (*tête d'ail; L'Europe sommeillait* = 'Tout en Europe avait l'air de sommeiller').

Les sens métonymiques sont donc sémantiquement plus proches du sens primaire que les sens métaphoriques; ceci nous permet d'énoncer la Règle 9.

### Règle 9 : Règle de priorité métonymique

- |   |
|---|
| La lexie $L_1$ dont le sens est métonymique par rapport à L doit précéder la lexie $L_2$ dont le sens est métaphorique par rapport à la même lexie L. |
|---|

Il s'ensuit que dans un vocable, toutes les lexies à sens figuré suivent toutes les lexies à sens littéral; parmi les lexies figurées, les lexies métonymiques précèdent les lexies métaphoriques. Par exemple, dans le vocable DOIGT, la lexie DOIGT 'partie terminale de la main' précède, en vertu de la Règle d'inclusion sémantique (Règle 7), toutes les autres lexies – puisque celles-ci, qui sont figurées, renvoient toutes à ce DOIGT. Parmi les lexies à sens figuré, DOIGT 'partie d'un gant destinée à recevoir un doigt' (contiguïté spatiale : métonymie) précède DOIGT 'partie mobile allongée d'un dispositif rappelant la forme ou la fonction d'un doigt' (contiguïté mentale : métaphore).

Il peut arriver que les deux lexies d'un même vocable que le lexicologue compare pour les ordonner soient toutes les deux soit métonymiques, soit métaphoriques. Dans un tel cas, la Règle 9 devient, bien entendu, inutilisable. Mais l'utilisation de la Règle 8 permet de trancher la question. Considérons, par exemple, les deux sens figurés du verbe BRÛLER présents dans *Il brûle de fièvre* ('Il a une très forte fièvre') et *Il brûle de passion pour Marie* ('Il a une très forte passion pour Marie'); les deux sont métaphoriques, de sorte que nous ne pouvons

pas les ordonner par la Règle 9. Cependant, nous pouvons avoir recours à la Règle 8. Le premier BRÛLER est plus proche du sens de base (*Le feu brûle*) : en effet, *brûler de fièvre* suppose une vraie chaleur, bien mesurable, qu'on compare à celle du feu qui brûle; par contre, *brûler de passion* suppose une émotion très intense, non mesurable, qu'on compare à la chaleur corporelle, elle-même comparée à la chaleur du feu. Ce dernier sens est donc plus éloigné du sens de base que le premier. Suivant la Règle 8, le lexème correspondant à ce sens – BRÛLERIV.2b – précédera dans le vocable BRÛLER le lexème exprimant le deuxième sens, c'est-à-dire BRÛLERV.2.

La règle 10, encore plus spécifique, ne s'applique qu'à des lexies figurées à sens métaphorique. Pour la formuler, nous devons d'abord dégager la notion de composante sous-jacente au transfert métaphorique. Une telle composante est le pont sémantique entre la lexie de base et la lexie métaphorique correspondante. Ainsi, la composante sous-jacente au transfert métaphorique pour BRAS 'parties latérales et symétriques d'un artefact...' (*bras d'une balance*) qui se trouve dans la définition de BRAS 'parties latérales supérieures et symétriques du corps...', allongées, articulées et mobiles...' est 'parties latérales... et symétriques', puisque c'est cette composante qui forme le pont sémantique entre BRAS 'parties du corps...' et BRAS 'parties d'un artefact...'. Pour BRAS 'parties allongées... d'un cours d'eau...', la composante sous-jacente au transfert métaphorique est bien entendu 'allongée'.

Soient deux lexies  $L_1$  et  $L_2$  dont le sens est métaphorique par rapport au sens de la lexie  $L$ , toutes les trois relevant du même vocable.

### Règle 10 : Règle de hiérarchie des composantes sémantiques

|| La lexie  $L_1$  dont la composante sous-jacente au transfert métaphorique précède, dans la définition de  $L$ , la composante sous-jacente au transfert métaphorique de la lexie  $L_2$  doit précéder  $L_2$ .

Dans l'exemple ci-dessus avec le vocable BRAS, la lexie BRAS 'parties d'un artefact...' précède donc la lexie BRAS 'parties... d'un cours d'eau...' puisque, dans la définition de BRAS 'parties du corps...', la composante de localisation, c'est-à-dire 'latéral', précède la composante de forme, c'est-à-dire 'allongé'.

Le résultat de l'application de cette règle dépend de la formulation de la définition de la lexie de base  $L_1$ , c'est-à-dire de l'ordre d'apparition des composantes définitoires. Or cet ordre n'est pas arbitraire : il est déterminé par la Règle d'ordonnancement par pertinence sémantique, voir Ch. III, 3.3.5.2, p. 106.

Pour revenir à BRAS, la composante '[parties] latérales... et symétriques' précède, dans la définition de cette lexie, la composante 'allongées, ...' parce que pour une partie d'un tout la caractérisation de sa localisation et de sa symétrie par rapport au tout est sémantiquement plus pertinente que la caractérisation de sa forme.

**NB :** Soulignons que tout ce qui vient d'être dit dans la Règle 10 à propos des composantes sémantiques s'applique également aux connotations. Nous ne l'avons pas mentionné dans la règle pour ne pas en alourdir la formulation.

La Règle 11, que nous abordons à présent, est un peu spéciale car elle traite du champ lexical, en faisant appel à la notion 'être typique pour le champ lexical' (le concept de champ lexical est introduit plus loin, en 2.2, p. 176). Une lexie L au sens figuré est dite typique pour un champ lexical si d'autres lexies au sens figuré ayant la même relation par rapport à leur lexie de base que L se retrouvent dans la plupart des vocables de ce champ. Ainsi, la lexie COUDE 'partie d'un artefact... rappelant la forme du coude [partie du corps]' est typique pour le champ lexical 'parties du corps' puisque la plupart des vocables de ce champ incluent des lexies ayant la définition 'partie de... rappelant la forme de... [partie du corps]' : *bouche d'un volcan, bras d'un fleuve, dents d'une roue, jambes d'un compas, langues de feu, nez d'une tuile, lèvres d'une plaie, œil d'une lettre, pied de cordonnier, ventre d'une cruche, ...* La propriété 'être typique pour le champ lexical' est quantitativement variable : un transfert métaphorique peut être plus ou moins typique pour le champ lexical donné selon la proportion des lexies dans lesquelles il apparaît.

### Règle 11 : Règle de caractère typique de la lexie

|| La lexie  $L_1$ , qui est plus typique pour le champ lexical en question que la lexie  $L_2$  précède  $L_2$ .

La lexie DOIGT 'petite quantité de boisson alcoolisée mesurée en terme d'épaisseur d'un doigt' n'est pas du tout typique pour le champ lexical des parties du corps : le vocable DOIGT est le seul à posséder une lexie avec le sens 'quantité... mesurée en terme de... [cette partie du corps]'; par conséquent, cette lexie est placée en dernier dans le vocable DOIGT.

#### 1.4 Distances sémantiques entre les lexies d'un même vocable

D'un point de vue purement logique, les articles de dictionnaire ordonnés à l'intérieur d'un superarticle (= vocable) peuvent tout simplement être numérotés de façon successive – 1, 2, ..., n; certains dictionnaires procèdent ainsi (par exemple, *Le Lexis*. Nous pensons cependant que les locuteurs perçoivent des regroupements HIÉRARCHIQUES de lexies au sein d'un vocable : certaines lexies sont senties comme étant plus liées (ou moins liées) entre elles que d'autres. Ainsi, les lexies TÊTE 'partie supérieure du corps humain' et TÊTE 'partie de la tête où poussent les cheveux et/ou les cheveux' '*se laver* (*se gratter*) *la tête*; *une tête blonde*' sont sans doute plus proches l'une de l'autre que chacune d'elles ne l'est de TÊTE dans *la tête de la colonne* ou dans *la tête d'un mouvement politique*. C'est au fond cette intuition que nous voulons refléter dans l'organisation de nos



vocables. Par conséquent, la numérotation (des articles de dictionnaire) que nous acceptons doit rendre compte du caractère hiérarchique d'un superarticle : un vocable est un regroupement de regroupements, dont chacun à son tour peut être un regroupement de lexies, et ainsi de suite. Autrement dit, la numérotation lexicale du DEC indique les DISTANCES entre les lexies, en donnant à un vocable une structure hiérarchique arborescente. Cette façon de faire est aussi courante dans les dictionnaires existants (par exemple, le *Petit Robert*).

La distance entre deux lexies quelconques correspond surtout à la proximité de leurs sémantismes; c'est pourquoi nous parlons de distances sémantiques. Signalons, cependant, que la proximité sémantique est en règle générale accompagnée par une similitude de comportement syntaxique et de cooccurrence lexicale.

Dans le Chapitre III (3.3.1.2, après la Règle 4, p. 87), nous avons déjà traité très brièvement de la représentation des distances sémantiques par des numéros lexicaux; ici nous reprenons la question de façon plus systématique.

Nous distinguons trois niveaux de distances sémantiques entre les lexies :

- les grandes distances, notées par des chiffres romains;
- les distances moyennes, notées par des chiffres arabes;
- et les petites distances, notées par des lettres minuscules.

Un superarticle, c'est-à-dire un vocable, apparaît donc, par exemple, sous la forme présentée dans la Fig. IV-1(A). Cette présentation correspond à la structure arborescente de la Fig. IV-1(B).

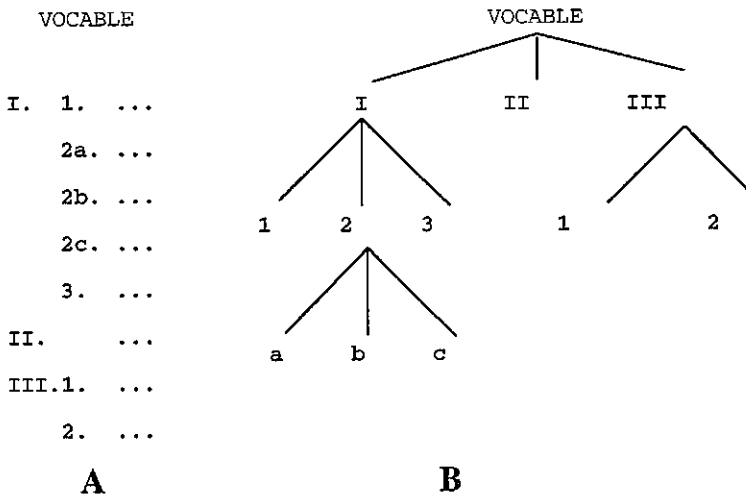


Figure IV-1 Distances sémantiques entre lexies d'un vocable

La distance sémantique entre les lexies  $L_1$  et  $L_2$  est mesurée, comme nous l'avons déjà dit, en fonction de deux facteurs : d'une part, l'IMPORTANCE de leur intersection sémantique et, d'autre part, la RÉGULARITÉ de la distinction sémantique entre elles.

La partie commune des définitions des lexies  $L_1$  et  $L_2$  est dite *importante* si elle est constituée par près de la moitié du matériel sémantique d'au moins une des deux définitions. Ainsi, les définitions de VITESSE 'grandeur<sup>III</sup> caractérisant le déplacement de X par la distance parcourue...' (*la vitesse de l'avion*) et de VITESSE 'grandeur<sup>III</sup> caractérisant la rotation de X par le nombre de tours...' (*la vitesse de la turbine*) ont une partie commune importante : 'grandeur<sup>III</sup> caractérisant le comportement continu de X par le nombre de...'.<sup>1</sup>

La distinction entre les définitions des lexies  $L_1$  et  $L_2$  est dite *régulière* si elle est présente dans beaucoup d'autres paires de lexies. Ainsi, la distinction entre les définitions de CHANGER 'devenir différent' et de CHANGER 'causer que X devient différent' – 'causer que...' – est régulière en français puisqu'elle se retrouve dans des centaines de paires de verbes : CASSER (*La branche casse ~ Jean casse la branche*), REMPLIR (*Le vin remplit la bouteille ~ Jean remplit la bouteille de vin*), CUIRE (*La viande cuit ~ Jean cuit la viande*), ...

La distance sémantique entre  $L_1$  et  $L_2$  est GRANDE si la partie commune de leurs définitions n'est pas assez importante et si la distinction sémantique entre elles n'est pas régulière. Cette distance est marquée par des chiffres romains. Dans ce cas, les lexies  $L_1$  et  $L_2$  sont sémantiquement apparentées, mais assez éloignées.

La distance sémantique entre  $L_1$  et  $L_2$  est RELATIVEMENT PETITE si la partie commune de leurs définitions est assez importante et si la distinction sémantique entre elles est assez régulière. Cette distance est marquée par des chiffres arabes. Dans ce cas, les lexies  $L_1$  et  $L_2$  sont sémantiquement proches.

La distance sémantique entre  $L_1$  et  $L_2$  est PETITE si la partie commune de leurs définitions est importante et si la distinction sémantique entre elles est régulière. Cette distance est marquée par des lettres minuscules. Dans ce cas, les lexies  $L_1$  et  $L_2$  sont sémantiquement très proches.

Pour conclure la discussion sur les distances sémantiques entre les lexies d'un même vocable, signalons que la distance sémantique entre les sens primaires (littéraux) faisant partie du vocable et tous les sens figurés, c'est-à-dire métonymiques et métaphoriques, appartenant au même vocable est postulée comme étant GRANDE. Par exemple, pour NEZ, les lexies à sens primaires sont regroupées sous NEZI (NEZI.1a = 'nez humain', NEZI.1b = 'visage' [*ne pas montrer son nez*], NEZI.1c = 'nez d'animaux', NEZI.2 = 'odorat' [*avoir du nez*]) et les lexies à sens métaphoriques, sous NEZII (NEZII.1 = 'partie antérieure...' [*nez d'un avion*], NEZII.2 = 'cap élevé...' [*nez de Jobourg*]). De façon similaire, dans le vocable DOS, nous avons DOSI.a et DOSI.b ('dos de personne, dos d'animal'), opposés à quatre lexèmes figurés : deux métonymiques (DOSII.1 de fauteuil et DOSII.2 de vêtement) et deux métaphoriques (DOSII.3a de cuillère et DOSII.3b de colline).

Pour illustrer ce que nous venons de présenter, prenons le vocable nominal MER, qui comprend les quatre lexèmes suivants (DEC-3)<sup>3</sup> :

## MER

**MER1.1a**, pas de pl, normalement avec l'ART<sub>déf</sub>

*Mer* = 'Immense [étendue d'eau] salée qui couvre une grande partie de la Terre – un des éléments de base de la Terre).

[*La mer couvrait la plupart de la surface de la planète; Encore au 17<sup>e</sup> siècle, la mer pulullait de pirates; Ce plancton se trouve partout dans la mer.*]

**MER1.1b**, pas de sg, normalement avec l'ART<sub>déf</sub>

*Mers* = 'Ensemble des parties de la mer1.1a).

[*Les mers couvraient la plupart de la surface de la planète; Encore au 17<sup>e</sup> siècle, les mers pulullaient de pirates; Ce plancton se trouve dans les mers tropicales.*]

**MER1.2**. *Mer* = 'Immense partie de la mer1.1a délimitée par une ou des terre(s) mais pas par un ou des continent(s) entier(s) ni par une échancre de la côte).

[*la mer Noire, la mer de Chine, la mer Baltique, ...; Nous naviguons sur une mer inconnue.*]

**MERII**. *Mer de X* = 'Immense étendue de X qui n'est pas de l'eau [comme si c'était une mer1.2)].

[*une mer de voitures; des mers végétales*]

Justifions d'abord l'ordonnancement proposé, puis les distances sémantiques.

### Ordonnancement

Le lexème MER1.1a ne peut être qu'en première position puisqu'il ne fait appel à aucun des autres lexèmes du vocable, alors que ceux-ci y renvoient tous. Comment ordonner ces derniers entre eux ? Le lexème MERII étant une acception figurée, plus précisément métaphorique, il doit être placé en dernière position – après toutes les acceptions littérales. Le lexème MER1.1b semble sémantiquement plus proche de MER1.1a que ne l'est MER1.2 : la différence sémantique est 'ensemble des parties', qui produit une quasi-synonymie (en effet, 'X' ≈ 'ensemble des parties de X'). Ceci est attesté par les exemples, qui, comme on le voit, sont presque interchangeables entre ces deux acceptions – avec quand même de

3 La version du vocable MER présentée ici est légèrement différente de celle du DEC : nous avons supprimé le lexème MER1.2a, qui a été jugé non pertinent. Comme l'expression 'MER LUNAIRE' constitue un phrasème (signifiant 'vaste région en dépression sur la surface de la Lune [comme si c'était une mer1.3]', nous ne l'avons pas incluse dans le vocable MER.

légères nuances. Par conséquent, **MERI.1b** suivra **MERI.1a** immédiatement, en précédant **MERI.2**.

**NB :** Cette différence est assez régulière en français; ainsi, on la retrouve dans *la forêt ~ les forêts* (*Les trois quarts du territoire sont couverts par la forêt tropicale / par les forêts tropicales*), *la montagne ~ les montagnes*, *le désert ~ les déserts, ...*, *la carotte ~ les carottes* (*La carotte est un légume riche en minéraux / Les carottes sont des légumes riches en minéraux*), *le haricot ~ les haricots*, *le navet ~ les navets, ...*

Les explications fournies permettent donc de justifier l'ordonnement proposé.

### **Distances sémantiques**

Les trois premiers lexèmes de **MER** désignent tous des <sup>1</sup>étendues d'eau<sup>1</sup>; nous les avons donc regroupés sous **MERI**; le lexème **MER** dans le sens 'immense étendue de X...)' devient **MERII**. Comme il s'agit d'une différence « sens littéraux ~ sens métaphorique », elle doit être marquée par des chiffres romains.

Parmi les lexèmes de **MERI**, deux lexèmes – **MERI.1a** et **MERI.1b** – nous paraissent plus rapprochés l'un de l'autre (voir ci-dessus), ce qui est marqué par des lettres.

Enfin, la distance entre **MERI.1a, b** et **MERI.2** est moyenne, de sorte que nous l'indiquons par des chiffres arabes.

Nous avons donc fourni tous les éléments nécessaires à la structuration d'un vocable et nous pouvons ainsi mettre un terme à la section **1**; mais avant de le faire, nous nous devons de signaler que le système de règles proposées renferme un défaut inhérent. En effet, nous avons parlé de l'ordonnement **LINÉAIRE** des lexies, alors qu'on ne peut nullement supposer que l'emmagasinement des lexies dans la langue, c'est-à-dire dans le cerveau des locuteurs, soit linéaire. Cela signifie que nos règles tentent de projeter sur une surface plane ce qui est essentiellement tridimensionnel, et on sait très bien que c'est impossible sans déformations et entorses sérieuses en ce qui a trait tant à l'ordonnement qu'aux distances sémantiques entre les lexies. Le cas de l'ordonnement des lexies dans un dictionnaire est en fait semblable à celui des cartes géographiques, qui représentent sur le papier, donc sur une surface plane, la surface sphérique de la Terre. Les déformations sont inévitables, et le seul choix qu'a le cartographe est d'explicitier la déformation survenue ainsi que de décider du type de déformation préférée (respecter plutôt les distances – la projection de Mercator – ou les dimensions).

La tâche du lexicologue est toutefois encore plus complexe à cause de l'absence de considérations **TÉLÉOLOGIQUES** : les ordonnements et les distanciations différentes des lexies à l'intérieur d'un vocable n'entraînent aucune conséquence dans l'utilisation des articles de dictionnaire correspondants, ce qui

supprime les critères pratiques. Mais même si l'ordre et les distances lexicales n'ont pas d'impact sur l'usage du DEC pour la synthèse des textes, il n'en reste pas moins qu'ils ont une valeur pédagogique et psychologique certaine pour l'utilisateur humain. Le cerveau est habitué à percevoir et à assimiler par regroupements; nous pensons que ce trait typiquement humain représente une justification suffisante de notre façon de faire.

### 1.5 Tableau synoptique du vocable

Les vocables sont souvent très riches, c'est-à-dire qu'ils comportent de nombreux lexèmes. Par exemple, le vocable FLAMBER (DEC-2) est constitué de 18 lexèmes différents; si ce nombre est quelque peu plus grand que la moyenne, il ne s'agit cependant pas d'une exception. Le lexicologue qui élabore le vocable et encore bien plus l'utilisateur du dictionnaire risquent d'être noyés par la diversité d'acceptions, de cooccurrents, etc. et ainsi de perdre de vue les ressemblances et les différences entre les lexies formant le vocable. Pour pallier cet inconvénient et rendre facilement accessible l'information fort pertinente sur la structure du vocable, nous proposons de mettre au tout début de chaque vocable un *tableau synoptique*.

Le tableau synoptique d'un vocable joue, pour ainsi dire, le rôle d'une table des matières de ce vocable. Il contient une liste complète de toutes ses lexies. Chaque lexie y est représentée par son numéro lexical et une définition abrégée, mais intelligible et distincte de toutes les autres définitions présentées dans le tableau synoptique. Un exemple bref et bien choisi facilite encore l'identification de l'acception voulue.

À titre d'exemple, nous présentons ci-dessous le tableau synoptique du vocable FLAMBER (repris au DEC-2, avec de légères adaptations).

#### FLAMBER, verbe

- I. 1a. X brûle<sup>I.1</sup> intensément... [*Un feu flambe*]  
 1b. X brûle<sup>I.2</sup> d'un feu qui flambe<sup>I.1a</sup>... [*Une allumette flambe*]  
 2a. Alcool... ajouté à une boisson X ou à un mets X flambe<sup>I.1b</sup> [*Le punch flambe*]  
 2b. X cause que Y flambe<sup>I.2a</sup>... [*Pierre flambe les crêpes*]  
 3. X soumet un objet Y... à l'action d'un feu flambant<sup>I.1a</sup>... [*Pierre flambe le cochon*]
- II. 1. Astre X brille<sup>I</sup> intensément... [*Le soleil flambe*]  
 2. X semble flamber<sup>II.1</sup> à cause d'une source de lumière Y... [*Le ciel flambe d'éclairs*]  
 3a. Astre X brille<sup>3</sup> sur Y de telle façon que X semble flamber<sup>I.1a</sup>... [*Le soleil flambe sur les collines*]  
 3b. X brille<sup>4</sup> de Y dans Z de telle façon que X semble flamber<sup>II.1</sup>... [*Le diamant flambe d'une lueur vive dans les rayons du soleil*]

4. X de Y brille<sup>5</sup> de telle façon que X semble flamber<sup>II.1</sup>... [*Une couleur flambe*]
- II. I. X brûle<sup>IV.2a</sup> ... [*La gorge de Pierre flambe*]
- IV. 1. Impulsion humaine X ou émotion X a lieu... [*La joie flambait dans son cœur*]  
 2. X est dans un état d'excitation en éprouvant une émotion Y... [*Pierre flambe de joie*]  
 3a. Émotion... intense se manifeste par l'éclat anormal des yeux<sup>I.1a</sup> Y... [*La colère flambe dans ses yeux*]  
 3b. (Partie du) visage<sup>I.a</sup> X a une teinte anormalement rouge... ou les yeux<sup>I.1a</sup> X ont un éclat anormal... [*Ses yeux flambent de colère*]
- V. Conflit militaire X a lieu... [*La guerre flambe*]
- VI. 1. X dépense très vite Y de façon immodérée... [*Pierre flambe sa fortune*]  
 2. X joue intensément à des jeux de hasard... [*Pierre flambe aux cartes*]

Ce tableau permet, d'un seul coup d'œil, de voir la structuration sémantique du vocable étudié. En effet, on constate l'organisation sémantique suivante :

FLAMBERI (de 1a à 3)	concerne	le FEU véritable
FLAMBERII (de 1 à 4)	concerne	la LUMIÈRE
FLAMBERIII	concerne	la SENSATION
FLAMBERIV (de 1 à 3b)	concerne	les ÉMOTIONS et les MANIFESTATIONS D'ÉMOTIONS
FLAMBERV	concerne	le CONFLIT MILITAIRE
FLAMBERVI (1 et 2)	concerne	les DÉPENSES (excessives)

Une telle présentation laisse apparaître les liens sémantiques de façon très claire et rend l'unicité du vocable plus visible.

## 2 Structure d'une famille d'articles de dictionnaire : champ lexical

Comme nous l'avons dit au début de ce chapitre, la macrostructure du DEC est formée à deux niveaux : d'abord, les lexies sont regroupées au sein des vocables (section 1); ensuite, les vocables sont regroupés en champs lexicaux; c'est ce dont il sera question dans cette section. Soulignons que la notion même de regroupement n'est pas tout à fait la même pour les vocables et pour les champs lexicaux; nous allons fournir les précisions nécessaires plus loin.

Le concept de champ lexical est fondé sur ceux de champ sémantique et de vocable. Le vocable ayant été défini plus haut, nous commencerons par le champ sémantique (2.1), pour parler ensuite du champ lexical (2.2) et aborder finalement le problème de son utilisation dans l'élaboration et la présentation du dictionnaire (2.3-4).

## 2.1 Champ sémantique

Pour être en mesure de répondre, de façon satisfaisante, aux exigences posées de rigueur, de formalité et de caractère explicite de la description lexicographique, le lexicologue se voit obligé de fragmenter la masse lexicale d'une langue et de constituer ainsi des regroupements de lexies qui présentent des caractéristiques sémantiques communes plus ou moins semblables ou même équivalentes. Cette façon de faire divise le lexique en ensembles déterminés et limités quant au nombre de leurs éléments. De tels ensembles, suffisamment maniables, qu'on appelle *champs sémantiques*, forment l'objet primaire d'étude du lexicologue. Au chapitre II, section 1.3, p. 40, nous avons déjà – sinon *de jure*, au moins *de facto* – érigé le traitement par champ sémantique en un des principes de base de rédaction du DEC; il s'agit ici de formuler les concepts indispensables pour énoncer ce principe explicitement.

L'acquisition du lexique par le locuteur est intimement liée au regroupement de lexies en familles qui manifestent une parenté sémantique plus ou moins évidente. La langue elle-même détermine de tels regroupements en fournissant des lexies qui désignent les composantes sémantiques sous-jacentes à des familles particulières. Ainsi, un francophone regroupera instinctivement les lexies TABLE, LIT, CHAISE, ARMOIRE, VAISSELIER, etc. en famille de 'meubles', pusique sa langue lui donne la lexie MEUBLE, qui exprime cette composante sémantique « familiale ». De la même façon, on regroupera les lexies DESSINER, PEINDRE, SCULPTER, TAILLER, GRAVER, etc. sous l'étiquette 'représenter', et ROUGE, BLEU, VERT-GRIS, etc. – sous 'couleur'. Ces étiquettes, qui correspondent à un générique commun, peuvent tout naturellement être désignées comme *identificatrices de champ* (sémantique). Nous pouvons donc maintenant aisément définir le champ sémantique.

### Définition IV.6 : Champ sémantique

|| Nous appelons *champ sémantique* l'ensemble des lexies qui ont une même composante sémantique identificatrice de champ.

Éclairons le concept introduit en nous appuyant sur l'exemple de champ sémantique 'phénomènes atmosphériques', formé des lexies PLUIE, NEIGE, GRÊLE, BROUILLARD, TEMPÊTE, ORAGE, ... (DEC-2). Pour ce faire, nous devons caractériser les trois aspects suivants de la composante identificatrice de champ.

Premièrement, le choix de la désignation de la composante sémantique identificatrice de champ n'est ni arbitraire ni sans conséquence. Elle est beaucoup plus qu'une simple étiquette qui identifie un champ sémantique; elle doit servir en tant que composante sémantique dans les définitions de tous les membres du champ et elle est donc par conséquent assujettie à tous les principes, règles et critères de construction de définitions. Ainsi, pour le champ 'phénomènes

atmosphériques), il faut retenir l'adjectif *atmosphérique* plutôt que *météorologique*, utilisé dans des dictionnaires courants pour définir PLUIE, etc. En effet, météorologique est lui-même sémantiquement dérivé de 'atmosphère' :

- météorologique* : 'relatif à la météorologie'
- météorologie* : 'science des phénomènes atmosphériques'
- atmosphérique* : 'relatif à l'atmosphère'
- atmosphère* : 'couche d'air qui entoure la Terre'.

Cette petite série de définitions montre à l'évidence que 'atmosphère' (et non 'météorologie') est le sens sous-jacent au champ en cause. La pluie et les tempêtes existaient bien avant l'apparition de la météorologie !

Deuxièmement, même si la composante identificatrice de champ fait partie de toutes les définitions des lexies du champ, elle n'y apparaît pas toujours de façon explicite. Ainsi, le lexème PLUIE est défini comme 'chute d'eau en tant que phénomène atmosphérique ou... cette eau', où la composante identificatrice de champ est explicite. Par contre, PLEUVOIR est défini comme 'la pluie a lieu'; dans cette dernière définition, la composante identificatrice de champ n'est présente qu'implicitement, par le biais de la définition de PLUIE.

Troisièmement, la composante sémantique identificatrice de champ n'est pas nécessairement la composante centrale de la définition où elle apparaît. PLUIE, comme on vient de le voir, n'est pas 'phénomène atmosphérique qui...', mais ('... en tant que phénomène atmosphérique'); la même formulation est valable pour les autres membres du champ. Cela s'explique par plusieurs raisons; nous n'en mentionnons ici que deux :

- Le lexème PLUIE (comme d'ailleurs les autres précipitations) désigne aussi une substance – l'eau qui tombe (cf. *La pluie qui durait depuis trois jours ruisselait le long des murs* : c'est la chute d'eau qui durait et c'est l'eau qui ruisselait). La composante 'phénomène atmosphérique' ne peut donc être la composante centrale.
- Certains lexèmes de ce champ sémantique ne sont pas comptables : \**trois neiges* [= *trois chutes de neiges*], \**quatre brouillards* [= *avoir du brouillard quatre fois*], ... Pourtant 'phénomène' l'est : *quatre phénomènes atmosphériques*; si nous définissions NEIGE comme 'phénomène qui...', il serait difficile d'expliquer la non-comptabilité de ce lexème.

La composante sémantique identificatrice de champ peut donc se trouver dans des positions différentes au sein d'une définition.

Le rôle de la composante identificatrice de champ est double. D'une part, sa présence permet de formuler des règles générales concernant le champ sémantique étudié dans sa totalité. Par exemple, on peut extraire les renseignements lexicographiques communs à la majorité des membres du champ et les réunir dans l'article de dictionnaire de la lexie  $L_{\text{idént}}$  qui correspond à la composante identificatrice de champ. Pour ce faire, on crée dans l'article de dictionnaire de



$L_{ident}$  une division spéciale – appelée *section publique* – destinée à accueillir toutes les informations communes aux membres du champ. De cette façon, on évite toute répétition fastidieuse et inutile tout en obtenant des formulations généralisantes. Ce procédé est connu sous le nom d'*héritage lexical* : la lexie hérite une partie de ses informations lexicographiques de la section publique de l'article de dictionnaire de la lexie  $L_{ident}$ . Dans le champ de phénomènes atmosphériques, les verbes comme COMMENCER, DURER ou FINIR ne doivent pas figurer séparément sous PLUIE, GRÊLE, NEIGE, etc., mais sont hérités de la lexie identificatrice de champ, qui est en l'occurrence PHÉNOMÈNE ATMOSPHÉRIQUE. (Pour la question d'héritage lexical dans le cadre du DEC, voir Mel'čuk & Wanner 1994.)

D'autre part, la composante sémantique identificatrice de champ aide à dégager les schémas généralisés de définition pour les membres de ce champ, ce qui contribue à la standardisation et à l'homogénéité de la description lexicographique.

Comme nous l'avons dit à maintes reprises, le lexicologue procède par champ sémantique – en décrivant en parallèle les lexies d'un même champ, en les comparant et en les opposant sans cesse. Le traitement simultané des lexies d'un même champ sémantique met en évidence leurs ressemblances ainsi que leurs dissemblances, donnant, de cette façon, une plus grande profondeur de vision et assurant, en même temps, une meilleure cohérence et une parfaite systématisme. Par exemple, pour décrire le lexème CŒUR, le lexicologue du DEC doit considérer les autres organes du corps et donc traiter CŒUR au sein du champ sémantique 'organes du corps' (soit POUMON, ESTOMAC, FOIE, etc.); de la même façon, le lexème LIT doit être étudié en même temps que les autres lexies du champ sémantique 'meuble', BISTOURI – avec le champ sémantique 'chirurgie', et [un] FRANÇAIS – dans le champ sémantique 'ethnies'. Un tel traitement garantit une cohérence descriptive pour les lexies sémantiquement apparentées.

La tâche du lexicologue ne s'arrête cependant pas là. Toute lexie appartient aussi à un vocable et nécessite une cohérence pas moindre avec les autres lexies de ce vocable (problème déjà traité dans la section précédente). Le lexicologue est ainsi forcé de « jongler » avec deux types de cohérence pour chaque lexie : la cohérence horizontale – avec les autres lexies de son champ sémantique, et la cohérence verticale – avec les autres lexies de son vocable. Cette situation nous amène à une entité plus complexe que le champ sémantique, soit le champ lexical.

## 2.2 *Champ lexical*

### Définition IV.7 : Champ lexical

|| Nous appelons *champ lexical* d'un champ sémantique l'ensemble des vocables dont les lexies de base appartiennent à ce champ sémantique.

Prenons le champ sémantique 'parties du corps' : il est formé par les lexèmes BRASI.1a, JAMBEL.1a, VENTREL.1a, TÊTEL.1a, POITRINE1a, etc. ; la lexie identificatrice de ce champ est, bien entendu, l'expression PARTIE DU CORPS. Chacun des lexèmes indiqués entre, à son tour, dans un vocable :

- le vocable BRAS comporte 14 lexèmes, dont BRASI.1a [de Jean], BRASI.1b [du singe], BRASI.1c [du poulpe], BRASI.2a [anatomique : entre le coude et l'épaule], BRASI.2b [du cheval], BRASII.1 [d'un brancard], BRASII.2 [d'une ancre], ..., BRASII.6 [d'une manivelle], BRASIII.1 [de l'industrie], ..., BRASIV [de justice];
- le vocable JAMBE en a 7, dont JAMBEL.1a [de Jeanne], JAMBEL.1b [de la girafe], ..., JAMBELI.1 [du pantalon], JAMBELI.2 [du compas], JAMBELI.3 [de force/de train d'atterrissage]; et ainsi de suite.

La réunion de tous ces vocables forme le champ lexical des parties du corps.

C'est le champ lexical qui est en réalité le nœud central du travail lexicographique. Si l'objet primaire du lexicologue – le point de départ pour une étude particulière – est un champ sémantique précis, alors l'objet final de cette étude, c'est-à-dire les données complètes à couvrir, est le champ lexical correspondant. Ce n'est qu'après avoir traité toutes les lexies de tous les vocables impliqués dans le champ sémantique choisi, que le lexicologue peut être plus ou moins certain de la qualité des résultats obtenus.

Schématiquement, on peut donc dire que le travail du lexicologue procède par les trois étapes successives suivantes :

UN CHAMP SÉMANTIQUE DONNÉ  
+  
LES VOCABLES CORRESPONDANTS  
+  
LE CHAMP LEXICAL QUI EN DÉCOULE

Cependant, il ne faudrait pas se laisser induire en erreur par la simplicité de ce schéma. La réalité linguistique est de loin plus riche, ce qui entraîne diverses complications. Nous devons admettre que le concept fondamental de champ sémantique n'est pas si rigoureux que l'on peut penser.

#### 1° Frontières mal démarquées

Le contenu même d'un champ sémantique n'est pas très bien défini : par exemple, les lexèmes DOIGT et JOUE appartiennent-ils au champ sémanti-

que 'parties du corps' ? En effet, DOIGT est une partie de la MAIN, qui est une partie du BRAS, et ce n'est que ce dernier qu'on peut définir sans difficulté comme partie du corps. Bien plus, les YEUX ou les OREILLES sont des organes plutôt que des 'parties du corps'; et pourtant, le lexicologue doit les prendre en charge lorsqu'il travaille sur ce champ.

### 2° Appartenance multiple de lexèmes

Un lexème peut appartenir à des champs sémantiques différents. Ainsi, le lexème BISTOURI appartient au champ sémantique 'outils coupants' (COUTEAU, SCIE, HACHE, ... – jusqu'aux armes blanches comme SABRE, ÉPÉE, etc.) et au champ sémantique 'chirurgie' (SCALPEL, OPÉRATION, AGRAFE, SERINGUE, ...). De la même façon, AVION relève des champs 'transport', 'guerre', 'sport', etc.

### 3° Chevauchement des champs sémantiques

Même si on considère des champs sémantiques relativement bien définis, on trouve beaucoup de cas d'inclusion et d'intersection de champs. Le champ 'meubles' inclut le champ 'sièges' (CHAISE, BANC, TABOURET, ...); le champ 'outils' se recoupe avec le champ 'armes blanches' (l'intersection contient, par exemple, COUTEAU et HACHE); le champ 'oiseaux' a une forte intersection avec 'élevage', 'chasse', 'sport' et 'gastronomie'.

Il est inutile de multiplier les exemples : on voit déjà qu'un concept lexicographique aussi fondamental dans notre optique que celui de champ sémantique reste relativement vague. Ce vague ne peut qu'être accentué au niveau du champ lexical.

Bien qu'un champ lexical soit déterminé par un champ sémantique particulier, ce champ lexical renferme de nombreuses lexies qui appartiennent à d'autres champs sémantiques, qui, à leur tour, déterminent eux aussi d'autres champs lexicaux. Il en résulte qu'un champ lexical, DE PAR SA NATURE MÊME, a des contours flous : il s'enchaîne à beaucoup d'autres champs lexicaux, dont chacun, à son tour, s'enchaîne à beaucoup d'autres, et ainsi de suite – de sorte qu'il est bien possible que les champs lexicaux forment une toile d'araignée englobant le lexique d'une langue dans son entier. C'est bien cela qui rend la tâche du lexicologue aussi fascinante mais aussi très difficile : ne pouvant considérer le lexique dans sa totalité, il doit accepter le meilleur compromis possible, donc œuvrer avec des champs lexicaux fragmentés mais maniables. Cette procédure exige une remise en question permanente des descriptions faites – en fonction des nouvelles données à considérer. L'étymologie latine du terme *travail* – *tripallium* 'trident utilisé pour faire avancer' – convient bien à la description du travail lexicologique !

### 2.3 Principe de traitement par champs sémantiques et lexicaux

Après toutes les discussions précédentes, il semble évident que, pour satisfaire les exigences de cohérence interne et externe des articles de dictionnaire, la seule façon de faire est d'examiner en parallèle et de traiter simultanément des lexies du même champ sémantique ou, encore mieux, du même champ lexical. Il est exclu que le lexicologue du DEC procède de la façon traditionnelle, c'est-à-dire par lettres; les énoncés du type « Nous sommes à la lettre G » ou « Jean rédige la lettre P et Marie, la lettre M » n'ont aucun sens dans le cadre du DEC. Le travail lexicologique ne peut être accompli que par champs lexicaux. Nous aimerions présenter cette constatation comme un autre principe de rédaction :

|| Dans un DEC, les descriptions des lexies se font obligatoirement par champs sémantiques ou lexicaux.

Tous les principes de rédaction du DEC étant présentés dans le Chapitre II, nous avons dû reporter la formulation de ce principe à ce point de notre exposé puisque nous devons auparavant introduire les concepts cruciaux de champs sémantique et lexical; cela ne diminue bien sûr en rien son importance.

### 2.4 Champs sémantiques et lexicaux dans le dictionnaire

Les champs sémantiques et lexicaux sont des regroupements de lexies – tout comme les vocables. Cependant, comme nous l'avons déjà dit, les fonctions de ces deux types de regroupements sont fondamentalement différentes quant à l'élaboration et la présentation des articles de dictionnaire.

Lors de l'ÉLABORATION des articles de dictionnaire, les champs et les vocables jouent le même rôle : ils garantissent le traitement de chaque lexie toujours en opposition avec ses partenaires « horizontaux » (= dans les champs) et ses partenaires « verticaux » (= dans les vocables). Le lexicologue applique ainsi le fameux principe voulant que la langue naturelle n'est qu'un système d'oppositions. Seulement la comparaison par opposition peut faire ressortir les propriétés sémantiques et syntaxiques spécifiques de lexies et en permettre une description suffisamment fine et tranchée. Par exemple, la description de la lexie IMPRIMER (*Cet éditeur imprimera ton roman à 3 000 exemplaires*) exige qu'on l'oppose, par exemple, à DACTYLOGRAPHIER et PHOTOCOPIER, c'est-à-dire qu'on la traite à l'intérieur du champ 'reproduction'. En même temps, il n'est pas moins indispensable d'opposer cet IMPRIMER aux autres IMPRIMER (*imprimer une passion folle dans le cœur de N, imprimer une énergie additionnelle à un faisceau d'électrons, imprimer un visa dans le passeport, ...*), c'est-à-dire de traiter la lexie à l'intérieur du vocable IMPRIMER.

Dans la PRÉSENTATION des articles de dictionnaire terminés, cependant, les champs et les vocables n'apparaissent plus sur le même plan. Les lexies appartenant à un même vocable sont physiquement réunies dans le dictionnaire pour

former un SUPERARTICLE qui correspond à ce vocable; il est inutile d'en parler ici, puisque nous en avons déjà discuté longuement (section 1). Par contre, les lexies appartenant à un même champ sémantique ne peuvent pas être physiquement réunies ! Cette constatation revêt une telle importance qu'il convient de fournir les explications nécessaires.

Du point de vue purement formel, n'importe quel regroupement physique d'articles dans un dictionnaire est pensable; il faut quand même prêter attention aux résultats – du point de vue de la substance. Or un premier obstacle au regroupement de lexies par champ dans un DEC imprimé serait un éclatement des vocables. En effet, les différentes lexies d'un vocable appartiennent le plus souvent à des champs différents. Par exemple, le lexème MERI.1a (*La mer s'étendait à perte de vue*) serait consigné dans le champ 'étendues d'eau', alors que MERII (*mer de voitures*) serait placé dans le champ 'grande quantité'; de cette façon, ces deux lexies se trouveraient à des endroits éloignés. En procédant ainsi, l'usager ne serait plus en mesure d'observer directement la polysémie, y compris la filiation sémantique de lexies apparentées, ce qui semble aller à l'encontre de l'intuition linguistique. Bien plus, toutes les informations syntaxiques et morphologiques communes aux lexies du même vocable ne pourraient plus être extraites et présentées une seule fois en tête du vocable : on devrait les répéter au complet pour chaque lexie. (Non seulement il s'agirait ici d'une redondance encombrante, mais aussi d'une perte importante de généralisations.)

Un deuxième obstacle au regroupement de lexies par champ est le caractère intrinsèquement flou des champs sémantiques et lexicaux. Si l'« effectif » d'un vocable est toujours déterminé, celui d'un champ sémantique ne l'est pas dans la plupart des cas. Pour un champ lexical, ce n'est jamais le cas. Si donc on décidait de présenter un DEC par champs, on se vouerait à des difficultés de classement sans fin : quelle lexie inclure/ne pas inclure dans quel champ ? Étant donné le chevauchement des champs eux-mêmes, le problème de présentation deviendrait insoluble.

Par conséquent, dans la présentation du DEC, nous optons pour l'expression de la proximité « intravocable » par la proximité physique – contiguïté sur la page; les lexies faisant partie d'un même vocable sont donc disposées les unes à la suite des autres. Quant à la proximité « intrachamp », elle n'est pas prise en charge par la disposition typographique.

Pour résumer, nous pouvons énoncer la stratégie à suivre dans la création du DEC :

- ÉLABORER PAR CHAMPS (sémantiques et lexicaux),
- PRÉSENTER PAR VOCABLE (dans l'ordre alphabétique).

**NB :** Ceci ne s'applique qu'à un DEC imprimé sur papier. Pour les DEC informatisés, la situation est différente. Si le principe d'élaboration par champs reste intact, la façon de présenter les articles change de façon radicale. On peut imaginer une présentation fluide, modifiable sur

demande par l'utilisateur, – de sorte que de multiples présentations soient aisément possibles : sorties par vocable ou bien par champ, les limites désirées ayant été spécifiées. Cette question est traitée de façon détaillée au Ch. VI.

Cependant, comme il est dommage de perdre les avantages du regroupement de lexies par champs, nous proposons d'introduire – en tant que partie intégrante du dictionnaire – un *indicateur de champs*. On peut envisager une hiérarchie de champs sémantiques (du style de Hallig & Wartburg 1952), avec, pour chaque champ, la liste de lexies qui en relèvent. Quant aux champs lexicaux, une solution de présentation reste encore à trouver.

## CHAPITRE V

# ILLUSTRATIONS LEXICOGRAPHIQUES

Nous pouvons à présent passer à l'illustration de la théorie lexicologique présentée dans les chapitres précédents. Plus spécifiquement, nous donnerons ci-dessous (en 4) 11 articles de dictionnaire élaborés en suivant le cadre du DEC. Ces articles, qui correspondent aux trois vocables – VENDRE, SE VENDRE et VENDEUR, ont été repris (avec des changements) au DEC-2 et retravaillés en consultant quelques centaines de phrases puisées dans TEXTUM (Ch. II, 3.2, p. 52). De plus, nous avons essayé de prendre en charge toutes les données correspondantes du *Petit Robert*.

Pour bien clarifier la procédure de description lexicologique, nous présentons en détail le cheminement du lexicologue en train de travailler sur les lexies du vocable VENDRE – de façon à permettre au lecteur de suivre pas à pas le déroulement de la réflexion.

Nous procéderons par les étapes suivantes :

- champ lexical (1);
- vocable (2);
- lexème (3).

### 1 Détermination d'un champ lexical

Notre choix des articles illustratifs respecte le *Principe de traitement par champs sémantiques et lexicaux* (Ch. IV, 2.3, p. 178), c'est-à-dire que nous traiterons des lexies du même champ lexical, à savoir, celui de la VENTE. (En fait, nous ne présenterons qu'une infime partie de ce champ.)

Examinons de plus près la démarche du lexicologue au tout début d'une recherche particulière. Supposons qu'il ait décidé, tout à fait arbitrairement, de décrire le vocable VENDRE. (Soulignons que le choix du point de départ relève de la seule volonté du chercheur et n'a aucune conséquence pour le résultat final.) Il constate immédiatement que ce vocable contient plusieurs acceptions, donc – dans la conception du DEC – plusieurs lexèmes. Sa première tâche est de déterminer le lexème de base, celui qui est, d'après son intuition, le plus général, le plus fréquent, ou le plus disponible. Dans le cas de VENDRE, c'est très certainement VENDREI.1 (voir 4), qu'on trouve dans *Jean a vendu sa maison à Pierre pour 50 millions*.

**NB : 1.** Il va de soi qu'il y a toujours risque d'erreurs dans la détermination du lexème de base, mais cela n'entraîne pas de décisions

irréversibles. Au fur et à mesure de l'analyse lexicologique, les principes logiques de rédaction, appliqués à d'autres exemples d'usage, mettront inévitablement en évidence les éventuelles faiblesses de la description initiale. Cela exigera la refonte d'une partie de la description; c'est le prix à payer dans tout travail lexicologique.

2. La question de la division des acceptions sera considérée plus loin; ici nous nous concentrerons uniquement sur le problème du champ lexical.

VENDREI.1, qui se définit, de façon abrégée, comme 'X transmet à Z les droits de possession permanente de Y [en échange] d'une quantité d'argent W', appartient à un champ sémantique dont la composante identificatrice est 'transmission des droits de possession permanente [en échange] d'argent' et qui inclut également des lexèmes comme ACHETERI, ARGENT4<sub>PR</sub>, MARCHÉ, MARCHANDISE, PRIXI, MARCHANDER, MAGASINI<sub>PR</sub>, COMMERCEI.1<sub>PR</sub>, COMMERÇANT, etc. (l'index PR signale la provenance du numéro lexical : le *Petit Robert*). Cette composante générique peut être exprimée par la lexie ACHAT-VENTE<sup>1</sup>, de sorte que le champ sémantique qui nous intéresse peut être dénommé champ sémantique de l'achat-vente<sup>1</sup>.

Le lexicologue qui veut décrire VENDREI.1 doit donc obligatoirement commencer par le champ sémantique de l'achat-vente et traiter plusieurs lexèmes à la fois, plutôt qu'un lexème isolé. Pour assurer la cohérence des descriptions, la comparaison constante des articles en élaboration s'impose; il faut les ajuster les uns aux autres, en y appliquant les mêmes schémas.

La question des limites de l'exploration du champ se pose immédiatement. Théoriquement, il serait souhaitable de considérer le maximum de lexèmes du champ; mais, pour des raisons pratiques, il faut se restreindre au minimum, surtout au tout début. Cependant, ce minimum doit être SUFFISANT : en tout cas, il doit contenir au moins deux lexèmes opposés (car la langue est organisée par opposition). Normalement, il est mieux de prendre de cinq à sept lexèmes exprimant les éléments centraux de la composante identificatrice; dans notre cas, ce serait, en plus de VENDREI.1, les lexèmes ACHETERI, MARCHANDISE, PRIXI, MAGASINI<sub>PR</sub>, COMMERCEI.1<sub>PR</sub>, avec leurs dérivés. On évitera, d'une part, de travailler au départ sur des lexèmes TROP SPÉCIFIQUES : on ne prend que les « têtes de séries ». Ainsi, on retiendra par exemple VENDEUR, mais pas REPRÉSENTANT, COMMIS, GROSSISTE, etc. (= espèces de VENDEUR); si on traite MAGASINI<sub>PR</sub>, on n'examinera pas BOUTIQUE, ÉCHOPPE, BOULANGERIE, LIBRAIRIE, etc. (= espèces de MAGASINI<sub>PR</sub>). D'autre part, on laissera de côté les lexèmes TROP GÉNÉRAUX, parce qu'ils débordent le champ à l'étude, en faisant charnière avec d'autres champs lexicaux. Ainsi, on évitera de

1 ACHAT-VENTE est une lexie d'usage restreint, qui apparaît surtout comme titre de rubrique dans la presse.



commencer par ARGENT<sup>4</sup><sub>PR</sub> ou PAYERI.1a, sans doute éminemment pertinents pour le champ de l'achat-vente, mais appartenant également au champ sémantique beaucoup plus général des transactions commerciales, qui recouvre, en plus, la location, les services, la rémunération, etc. Bien entendu, tous ces lexèmes seront traités à leur tour, mais lors d'une étape ultérieure. Nous voulons seulement souligner l'importance du choix de départ : un groupe de lexèmes du champ sémantique voulu, qui soit assez riche pour bien révéler la problématique lexicologique du champ et assez circonscrit pour une appréhension cognitive globale.

Une fois la sélection des lexèmes de départ pour le champ sémantique effectuée, le lexicologue réunit tous les vocables auxquels les lexèmes de départ appartiennent. Dans notre cas, on aurait, par exemple, l'ensemble des vocables suivants :

VENDRE, SE VENDRE, VENDEUR, VENTE, ACHETER, ACHAT, ACHETEUR, MARCHANDISE, PRIX, MAGASIN, COMMERCE.

Cet ensemble forme une partie du champ lexical de l'achat-vente, sur laquelle le chercheur voulant décrire VENDREI.1 se doit de travailler simultanément. Il a à considérer à la fois la cohérence d'un lexème du champ sémantique de l'achat-vente par rapport aux autres lexèmes du même champ – démarche « horizontale » (c'est-à-dire VENDREI.1 en relation avec ACHETERI, PRIXI, etc.), et par rapport aux autres lexèmes du même vocable – démarche « verticale » (c'est-à-dire VENDREI.1 en relation avec VENDREI.2 et VENDREII/III; ACHETERI en relation avec ACHETER2a [*Jean a acheté le silence du témoin*], ACHETER2b [*Cette somme achètera le silence du témoin*], ACHETER3a [*Jean a acheté sa tranquillité d'esprit grâce à son travail*] et ACHETER3b [*Ce cadeau nous a acheté la paix*]; PRIXI en relation avec PRIXII [≈ 'importance'] et PRIXIII [≈ 'récompense']; etc.).

**NB** : Il nous a été malheureusement impossible de réaliser dans l'ILEC l'ensemble du programme énoncé ci-dessus. Les contraintes de temps et d'espace nous ont forcés à renoncer à l'analyse des vocables MARCHANDISE, MAGASIN, COMMERCE etc. et à ne présenter, à titre d'illustration, que les trois premiers vocables de notre champ.

Le traitement du lexique par champs assure une description satisfaisante pour les lexèmes du champ sémantique examiné : chacun est étayé HORIZONTALEMENT, par ce champ sémantique justement, ainsi que VERTICALEMENT, par son vocable. La démarche est similaire à celle de la résolution de mots croisés : une solution hypothétique doit se confirmer horizontalement et verticalement. Par contre, les lexèmes du même champ lexical qui n'appartiennent pas au même champ sémantique (et qui ont donc été introduits du fait que l'on considère des vocables entiers) n'ont pas encore une description complètement étayée : leurs champs sémantiques n'ont pas été étudiés. Après l'épuisement du champ sémantique visé, il faut recommencer – de la même façon ! – avec un autre champ sémantique du

même champ lexical. Dans notre exemple, cela serait le champ sémantique de VENDREII, c'est-à-dire le champ de la trahison, le champ sémantique de PRIXIII, c'est-à-dire le champ des récompenses-honneurs, etc.

La caractérisation détaillée du cheminement du lexicologue du DEC dans la description du lexique présentée en 2 s'appuie sur l'exemple du vocable VENDRE; deux autres vocables du champ 'achat-vente' (SE VENDRE et VENDEUR) sont donnés sans commentaire.

## 2 Détermination d'un vocable

Dans la détermination d'un vocable, en application des critères et des règles énoncés aux Ch. III et IV, il faut distinguer les trois tâches suivantes :

- division des acceptions (2.1);
- regroupement des acceptions (2.2);
- organisation du regroupement obtenu (2.3).

### 2.1 Division des acceptions

La première analyse des exemples disponibles fait apparaître trois grands types d'emplois de VENDRE :

- (1)
  - a. *Jean a vendu son ordinateur à Pierre pour 10 000 francs.*
  - b. *Les Indiens ont vendu l'île de Manhattan aux Hollandais.*
  - c. *Le professeur Baudot a vendu ses droits d'auteur à la maison Hachette.*
  - d. *On a vendu mon pauvre ami en esclavage pour trois cents piastres.*
- (2)
  - a. *Gaston a vendu tous ses complices aux policiers.*
  - b. *On trouve toujours un citoyen prêt à vendre sa patrie.*
  - c. *— Je ne vends pas mes principes !*
- (3)
  - a. *Gaston a facilement vendu sa proposition au comité.*
  - b. *Il ne sait pas vendre ses idées !*
  - c. *— Il me reste à vendre mon projet au directeur.*

Les exemples (1) correspondent au champ sémantique de l'achat-vente (transmission des droits de possession [en échange<sup>1</sup> d'argent), les exemples (2), au champ sémantique de la trahison (abandon d'obligations [en échange<sup>1</sup> d'avantages); et les exemples (3), au champ sémantique de la persuasion. Comme dans le deuxième et le troisième cas il ne s'agit pas de vente au sens propre, nous décidons de décrire les emplois correspondants par des lexèmes à part. Les distances sémantiques entre les trois groupes d'exemples nous semblent à peu près égales; nous postulons donc VENDREII pour couvrir les exemples (2), et VENDREIII, pour les exemples (3). On notera que VENDREII n'a pas de conversif ACHETER (*vendre sa patrie* mais \**acheter sa patrie*), celui qui vendII

quelque chose n'est pas un \**vendeur* (mais plutôt un *traître*), etc. VENDREIII, par contre, possède son ACHETER et son VENDEUR, mais il exclut l'échange et n'a donc pas d'actant signifiant 'pour W'; somme toute, VENDREIII n'est pas plus près de VENDREI que VENDREII.

Les exemples du premier type, tout en étant assez homogènes, c'est-à-dire référant à la vente proprement dite, laissent quand même entrevoir deux cas particuliers :

- (4) *Elle vendait son corps (ses charmes, ses faveurs) à tous les marins du port pour des clopinettes.*
- (5) *La ville m'a vendu pour récupérer les impôts.*

En (4), la particularisation se limite uniquement à la chose vendue : ce sont les faveurs sexuelles; aucune contrainte ne s'exerce sur les autres participants de la transaction. De plus, toute la cooccurrence syntaxique et lexicale du verbe VENDRE reste la même : tout ce qu'on peut dire avec *vendre des journaux aux passants* peut s'appliquer à *vendre ses caresses au premier venu*.

Par contre, en (5), la particularisation concerne le vendeur (il ne peut être qu'une autorité administrative ou juridique), le vendu (cela ne peut être que les biens de la personne désignée par le complément d'objet direct) et le but de la vente (il ne peut être que le paiement des dettes encourues par la personne en question envers le vendeur ou une tierce personne). Par conséquent, même si la cooccurrence de ce VENDRE est identique, on a l'impression qu'il s'agit ici quand même d'un sens à part.

La conclusion qui se dégage est que les emplois du type (4) peuvent – et par le principe du rasoir d'Occam<sup>2</sup> doivent – être pris en charge par le même lexème que les autres ventes « normales ». Quant aux emplois du type (5), ils ne peuvent pas, sauf au prix d'un nombre trop élevé de restrictions dans la description, être couverts par le même lexème. Nous décidons donc d'introduire encore deux lexèmes de VENDRE : VENDREI.1 pour les ventes régulières (y inclus la vente des faveurs sexuelles) et VENDREI.2 pour les ventes administratives. (VENDREI.2 est donc un cas particulier de VENDREI.1.)

Sous-jacent au raisonnement ci-dessus, il y a le Critère 2 : Critère de différence sémantique locale/globale (Ch. III, 2.2, p. 63). Une différence sémantique locale, telle '[possession] sexuelle temporaire' dans VENDREI.1, n'entraîne aucune conséquence pour la définition de ce lexème. Elle constitue simplement un supplément d'information à la définition qui ne touche en aucune façon les autres parties de la définition.

2 Guillaume d'Occam [= Ockham], célèbre philosophe anglais du XIV<sup>e</sup> siècle, a formulé ce qui est universellement connu comme Ockham's Razor – un des principes majeurs de la science moderne. C'est le Principe d'économie de description : *Entia non sunt multiplicanda* 'Les entités ne doivent pas être multipliées [inutilement]', ce qui signifie que tout ce qui peut être rasé dans une description doit l'être !

Par contre, une différence sémantique globale, comme 'autorités administratives ou juridiques' en tant qu'actant X dans VENDREI.2 entraîne nécessairement d'autres changements dans la définition, à savoir, ce qui est vendu doit être 'les biens de la personne Y', alors que le but de cette vente est 'pour payer les dettes de Y...'. Cela nous oblige à diviser les acceptions correspondantes et à les présenter comme deux lexèmes : VENDREI.1 (ventes régulières) vs VENDREI.2 (ventes administratives). Cependant, pour ces deux lexèmes, il s'agit d'une vente véritable, de sorte que VENDREI.2 n'est qu'un cas particulier de VENDREI.1; leur différence est donc indiquée par des chiffres arabes.

## 2.2 Regroupement des acceptions

Après avoir distribué tous les emplois observés de VENDRE sous les quatre acceptions ci-dessus – VENDREI.1 et VENDREI.2 (= achat-vente), VENDREII (= traîtrise) et VENDREIII (= persuasion), on doit se demander si ces acceptions sont réunissables dans un même vocable. L'intuition linguistique les perçoit comme telles, et il s'ensuit que nous devons trouver des ponts sémantiques entre ces quatre VENDRE. Pour VENDREI.1 et VENDREI.2, le pont est évident : VENDREI.2 est '≈ vendreI.1 – dans des conditions très spéciales'. La situation est différente pour VENDREII et VENDREIII : ici le lien sémantique se fait par comparaison, puisque VENDREII et VENDREIII signifient '≈ vendreI.1, mais au sens figuré'. Les ponts sémantiques exprimant ce lien se présentent sous la forme de comparaison : 'comme si X vendaitI.1 Y à Z', cette composante étant incluse dans les définitions de VENDREII et VENDREIII. On doit donc en conclure que les quatre lexèmes VENDRE appartiennent au même vocable.

Il faut noter que la composante 'comme si...' est une façon commode de représenter un lien MÉTAPHORIQUE. En fait, une métaphore liant A à B est toujours fondée sur la similitude de A et de B, c'est-à-dire que 'A RAPPELLE B par sa fonction/son aspect/...'. On devrait donc exprimer le pont sémantique métaphorique par la composante 'rappeler', ce que nous faisons d'ailleurs pour les noms : voir Ch. IV, 1.3, Règle 9, p. 164, DOIGT. Cependant, pour les verbes, la formulation avec 'rappeler' serait trop encombrante, de sorte que nous préférons utiliser son synonyme 'comme si'.

Ajoutons que s'il y a comparaison, il y a nécessairement une BASE de comparaison. Dans notre cas, les bases de comparaison de VENDREII et VENDREIII avec VENDREI.1 seront les CONNOTATIONS de ce dernier, voir ci-dessous, 3.1.2, p. 190.

## 2.3 Organisation du regroupement obtenu

Le regroupement de lexèmes nous conduit aux vocables. Chose certaine, un vocable est une structure hiérarchisée qui découle de l'existence des ponts

sémantiques. Dans le cadre du DEC, dictionnaire imprimé, cette hiérarchie, qui n'est nullement linéaire, doit être projetée sur du papier, tout en respectant les conventions typographiques pour les dictionnaires, donc être représentée linéairement. (Rappelons à nouveau – voir p. 170 – que le lexicologue fait face ici aux mêmes problèmes que le géographe qui veut représenter sur une carte la surface du globe terrestre.) Ces problèmes de représentation sont traités, comme nous l'avons démontré, selon deux axes : l'ordonnement des articles de dictionnaire au sein d'un superarticle et l'indication des distances sémantiques entre les lexies appartenant au même vocable (au moyen de numéros lexicaux).

### 2.3.1 Ordonnement des articles de dictionnaire au sein d'un superarticle

Conformément aux règles d'ordonnement de lexies (Ch. IV, 1.3, p. 162 *ssq.*), nous obtenons l'ordre suivant :

- D'après la *Règle d'inclusion sémantique* [= Règle 7], la lexie VENDREI.1 doit précéder les deux autres lexies, puisque son sens est inclus dans leurs définitions.
- D'après la *Règle de proximité sémantique* [= Règle 8], la lexie VENDREI.2 (spécialisation du sens de VENDREI.1) doit précéder VENDREII (sens métaphorique).
- D'après la *Règle de hiérarchie des composantes sémantiques* [= Règle 10], mais interprétée par rapport aux connotations, la lexie VENDREII doit précéder VENDREIII, puisque la connotation sous-jacente à VENDREII (« Primauté du profit sur les considérations morales ») précède celle sous-jacente à VENDREIII (« Capacité de convaincre », qui est plus éloignée de l'idée de vente).

C'est pour respecter cet ordre que nous avons numéroté ainsi nos trois lexèmes.

### 2.3.2 Indication des distances sémantiques entre les lexies

Dans le cas à l'étude, les distances sémantiques s'établissent aisément. Comme VENDREII et VENDREIII sont métaphoriques par rapport aux deux autres lexèmes et que les métaphores ne sont jamais trop régulières, le recours aux chiffres romains s'impose. Au sein de VENDREI, la distance sémantique entre les deux lexèmes est faible mais cette distinction est loin d'être régulière, ce qui nous force à choisir les chiffres arabes (plutôt que des lettres).

Ceci termine nos commentaires sur l'établissement du vocable VENDRE, c'est-à-dire du superarticle, formé de quatre articles de dictionnaire; ces derniers sont annoncés dans le tableau synoptique de ce vocable, voir ci-dessous. À présent, nous pouvons passer aux commentaires traitant d'un article de dictionnaire individuel : celui de la lexie de base VENDREI.1.

### 3 Description d'un lexème : VENDREI.1

Comme nous l'avons déjà signalé, pour des raisons d'espace, nous ne traiterons ici que de la description d'un seul lexème, à savoir VENDREI.1. Il est pourtant évident que la méthodologie d'analyse lexicologique présentée reste valable pour n'importe quelle lexie.

Notre exposé suivra fidèlement l'ordre des zones constituant un article de dictionnaire du DEC :

- zone sémantique (3.1);
- zone syntaxique (3.2);
- zone de combinatoire lexicale (3.3);
- zone d'exemples (3.4);
- zone phraséologique (3.5).

#### 3.1 Zone sémantique

Comme on le sait déjà, la zone sémantique de l'article de dictionnaire de L inclut la définition lexicographique de L et ses connotations.

##### 3.1.1 Définition lexicographique

Conformément à la Règle 1 (*Règle de la forme propositionnelle*), nous établissons tout d'abord pour VENDREI.1 sa *forme propositionnelle* :

*X vend Y à Z pour W,*

qui représente par les quatre variables les quatre ASém de la transaction ACHAT-VENTE. Soulignons que l'ASém W (= argent ou quelque chose d'équivalent) est nécessaire puisque sans W, la transaction ne s'appelle plus *vendre* en français, mais *donner*.

Suivant les prescriptions des Règles 2 et 3 (*décomposition sémantique et bloc maximal*), nous cherchons des composantes sémantiques qui soient plus simples que le sens 'vendreI.1' mais, en même temps, les plus englobantes possibles.

En tant que composante générique et centrale, nous retenons '...transmet à Z les droits de possession de Y...'; cela signifie que VENDREI.1 est décrit comme un cas particulier de transmission des droits de possession. Nous ajoutons les modificateurs '...complète et permanente...' à '...possession de Y...' afin de différencier VENDREI.1 de PRÊTER et LOUER, où la possession est incomplète et temporaire.

La composante '...[en échange] de...' départage les verbes comme VENDREI.1, PRÊTER (au sens bancaire) et LOUER et les verbes comme DONNER ou FOURNIR.

La composante ('... une quantité d'argent...') est nécessaire pour marquer VENDRE<sub>L.1</sub> par rapport à TROQUER / ÉCHANGER. En effet, la phrase (6a) est incorrecte, et il faut dire (6b) :

- (6) a. \*Alain a vendu son ordinateur pour trois bouteilles de Château-Margaux.  
 b. Alain a  $\left\{ \begin{array}{l} \text{troqué} \\ \text{échangé} \end{array} \right\}$  son ordinateur contre trois bouteilles de Château-Margaux.

La précision ('... ou l'équivalent') permet de rendre compte des cas de vente où W n'est pas de l'argent même si les négociations préalables sont faites en termes d'argent. Ainsi, on peut dire :

- (7) L'Irak vendait son pétrole pour de l'armement.

Enfin, la composante ('... que X demande à Z pour Y') fait ressortir l'équivalence entre la marchandise Y et la somme W payée qui est présumée par le vendeur X et l'acheteur Z.

En vertu de la Règle 4 (*standardisation*), nous avons désambiguïsé tous les éléments lexicaux dans le définissant; comme les lexies correspondantes ne sont pas encore élaborées dans le DEC, nous avons emprunté tels quels les numéros lexicaux au *Petit Robert*. Le résultat illustre bien la définition à laquelle on aboutit en observant scrupuleusement les règles d'élaboration d'une définition lexicographique rigoureuse et formelle.

Quant à la Règle 5 (*substituabilité absolue du défini et du définissant dans tous les contextes possibles*), nous ne pouvons donner ici que deux exemples :

- (8) a. Substitution du défini par le définissant :

*Au marché, on vendait des vélos neufs pour cent dollars* = 'Au marché, on transmettait<sub>L.1</sub> à des gens les droits<sup>3</sup><sub>2</sub> de possession<sub>L.1</sub> complète<sup>1</sup><sub>1</sub> et<sub>L.1</sub> permanente<sub>L.1</sub> des vélos neufs<sup>2</sup><sub>L.1</sub> [en échange<sup>1</sup> de 100 dollars qu'on demandait<sub>L.4</sub> à des gens pour ces vélos]'

Il nous semble que cette substitution est concluante : tout ce que le locuteur peut extraire de la phrase de départ, il peut également l'extraire de la représentation sémantique résultante.

- b. Substitution du définissant par le défini :

Imaginons que Jean entre dans la librairie Gibert avec un livre dans la main; il le remet à un commis, qui lui passe une carte intitulée « Bon d'achat de 100 francs ». À cette situation, nous pouvons faire correspondre le sens suivant :

'Jean a transmis<sub>L.1</sub> à la librairie Gibert les droits<sup>3</sup><sub>2</sub> de possession<sub>L.1</sub> complète<sup>1</sup><sub>1</sub> et<sub>L.1</sub> permanente<sub>L.1</sub> de son livre [en échange<sup>1</sup> d'un équivalent de 100 francs]'

La définition proposée de VENDRE<sub>L.1</sub> permet d'exprimer ce sens en français comme suit :

*Jean a vendu son livre à la librairie Gibert pour 100 francs.*

Cette phrase est tout à fait légitime dans les circonstances.

Les exemples (8) prouvent seulement que notre définition PEUT bien fonctionner, au moins dans les contextes cités. Il nous est impossible, comme nous l'avons déjà affirmé, de prouver qu'elle POURRA fonctionner dans tous les contextes. Il appartient à nos lecteurs d'apporter la contre-preuve !

### 3.1.2 Connotations

Le verbe VENDREI.1 manifeste, dans notre description, les deux connotations suivantes.

Premièrement, il connote la disposition de privilégier le gain au détriment des considérations morales. C'est sur cette connotation que se fondent le lexème VENDREII et plusieurs phrasèmes du type « VENDRE LA MÈCHE » ou « VENDRE PÈRE ET MÈRE ». Cette connotation communique à VENDREI.1, dans ses emplois figurés, un certain caractère péjoratif.

Deuxièmement, il connote la capacité de persuader en évoquant les avantages ou gains possibles. Cette connotation se retrouve dans le lexème VENDREIII et dans le phrasème « VENDRE DES FRIGIDAIRES AUX ESQUIMAUX ».

## 3.2 Zone syntaxique

Le Tableau de régime ayant déjà été expliqué (Ch. III, 3.4, p. 119), nous nous contenterons ici des deux observations suivantes concernant le Tableau de régime de VENDREI.1, que nous reproduisons ci-dessous pour faciliter la compréhension des explications.

### VENDREI.1

#### Régime

X = I	Y = II	Z = III	W = IV
1. N	1. N  obligatoire	1. à N	1. Num N 2. à Num N 3. pour N

Dans la colonne  $C_{II}$  [celle de l'Y : la marchandise], on voit l'indication « obligatoire ». Elle signifie que, pour ne pas produire une phrase agrammaticale, il faut nécessairement faire apparaître l'expression de cet actant à la surface. Les autres actants syntaxiques profonds – III et IV – ne sont pas obligatoires :



(9) a. *Il vend du chocolat.*

est une phrase normale, alors que

b. \**Il vend aux étudiants.*

et

c. \**Il vend pour trois cents dollars.*

sont impossibles – sauf comme énoncés elliptiques dans un contexte qui les complète.

**NB :** L'actant II peut perdre son caractère obligatoire dans des conditions particulières, par exemple, dans une phrase générique où le verbe en question se trouve à l'infinitif :

(10) *Vendre, c'est tout un métier !*

De plus, l'actant II cesse d'être obligatoire si VENDRE est utilisé pour marquer l'occupation habituelle : *Notre compagnie vend surtout à l'étranger.* Cette particularité est couverte par la contrainte 2 sur le régime de VENDREI.1, p. 196.

Dans C<sub>IV</sub> [= colonne 4, celle du W : le prix], nous avons noté la possibilité d'exprimer l'ASyntP IV à la surface soit tout simplement par un syntagme numérique, soit par un syntagme prépositionnel avec POUR : *vendre sa maison 50 000 francs/pour 50 000 francs*. Comme ces deux expressions sont toujours mutuellement substituables, il n'y a pas lieu d'indiquer les conditions de leur distribution. Pourtant, la préposition À, qui est aussi toujours possible, exprime une nuance sémantique : le prix est avantageux pour l'acheteur (*Je te vends cette bouteille 100 francs vs à 100 francs*). Cela est pris en charge par la contrainte 3, p. 196.

Quant à la contrainte 1, nous avons déjà précisé que nous voulons couvrir par VENDREI.1 aussi le sens de prostitution. Cette contrainte porte donc sur l'actant II (= C<sub>II</sub>) et stipule que, si le sens à exprimer est '[vendreI.1] les droits de possession SEXUELLE TEMPORAIRE de sa personne', alors l'actant II doit être un des noms listés (CORPS<sup>1</sup>II.1a, CHARMES, FAVEURS, CARESSES, CUL, ...).

L'exemple qui suit – *Jean a vendu son automobile (à Pierre)...* – illustre simplement la combinabilité des quatre actants, en précisant, par des parenthèses, le caractère optionnel de l'expression des actants SyntP III et IV à la surface.

### 3.3 Zone de combinatoire lexicale

Lors de la discussion des fonctions lexicales dans le Ch. III, 3.5, p. 125 *ssq.*, nous n'avons pas précisé comment dégager le jeu des valeurs des FL pertinentes pour la lexie L. Même si nous ne connaissons pas de méthodologie formelle qui permettrait de le faire de façon mécanique en garantissant de bons résultats, nous croyons quand même utile de conseiller trois approches à utiliser en parallèle :

- Partir de la liste préétablie des FL en les appliquant une à une à L et en essayant de trouver les valeurs correspondantes.
- Compiler une liste d'expressions figées, semi-figées et quasi figées qui renferment L, en s'appuyant sur l'intuition des locuteurs et sur les documents lexicologiques (dictionnaires de toutes sortes); analyser cette liste, en en dégageant les valeurs des FL.
- Consulter les bases de données textuelles, en recherchant les expressions avec L qui ne semblent pas être libres mais qui pourtant ne sont pas encore couvertes par la description en cours; les expressions ainsi trouvées fourniront d'autres valeurs des FL (surtout, des FL complexes et non standard).

Ces trois procédés sont complémentaires, et le lexicologue les utilisera de façon cyclique – en vérifiant la description obtenue pour sa cohérence formelle interne (au sein de la lexie) et externe (dans le vocable et dans les champs sémantique et lexical).

Commentons à présent quelques FL de VENDREI.1.

1. La FL **Syn**<sub>1</sub> indique un synonyme plus riche (Ch. III, 3.5.2, n°1, p. 129); en effet, *écouler*<sup>II</sup> et *débiter*<sup>2</sup>, ainsi que *laisser*<sup>III.4</sup> sont des cas particuliers de *vendre*<sup>I.1</sup> : chaque fois qu'on écoule<sup>II</sup> Y, on le vend<sup>I.1</sup>, l'inverse n'étant pas vrai, etc.
2. Par contre, *céder*<sup>I.2</sup> n'est pas un synonyme plus riche ni moins riche que *vendre*<sup>I.1</sup> : dans certains cas, on vend<sup>I.1</sup>, mais on ne cède<sup>I.2</sup> pas (– *Où vend-on l'\*cède-t-on des skis ?*), dans d'autres, on cède<sup>I.2</sup> sans vendre<sup>I.1</sup> (*J'ai cédé l'\*vendu mon bail à Jean*). C'est donc un synonyme à intersection de sens : **Syn**<sub>1</sub>.
3. ACHETER<sup>1</sup> est le **Conv**<sub>3214, 1</sub>(VENDREI.1) puisqu'on a :  
*Pierre [= I] a vendu un vélo [= II] à Marie [= III] pour cent francs [= IV] ⇔ Marie [= I] a acheté un vélo [= II] à Pierre [= III] pour cent francs [= IV].*  
 Cette transformation illustre la permutation des actants. Quant à l'intersection sémantique, elle s'explique par l'attitude divergente de l'actant I de chacun de ces deux verbes : dans VENDREI.1, il veut l'argent, dans ACHETER<sup>1</sup>, la marchandise.
4. Le nom d'action **S**<sub>0</sub> –VENTE1a – remplace le verbe sans affecter le sens dans toutes les nominalisations : *La vente du vélo par Pierre à Marie a été rapidement conclue ⇔ Pierre a rapidement vendu le vélo à Marie*, etc.
5. Les noms des quatre actants de VENDREI.1 sont :  
 I – celui qui vend<sup>I.1</sup> (= **S**<sub>1</sub>) s'appelle *vendeur*<sup>I.a</sup> : *Pierre, le vendeur du vélo, sera responsable pour son entretien pendant un an*. Cependant, la personne qui vend professionnellement est un *vendeur*<sup>I.b</sup>, ce qui est un autre lexème. (Cf. *Jean est vendeur de vélos l'\*des vélos, \*du vélo*; VENDEURI.b a un actant (optionnel) de plus, l'employeur, puisqu'on dit *Jean est vendeur chez Tati*). Les deux lexèmes sont sémantiquement

très proches, ce qui est exprimé par les minuscules dans le numéro lexical (qui indiquent la distance minimale). Quant à MARCHANDI, ce lexème désigne aussi un vendeur professionnel, mais avec des composantes sémantiques supplémentaires (un marchand est propriétaire des marchandises à vendre, il les achète, etc.), ce que marque l'indice  $\supset$ .

II – la chose vendue est une *marchandiseI*; mais, comme certaines choses qui se vendentI.1 ne sont nullement des marchandises, les services, par exemple, c'est un nom d'objet plus spécifique, donc plus riche :  $S_{2\supset}$ .

III – celui à qui on vendI.1 (=  $S_3$ ) s'appelle *acheteurI*.

IV – l'argent reçu pour la marchandise vendue (=  $S_4$ ) s'appelle *prixI*.

Les phrasèmes « une bouchée de pain » et « des clopinettes » désignent un prix [=  $S_4$ ] très bas [= *AntiMagn*], ce qui explique leur description. Ce sont sémantiquement des désintensificateurs qui s'utilisent syntaxiquement comme des compléments.

6. Le nom de lieu abstrait  $S_{loc}$  est le lexème MARCHÉI, désignant « là où l'on vendI.1 dans l'abstrait » : *conquérir le marché asiatique, les marchés financiers, prix du marché, ...* La FL  $S_{loc\supset}^{usual}$  (nom de lieu usuel, c'est-à-dire concret) donne les lexies qui désignent les lieux aménagés où l'on vendI.1 : dans un premier temps, la lexie la plus générale, presque terminologique « POINT DE VENTE », puis les lexies plus spécifiques – mais quand même encore très générales. En effet, nous ne citons, dans des cas semblables, que des « têtes de séries » quasi synonymiques – les termes génériques ou archilexies; les termes spécifiques se retrouvent dans les articles de dictionnaire de ces derniers. Nous ne citons ici que, par exemple, COMMERCEI.3; les lexies désignant les espèces de commerceI.3 (MAGASINII, « KIOSQUE À JOURNAUX », CONCESSIONNAIRE, ...) sont listées dans l'article de dictionnaire de COMMERCEI.3, alors que celles désignant les espèces de magasinII (BOUTIQUE, LIBRAIRIE, CRÈMERIE, « BUREAU DE TABAC », ...) le sont dans l'article de MAGASINII, etc.
7. L'adjectif  $A_2$  qualifie la chose qu'on vendI.1 : *des salades à vendre, Ce n'est pas à vendre, ...* L'adjectif potentiel  $Able_2$  qualifie la chose qu'on peut vendreI.1 : *des salades vendables, Ce n'est pas vendable, ...*, *AntiAble\_2* en étant l'antonyme. Pour exprimer le sens « être [= *Pred*] invendable [= *AntiAble\_2*] », le français a le phrasème « ne pas trouver acheteur (acquéreur) » [= *PredAntiAble\_2*].
8. *Magn\_4* et *AntiMagn\_4* se rapportent à l'actant IV, donc au prix. « VendreI.1 « à vil prixI » se dit BRADERI, qui est un élément fusionné (/) de la valeur de la FL *AntiMagn\_4*.
9.  $A_2nonPerf$  = « tel que l'on ne l'a pas venduI.1 » (*des oranges invendues*),  $S_2nonPerf$  = « chose que l'on n'a pas vendueI.1 » (– *Montre-moi un de tes invendus, qu'il faudra solder*).

10. À la suite des FL standard, nous listons les FL non standard; comme le lecteur le voit, chacune est formulée en français. Nous tenons à attirer l'attention sur la toute dernière FL non standard : l'expression «à vendre», qu'on met sur les pancartes, n'est pas libre et doit être listée. En effet, on ne dit pas \**pour vente* ou \**en vente* (cf. anglais *for sale*) ou bien \**X se vend* (cf. espagnol *Se vende caballo = Cheval à vendre*). L'expression «à vendre» comme annonce commerciale ne peut être construite par des règles générales du français, et le DEC se doit de la présenter dans l'article de VENDREI.1 comme un semi-phrasème, au moyen d'une FL non standard.

### 3.4 Zone d'exemples

Dans ce cas particulier, la zone illustrative n'appelle aucun commentaire.

### 3.5 Zone phraséologique

Les phrasèmes listés dans cette zone forment tous des vedettes indépendantes dans le DEC : «VENDRE CHÈREMENT», *X vend chèrement Y*, où *Y* est fortement contraint, sans être quand même unique (*vendre chèrement sa vie, son honneur, sa vertu, sa liberté, son indépendance, ...*); «VENDRE LA PEAU DE L'OURS AVANT DE L'AVOIR TUÉ», etc. Leur présence dans la zone phraséologique se justifie par la commodité de référence et par l'indication explicite de l'importance des connotations (pour chaque phrasème lié à la lexie vedette par une connotation, on indique cette connotation).

Le phrasème «SAVOIR SE VENDRE» (*X* sait se présenter de façon avantageuse – comme si *X* se vendaitI.1 «en tant que» marchandise) mérite un commentaire : il n'est PAS construit librement à partir des deux verbes SAVOIR et SE VENDRE; si SAVOIR garde en effet son sens propre, SE VENDRE ne signifie pas 'se présenter de façon avantageuse...' sauf dans cette expression ('*Il sait comment se vendre, \*Il se vend bien, \*Il est capable de se vendre de la meilleure façon, ...*), ce qui revient à dire que \*SE VENDRE 'se présenter de façon avantageuse ...' n'existe pas en français.

**4 Articles de dictionnaire illustratifs : VENDRE, SE VENDRE, VENDEUR**

**VENDRE, verbe.**

- I.1.** X transmet à Z les droits de possession... de Y<sup>1</sup> en échange<sup>1</sup> d'une quantité d'argent ... [*Pierre vend son automobile à Jean pour trois mille dollars*]
- 2.** Autorité juridique X vend<sup>1.1</sup> les biens de... Y pour payer<sup>1.2a</sup> les dettes de Y... [*Les créanciers de Jean avaient envoyé un huissier pour le vendre*]
- II.** X trahit Y auprès de Z<sup>1</sup> en échange<sup>1</sup> d'un bénéfice W... [*Cet homme a vendu ses camarades à l'ennemi*]
- III.** X convainc Z d'accepter<sup>1.1b</sup> la proposition<sup>2</sup> Y<sup>1</sup> de Y<sup>2</sup> de X... [*Cet homme a su vendre son idée au ministre*]

**I.1.** passif en *se*<sup>3</sup>. X vend Y à Z pour W = X transmet<sup>1</sup> à Z les droits<sup>2</sup> de possession<sup>1.1</sup> complète<sup>1</sup> et<sup>1.1</sup> permanente<sup>1</sup> de Y<sup>1</sup> en échange<sup>1</sup> d'une quantité<sup>1</sup> d'argent<sup>4</sup> (ou l'équivalent<sup>21</sup>) W que X demande<sup>1.4</sup> à Z pour Y [= Conv<sub>3214, 6</sub> (acheter<sup>1</sup>)].

**Connotations**

- 1) primauté du profit sur les considérations morales [cf. VENDRE<sup>II</sup>]
- 2) capacité de convaincre [cf. VENDRE<sup>III</sup>]

**Régime**

X = I	Y = II	Z = III	W = IV
1. N	1. N  obligatoire	1. à N	1. Num N 2. à Num N 3. pour N

- 1) C<sub>II</sub> : si (... les droits de possession sexuelle temporaire de la personne Y), alors N = CORPS<sup>II.1a</sup>, CHARMES, FAVEURS, CARESSES, CUL, ...

3 Nous n'avons pas commenté les indications grammaticales puisqu'elles sont transparentes. Ainsi, l'écriture «passif en *se*» signifie que la lexie vedette possède le passif réfléchi (*Le jambon se vend dans notre ville chez deux bouchers à cent francs le kilo*).

- 2)  $C_{II}$  : si V. est utilisé au sens habituel (se consacrer aux activités de vente), alors  $C_{II}$  n'est pas obligatoire
- 3)  $C_{IV.2}$  : si (... W étant avantageux pour Z) alors  $C_{IV} = C_{IV.2}$
- $C_I + C_{II} (+ C_{III}) (+ C_{IV})$  : *Jean a vendu son automobile (à Pierre) (trois mille dollars (à /pour trois mille dollars, pour une bouchée de pain))*

### Fonctions lexicales

- $Syn_{\supset}$  : écouler<sup>II</sup>, débiter<sup>2</sup>; laisser<sup>III.4</sup>
- $Syn_{\cap}$  : céder<sup>I.2</sup>
- $Conv_{3214\cap}$  : acheter<sup>1</sup>
- $S_0$  : vente<sup>1a</sup>
- $S_1$  : vendeur<sup>I.a</sup>
- professionnel- $S_1$  : vendeur<sup>I.b</sup>
- professionnel- $S_{1\supset}$  : marchand<sup>I</sup>
- $S_{2\supset}$  : marchandise<sup>1</sup>
- $S_3$  : acheteur<sup>1</sup>
- $S_4$  : prix<sup>I</sup>
- AntiMagn $S_4$  : fam<sup>1</sup> «une bouchée de pain», fam<sup>1</sup> «des clopinettes»<sup>1</sup>
- $S_{\substack{loc \\ usual \\ loc\supset}}$  : marché<sup>1</sup>
- $S_{loc\supset}$  : «point de vente»<sup>1</sup>; marché<sup>2</sup>; commerce<sup>I.3</sup>; pharmacie<sup>2</sup>
- $A_{2\supset}$  : en vente | Y est une marchandise régulière [*Les billets de loterie sont en vente chez tous les marchands de journaux*]
- Able<sub>2</sub> : vendable
- AntiAble<sub>2</sub> : invendable
- PredAntiAble<sub>2</sub> : «ne pas trouver acheteur (acquéreur)»<sup>1</sup>
- Magn<sub>4</sub> : cher
- AntiMagn<sub>4</sub> : «bon marché», «à vil prix<sup>I</sup>» // brader<sup>1</sup>
- $A_2$ nonPerf : invendu [Adj]
- $S_2$ nonPerf : invendu [Nom]
- $F_1 = V.$  Y en grandes quantités à un seul Z : spéc ~ en gros > spéc ~ en demi-gros
- professionnel- $S_1F_1$  : grossiste
- $F_2 = V.$  séparément les éléments composant un lot de marchandises<sup>1</sup> : spéc ~ au détail // spéc faire le détail
- professionnel- $S_1F_2$  : détaillant

- V. à un prix **I** réduit : ~ au rabais // solder<sup>3</sup>, liquider<sup>2</sup>
- V. Y à un prix **I** qui est plus bas que le prix **I** coûtant : ~ à perte // sacrifier **I.3**
- V. Y que X a acheté **I**  
dans le but de le V. avec un bénéfice : // revendre
- essayer de façon très active de V. Y en faisant la réclame à Z : // pousser à la vente **1a**
- V. dans le cadre d'enchères : ~ aux enchères
- V. en acceptant **I.4a** des paiements **I.1a** différés : ~ à crédit (à tempérament)
- V. en demandant d'abord un engagement écrit de l'acheteur **I** : ~ par souscription
- le propriétaire de l'objet en cause communique qu'il veut le V. : à ~ [formule écrite]

### Exemples

On vendait des galettes anisées, des pains de seigle d'un sou, des pommes vertes. Les producteurs agréés sont ceux qui vendent au moins 10 000 litres de lait par année. Sur le timbre de trois francs émis récemment par le Maroc, on voit la Super-Caravelle, le type d'avion que la compagnie Sud-Aviation a tenté de vendre à Air Canada. Étant donné les travaux qu'il y a à faire, je ne la vends pas cher, cette maison, je la vends 80 000 francs. Ses parents vendaient en gros des produits chimiques. Les conseillers financiers vendent leurs services à des investisseurs. Alice Hébert vendait trois jours par semaine sur les divers marchés. Ce produit se vend dans toute la France, aux adultes seulement, plus de 500 francs le paquet. Devant chaque maison, on voyait l'écriteau « À VENDRE ».

◇

Vendre chèrement [N]

Vendre la peau de l'ours avant de l'avoir tué

**Rattachés à VENDREI.1 par la connotation « primauté du profit » :**

Vendre la mèche

Vendre père et mère

Vendre son âme (au diable)

Vendre son droit d'aïnesse pour un plat de lentilles

**Rattachés à VENDREI.1 par la connotation « capacité de convaincre » :**

Savoir se vendre

Vendre des frigidaireaux aux Esquimaux

**I.2.** pas de passif en *se*, **spéc.** *X vend Y à Z pour W* = Autorité juridique X vend I.1 les biens d'une personne Y à Z pour W dans le but de payer I.2a les dettes de Y à α avec W.

**Régime**

X = I	Y = II	Z = III	W = IV
1. N	1. N obligatoire	1. à N	1. Num N 2. pour N

$C_I + C_{II} (+ C_{III}) + (C_{IV})$  : *Les autorités judiciaires ont vendu cet homme (à la compagnie Tartempion) (pour 3 000 dollars)*

**Fonctions lexicales**

officier ministériel dont la fonction est de saisir et V. les biens de Y : // huissier

**Exemples**

Les créanciers de Jean avaient envoyé un huissier pour le vendre. – Si vous faites faillite, ils viendront vous vendre.

**II.** passif en *se*. *X vend Y à Z pour W* = X trahit Y auprès de Z [en échange] d'un bénéfice W [comme si X vendait I.1 Y à Z pour W].

**Régime**

X = I	Y = II	Z = III	W = IV
1. N	1. N obligatoire	1. à N	1. [en échange] de N 2. pour N

1)  $C_{IV,1}$  : N ne désigne pas un avantage matériel  
 2)  $C_{IV,2}$  : N désigne un avantage matériel  
 $C_I + C_{II} (+ C_{III}) + (C_{IV})$  : *Cet homme a vendu ses complices (à la police) (en échange de sa liberté (pour des clopinettes))*



### Fonctions lexicales

Syn <sub>o</sub>	: trahir, donner
Syn <sub>n</sub>	: dénoncer
S <sub>1</sub>	: vendu
S <sub>1c</sub>	: traître; <b>fam</b> cafard, <b>fam</b> mouchard
AntiMagnS <sub>4</sub>	: <b>fam</b> «une bouchée de pain», <b>fam</b> «des clopinettes», rien, «trois fois rien»

### Exemples

Judas a vendu Jésus aux prêtres pour trente deniers. À propos, je vais te raconter une histoire, tu ne me vendras pas [H. de Montherlant].

III. passif en *se*. *X vend Y à Z = X convainc Z d'accepter*. 1.1b la proposition 2 Y<sup>1</sup> de Y<sup>2</sup> de X [comme si X vendait 1.1 Y à Z]<sup>4</sup>.

### Régime

X = I	Y = II	Z = III
1. N	1. N obligatoire	1. à N

C<sub>I</sub> + C<sub>II</sub> (+ C<sub>III</sub>) : *Cet homme a su vendre son idée (au ministre)*

### Fonctions lexicales

Conv <sub>321</sub>	: <b>fam</b> acheter III
S <sub>1</sub>	: vendeur II

### Exemples

Pour bien vendre sa réforme électorale, le sénateur McCloskey a dû faire des concessions à l'opposition. Depuis trois mois, le parti essaie de vendre son programme à la population. Sa proposition s'est très bien vendue !

4 Dans cette définition, on a ce que nous appelons une *variable scindée* : Y<sup>1</sup> et Y<sup>2</sup>. De telles variables ne sont jamais exprimées ensemble puisqu'elles représentent en fait deux «aspects» d'une même entité. Ainsi, dans le cas de VENDRE III, Y est toujours la proposition [= Y<sup>1</sup>] de faire quelque chose [= Y<sup>2</sup>], et on peut exprimer dans une phrase – en tant qu'ASyntP du verbe – soit Y<sup>1</sup>, soit Y<sup>2</sup> : *Jean essaie de vendre au directeur son projet [= Y<sup>1</sup>] de voyage = ... de vendre au directeur son voyage [= Y<sup>2</sup>]*.

**[SE] VENDRE, verbe, pron-acc.**

- I. X est tel qu'on l'achète<sup>1</sup> au rythme Y [*Les livres se vendent bien*]
- II.1. X... passe du côté de Y, qui est un ennemi... [en échange] d'un bénéfice Z... [*Il s'est vendu aux boches*]
- 2. X rend à Y des services [en échange] d'un bénéfice Z, ce qui va [à l'encontre] des normes morales... [*Ces fonctionnaires se vendent aux hommes d'affaires*]

I. *X se vend Y-ment* = X est tel qu'on l'achète<sup>1</sup> au rythme Y.

**Régime**

X = I	Y = II
1. N	1. Adv obligatoire

- 1)  $C_2$  : si Y = 'bien', alors  $C_2 = \Lambda$  est possible
- $C_I + C_{II}$  : *Ces livres se vendent bien*

**Fonctions lexicales**

- $S_2$  : rythme [de vente]
- $Magn_2$  : beaucoup, bien, facilement < **fam** [comme des petits pains]
- Anti $Magn_2$  : peu, mal, difficilement
- produit qui SE V.  $Magn_2$  : produit vendeur

**Exemples**

Les billets pour ce spectacle se vendent très bien (comme des petits pains). Ce vin se vend de mieux en mieux en dehors du pays.

II.1. *X se vend à Y pour Z* = X, qui appartient à une communauté, passe du côté de Y, qui est un ennemi de cette communauté, [en échange] d'un bénéfice Z [comme si X vendait I.1 lui-même à Y pour Z] [= Conv<sub>213 $\alpha$</sub>  (*acheter2a*)].

**Régime**

X = I	Y = II	Z = III
1. N	1. à N	1. [en échange] de N 2. pour N

$C_I (+ C_{II}) (+ C_{III})$  : *Ils se vendent (aux Allemands) (pour un plat de lentilles)*

**Fonctions lexicales**

Conv<sub>213</sub> : acheter2a  
 S<sub>I</sub> : vendu  
 S<sub>I</sub> : traître, renégat  
 AntiMagnS<sub>3</sub> : **fam** «une bouchée de pain», **fam** «des clopinettes», «un plat de lentilles», rien, «trois fois rien»

**Exemples**

C'est un traître, un type qui s'est vendu aux boches. – Fais attention à Pierre, il est prêt à se vendre à n'importe qui.

**II.2.** *X se vend à Y pour Z = X rend à Y des services «en échange» d'un bénéfice Z, ce qui va «à l'encontre» des normes morales [comme si X vendait I.1 lui-même à Y pour Z] [= Conv<sub>213</sub>(acheter2a)].*

**Régime**

X = I	Y = II	Z = III
1. N	1. à N	1. «en échange» de N 2. pour N

$C_I (+ C_{II}) (+ C_{III})$  : *Ces députés se vendent (aux hommes d'affaires) (pour quelques avantages)*

**Fonctions lexicales**

Syn : se prostituer  
 Conv<sub>213</sub> : acheter2a  
 AntiMagnS<sub>3</sub> : **fam** «une bouchée de pain», **fam** «des clopinettes», rien, «trois fois rien»  
 Qual<sub>I</sub> : corrompu  
 Magn<sub>3</sub> : cher, chèrement  
 AntiMagn<sub>3</sub> : «à vil prix II.2»  
 Caus<sub>2</sub> : // corrompre

**Exemples**

Il est des gens qui crient très haut qu'ils n'ont jamais voulu se vendre, c'est peut-être qu'on ne se serait jamais soucié de les acheter [G. de Nerval]. Ayant perdu ses parents, Jeanne se vendait dès l'âge de 15 ans.

**VENDEUR**

- I. a. Personne (de sexe masculin) qui vendI.1 ... Y... [*le vendeur de cette voiture*]
- I. b. Personne (de sexe masculin) qui vendI.1 [en tant que] profession des marchandises1 Y... [*un vendeur de vêtements dans un grand magasin*]
- II. Personne (de sexe masculin) qui est apte à vendreIII n'importe quoi à n'importe qui [*un bon vendeur d'idées*]
- III. Qui aide ... à vendreIII à n'importe qui [*une idée vendeuse*]

I. nom, masc [fém -euse]

I.a. *Vendeur de Y* = Personne (de sexe masculin)<sub>gén</sub> qui vendI.1 ou a venduI.1 Y [= S<sub>1</sub>(vendreI.1) = S<sub>3</sub>(acheterI)].

**Régime**

Y = II
1. de N 2. A <sub>poss</sub>

C<sub>II</sub> : *le vendeur de cette voiture, son vendeur*

**Fonctions lexicales**

femme : vendeuseI.a

**Exemples**

Le vendeur doit à l'acheteur la possession entière de la chose vendue.

I.b. *Vendeur de Y (chez U)* = Personne (de sexe masculin)<sub>gén</sub> qui vendI.1 [en tant que] profession des marchandises1 Y (étant employé par l'entreprise U) [= professionnel-S<sub>1</sub>(vendreI.1)].

**Régime**

Y = II	(U = III)
1. de N	1. chez N 2. dans N 3. de N 4. pour N

- 1) C<sub>III.1</sub> : N est un nom propre de personne  
 (C<sub>II</sub>) (+ C<sub>III</sub>) : un vendeur (de vêtements) (chez ⟨de, pour⟩ Eaton), un vendeur de manomètres dans une société de construction automobile

**Fonctions lexicales**

- femme : vendeuse**I.b**  
 qui se déplace à divers endroits pour vendre**I.1** : ambulant, itinérant

**Exemples**

Dans ce magasin, les vendeurs doivent porter l'uniforme. Le vendeur était très aimable, il m'a bien conseillé dans mon choix. Le principal vendeur brestois de poissons, M. Letulier, a été arrêté hier pour possession de drogue. Les vendeurs d'armes peu scrupuleux profitent de la tension au Proche-Orient.

◇

Vendeur de rêves

**II.** nom, masc [fém -euse]. *Vendeur* [de Y] = Personne (de sexe masculin)<sup>gén</sup> qui est apte à vendre**III** n'importe quoi à n'importe qui [= S<sub>1</sub>(vendre**III**)].

**NB** : S'utilise surtout avec (Anti)Magn.

**Fonctions lexicales**

- Magn : bon | prépos < excellent | prépos et postpos  
 AntiMagn : mauvais, piètre | prépos  
 femme : vendeuse**III**

**Exemples**

-- Pour obtenir la majorité, il nous faut un bon vendeur.

**III.** adjectif [fém -euse]. [X] *vendeur* = [X] qui aide α à vendre**III** β à n'importe qui.

**Exemples**

Avec ce programme vendeur, le parti va gagner les élections. – C'est une idée vendeuse, elle sera acceptée par tous.



# CHAPITRE VI

## INFORMATISATION

### DU DICTIONNAIRE EXPLICATIF ET COMBINATOIRE

On peut considérer qu'à ce stade de notre exposé, nous avons terminé la présentation des bases de la lexicologie explicative et combinatoire. Il ne nous reste plus maintenant qu'à traiter le problème fort important de l'*informatisation* du dictionnaire postulé par notre approche – du DEC. Nous le faisons en quatre étapes :

- présentation des différents aspects de l'informatisation du DEC considérés ici et justification du travail sur un DEC simplifié, appelé DiCo (1) ;
- description détaillée du DiCo (2) ;
- présentation d'un éditeur spécialisé, l'ÉDiCo, pour l'écriture du DiCo (3) ;
- examen des possibilités d'utilisation du DEC ou DiCo informatisés pour le traitement automatique du langage naturel (4).

#### 1 Qu'est-ce qu'un DEC informatisé ?

Une mise au point s'impose, avant de passer à l'étude du problème de l'informatisation du DEC. Une telle étude présuppose que l'on se situe à cheval entre deux disciplines : la linguistique comme telle et l'informatique. Ce dont nous allons parler maintenant relève en fait de ce qu'il convient d'appeler la *linguistique informatique*. Malheureusement, cette expression est souvent employée pour référer à des activités très diverses, allant de la linguistique pure, qui utilise l'outil informatique pour aider à la recherche, aux applications tout à fait pratiques du type systèmes de traduction automatique commerciaux, interfaces aux bases de données en langue naturelle, etc. Il est bien évident que, dans le cadre du présent ouvrage, nous nous limitons au premier type d'activité mentionné ci-dessus : c'est-à-dire l'utilisation de l'informatique pour la recherche fondamentale en linguistique. Nous sommes cependant convaincus que ce que nous présentons trouve des débouchés directs dans le cadre d'applications pratiques – c'est-à-dire pour ce que l'on appelle parfois les industries de la langue.

Si l'on considère le point de vue du lecteur, il est clair que nous nous livrons à un exercice périlleux. Certains lecteurs, dont les champs de compétence et les centres d'intérêt se situent plus en lexicologie pure pourront trouver bien des aspects de ce qui suit inutilement formels et techniques; d'autres lecteurs, plus orientés vers l'informatique et alléchés par le titre de ce chapitre, seront vraisemblablement déçus par le manque de propositions suffisamment détaillées et la relative naïveté de notre exposé – pour ce qui a trait aux problèmes concrets de

programmation. Mais nous n'avons pas le choix : il est crucial d'établir la connexion entre le DEC et son informatisation, même si nous courons le risque de nous retrouver ici un peu le cul entre deux chaises.

Cette mise au point étant faite, passons maintenant à la question qui nous préoccupe ici :

|| Que faut-il entendre par *informatisation du DEC* ?

En théorie, cette expression pourrait référer au moins aux trois types de tâches suivants :

1. informatisation des SOURCES LINGUISTIQUES À PARTIR DESQUELLES S'EFFECTUE LE TRAVAIL DE LEXICOLOGIE – c'est-à-dire utilisation de corpus informatisés et de programmes d'analyse de tels corpus (générateurs de concordances, programmes statistiques de repérage automatique des collocations, etc.);
2. traduction de l'INFORMATION CONTENUE DANS LE DEC sous forme de bases de données lexicales informatisées – c'est-à-dire sous forme de fichiers informatiques structurés de telle façon qu'ils permettent une gestion automatique de l'information qu'ils contiennent;
3. informatisation du TRAVAIL DE RÉDACTION du DEC – c'est-à-dire utilisation de programmes informatiques aidant à la construction de dictionnaires du type DEC.

Nous laisserons de côté, dans ce chapitre, la tâche 1 – les problèmes de construction et d'utilisation de corpus informatisés, pour nous concentrer sur les aspects de l'informatisation du DEC relevant des tâches 2 et 3. En effet, les problèmes liés à l'utilisation de corpus informatisés sont les mêmes, *grosso modo*, que l'on travaille sur un DEC ou sur un dictionnaire conventionnel (commercial)<sup>1</sup>.

Examinons maintenant plus en détail ce qui est mis en jeu dans les deux tâches d'informatisation qui nous intéressent ici. Pour ce qui est de la mise sur support informatique des DONNÉES CONTENUES DANS LE DEC, il convient de noter d'ores et déjà que cette tâche n'en est qu'à ses balbutiements et que, pour l'essentiel, tout

1 Deux remarques s'imposent ici.

1. Nous voudrions tout d'abord rappeler qu'il a déjà été fait référence à la fin du Chapitre II (3.2, p. 52) au travail sur le dictionnaire de l'anglais COBUILD, qui est entièrement construit à partir de l'information linguistique extraite de corpus informatiques. On trouvera dans Sinclair 1987, qui est un compte-rendu du projet COBUILD, bon nombre d'observations très pertinentes sur l'utilisation de corpus informatisés en lexicologie. Il faut mentionner, cependant, que la démarche du DEC se distingue radicalement de celle adoptée pour le COBUILD dans la mesure où, pour le DEC, le recours à l'intuition linguistique du lexicologue prime toujours sur l'utilisation des corpus, alors que pour le COBUILD la méthodologie est inverse.
2. Il est intéressant de noter qu'on assiste depuis quelques années à l'apparition d'outils informatiques permettant le repérage automatique des phénomènes collocationnels. De tels programmes, qui font appel à un traitement statistique des corpus, ne sont bien entendu pas capables de tout repérer, et surtout, ils ne sont pas capables de faire le tri de façon sûre entre les véritables collocations et les phénomènes de cooccurrence non restreinte (contrôlés par la sémantique et la syntaxe de la langue). Ce sont néanmoins de très bons outils de dégrossissage et leur utilisation est appelée à se développer.



reste à faire pour la mener à bien (pour une première tentative d'informatisation du DEC dans ce sens, on peut consulter Décary 1986). L'informatisation des données du DEC est un problème fort complexe et nous présenterons – comme première étape vers une solution – une version simplifiée du DEC visant avant tout la description des phénomènes de cooccurrence et pour laquelle un réel effort d'informatisation a déjà été réalisé. Nous avons appelé ce DEC simplifié *DiCo* – acronyme de Dictionnaire de Cooccurrences. La simplification opérée est de trois ordres :

- Les descriptions lexicographiques consignées dans le DiCo ont une moins grande couverture que celles du DEC, puisque, comme nous le verrons plus loin, un DiCo ne possède pas de véritable définition lexicale, la priorité étant donnée aux phénomènes de cooccurrence. Cela nous permet d'éviter d'avoir à rédiger et manipuler les zones sémantiques du DEC, qui sont sans doute les composantes de la description les plus difficiles à formaliser complètement.
- La syntaxe (au sens de 'syntaxe d'un langage formel') des articles de DiCo est encore plus formalisée que celle des articles du DEC. On peut ainsi considérer que le DiCo est formalisé quasiment à 100 %. Cela nous a bien sûr conduits à effectuer des simplifications, quand il était difficile de formaliser tout en gardant à la description la finesse et la précision requise dans un DEC.
- Du fait qu'il n'est pas imprimé mais correspond à une base de données en continuelle évolution et amélioration, le DiCo ne prétend pas à une aussi grande exhaustivité que le DEC, dans la description des lexies individuelles – même si en principe le but ultime est d'obtenir des descriptions complètes. Cet aspect de la simplification opérée dans le DiCo est particulièrement sensible au niveau de la description des fonctions lexicales : il est acceptable, dans une première phase de construction, de ne pas encoder toutes les fonctions lexicales pour une lexie donnée, et de sélectionner celles qui paraissent essentielles.

La tâche d'informatisation du TRAVAIL DE RÉDACTION du DEC est aussi un problème extrêmement difficile à résoudre au niveau d'un DEC pur et dur, et nous proposons d'examiner, comme solution intermédiaire et provisoire, le travail fait sur l'informatisation de la rédaction du DiCo (voir 3). Il s'agit essentiellement de donner un aperçu des avantages que peut présenter l'utilisation d'un *éditeur de dictionnaire* : c'est-à-dire un programme informatique dédié à la rédaction des articles de dictionnaire. Le programme particulier que nous utilisons pour la rédaction du DiCo est appelé *ÉDiCo*; nous ferons désormais référence à l'*ÉDiCo* quand il s'agira de donner des exemples concrets d'informatisation du travail de rédaction de dictionnaire.

Comme on le voit, notre traitement du problème de l'informatisation du DEC repose sur l'étude d'un dictionnaire simplifié de type DiCo. C'est un peu dommage, dans la mesure où le DiCo est un dictionnaire présentant moins d'intérêt théorique que le DEC. D'un autre côté, dès que l'on parle d'informatisa-

tion, on est forcé de prendre en considération des impératifs matériels qui sont parfois en conflit avec nos visées purement théoriques. Le rôle du présent chapitre est précisément de mettre en évidence ces conflits potentiels et de montrer comment on peut faire un premier pas vers la résolution du problème ardu d'informatisation du DEC en simplifiant ce problème. Toutefois, tout en réduisant les formats et le contenu du DEC, le DiCo en reprend les caractéristiques essentielles. Un article de DiCo est ainsi un premier pas vers un article de DEC et, en conséquence, l'informatisation d'un DiCo est un premier pas vers une informatisation du DEC. Cette façon de procéder – travailler sur une simplification d'un problème très difficilement soluble afin d'obtenir la solution de ce problème lui-même dans un second temps, c'est-à-dire *PROGRESSER PAR SIMPLIFICATION* – est une démarche tout à fait courante en sciences.

De plus, l'informatisation du DEC complet, si elle pose de réels problèmes, n'est ni une tâche utopique, ni quelque chose de non naturel. LE DEC, DANS SA CONCEPTION, EST FAIT POUR ÊTRE INFORMATISÉ. En ce sens, le DEC n'est pas différent de toute description relevant de la linguistique formelle. Le seul intérêt véritable de formaliser une description linguistique est de se donner les moyens de « faire tourner » un appareillage logico-déductif sur cette description afin (i) de vérifier sa validité, et (ii) d'en déduire plus de connaissances sur la langue. Un dictionnaire est un monstre, au niveau de la somme d'information qu'il contient, et il est évident que, pour obtenir le plein rendement d'un dictionnaire formel comme le DEC, le recours à l'ordinateur est essentiel. Informatiser le DEC n'est donc pas une bonne idée qui a germé soudainement dans notre cerveau en observant ce que l'informatique permet aujourd'hui de faire – notamment au niveau du traitement des données non numériques. C'était un des buts ultimes pris en compte il y a plus de vingt-cinq ans lors de l'élaboration de la théorie lexicologique explicative et combinatoire elle-même. En ce sens, un DEC est prédestiné à être informatisé : en fin de compte, il ne peut y avoir de DEC véritable et complet que si ce dictionnaire est informatisé !

Ce que nous venons de dire correspond en quelque sorte aux raisons théoriques qui nous poussent à donner la priorité à l'informatisation du DiCo plutôt qu'à celle du DEC. À cela, nous pouvons ajouter une raison pratique. La tâche théorique de rédaction d'un DEC présente des implications pratiques très importantes, si l'on considère le domaine du Traitement Automatique du Langage Naturel [= TALN]. Un DiCo, du fait de son extrême formalisation, peut être directement utilisé comme dictionnaire formel par les systèmes de TALN. (C'est d'ailleurs un point que nous examinerons en 4, p. 236 *ssq.*) Considérer le DiCo comme un outil potentiel pour le TALN apporte une justification additionnelle au fait de mettre l'emphase sur les phénomènes de cooccurrence lexicale. C'est en effet ce qui est généralement le moins bien décrit dans les dictionnaires courants, tout en étant ce qui est le plus directement exploitable en TALN – dans l'état actuel des connaissances et de la maîtrise technique dans ce domaine.

Cette brève présentation du problème de l'informatisation du DEC et du DiCo montre qu'il ne s'agit pas là d'une simple mise sur support informatique. Ainsi, pour nous, un DEC informatisé n'est pas seulement un DEC qui est disponible sur CD-ROM au lieu d'être imprimé. C'est un DEC sur support informatique à partir duquel des programmes peuvent faire tous les calculs logiques rendus possibles par la théorisation Sens-Texte du lexique. En effet, si l'on se contente de rendre le DEC accessible comme un simple fichier informatique, les seules opérations qui peuvent être effectuées sont, lors de la consultation, des opérations de recherche simple. On peut ainsi, comme dans le cas d'un traitement de texte, rechercher toutes les occurrences d'une forme donnée dans le DEC : par exemple, tous les endroits où la chaîne de caractères *professeur*<sup>2</sup> apparaît. Un DEC véritablement informatisé, au sens où nous l'entendons, permet d'effectuer des opérations autrement plus sophistiquées. En voici quelques illustrations.

1. Recherche « intelligente » dans le DEC :

- Quelles lexies sont des prédicats sémantiques trivalents impliquant deux actants de type 'humain' et un actant du type 'acte' ?
- Quelles lexies verbales réalisent de façon obligatoire leur second actant syntaxique profond sous la forme d'une expression nominale introduite par la préposition À ?
- Quelles sont toutes les lexies qui ont *lourdement* comme un élément de la valeur de la fonction lexicale **Magn** ?
- etc.

2. Mise à jour et vérification automatique :

- Décélérer une mauvaise correspondance entre le nombre d'actants sémantiques introduits dans la définition et les réalisations syntaxiques de ces actants présentées dans le tableau de régime.
- Si le lexicologue introduit ÉCOULERII comme valeur de **Syn**<sub>u</sub> (= plus riche) de VENDREI.1, mettre automatiquement VENDREI.1 comme valeur de **Syn**<sub>c</sub> (= moins riche) de ÉCOULERII.
- etc.

3. Déduction logique :

- Y a-t-il une corrélation entre la présence de *lourdement* comme élément de la valeur de **Magn** d'une lexie et la présence d'une composante sémantique particulière dans la définition de cette lexie ?
- Quelles similitudes retrouve-t-on dans le régime de tous les verbes dont la définition contient la composante 'transaction' ?
- etc.

---

2 Dans ce chapitre, nous utilisons la fonte «courier» pour indiquer toute chaîne de caractères stockée sur support informatique. Cela peut être un article de DiCo, une instruction donnée en entrée à un programme informatique, la sortie de l'exécution d'un tel programme, etc.

Nous ne donnons ci-dessus que quelques exemples; mais l'on comprend aisément tout ce qu'une véritable informatisation du DEC peut offrir, tant au niveau du contrôle du processus de rédaction du DEC, qu'au niveau de son utilisation en tant que base de données lexicales.

Le point 1 ci-dessus concerne l'accès à l'information au sens large. Pour ce qui concerne la consultation du dictionnaire proprement dite, un DEC informatisé est bien entendu beaucoup plus maniable que sa version imprimée. Ainsi, le problème de l'ordonnement linéaire de l'information, mentionné au Chapitre IV (début de 1.3), disparaît dans le cas du DEC informatisé. En effet, un tel dictionnaire permet l'accès instantané à n'importe quelle donnée lexicologique; il permet aussi l'affichage simultané à l'écran de plusieurs types d'information distincts (au moyen du multi-fenêtrage). Cela optimise la consultation du dictionnaire, et par là même sa rédaction en permettant de mieux prendre en compte la cohérence « verticale » (dans le vocable) et « horizontale » (dans le champ sémantique) – puisque plusieurs articles d'un même vocable, d'un même champ sémantique, etc. peuvent être consultés en parallèle. De plus, si le DEC est encodé sous forme d'*hypertexte*<sup>3</sup>, comme le serait n'importe quel bon dictionnaire un peu formalisé, l'accès aux articles peut se faire instantanément à partir de n'importe quelle zone d'un article donné. Par exemple, en cliquant sur une composante simple (un sens lexical) d'une définition, ou sur le nom d'un élément d'une valeur de fonction lexicale, on fait automatiquement ouvrir l'article de DEC correspondant. Dans le cas des définitions, on voit bien que seul le DEC permet aisément ce genre d'encodage hypertexte puisqu'il impose au lexicologue de définir les lexies avec des composantes clairement identifiées comme sens d'autres lexies. Chaque composante d'une définition de DEC est donc intrinsèquement un pointeur vers un article de dictionnaire lui correspondant. Cette caractéristique du DEC rend son encodage hypertexte particulièrement justifié. Bien entendu, tout cela relève de l'« habillage » du DEC informatisé, et ne doit pas nous concerner plus ici.

## 2 Description détaillée du Dictionnaire de cooccurrences [= DiCo]

Dans cette section, nous présentons en détail les formalismes utilisés dans le DiCo (2.1). Nous donnons ensuite (2.2) deux superarticles de DiCo : la traduction en format DiCo du superarticle de DEC pour VENDRE (qui a été présenté au Chapitre V) et le superarticle « original » de DiCo pour le vocable ADHÉSION.

3 Rappelons qu'un *hypertexte* est un texte sur support informatique encodé de telle façon qu'il contient des branchements directs entre des éléments du texte (mots, phrases, graphiques, etc.) et d'autres textes : il suffit d'effectuer une opération donnée sur un de ces éléments – cliquer, sélectionner, etc. – pour faire automatiquement afficher un texte qui y est associé (ou faire exécuter une procédure quelconque). Grâce à l'hypertexte, on peut ainsi naviguer beaucoup plus aisément à travers une collection de textes entre lesquels des connexions (des renvois, en quelque sorte) ont été établies au préalable.

## 2.1 Formalismes du DiCo

Nous commençons par la caractérisation de la structure canonique d'un superarticle de DiCo, pour passer ensuite aux composantes individuelles de cette structure.

### 2.1.1 Structure générale d'un fichier de DiCo décrivant un vocable = un superarticle de DiCo

Chaque superarticle de DiCo est stocké dans son propre *fichier texte* (anglais *text file*). Rappelons qu'un fichier texte est un fichier informatique ne contenant qu'une seule fonte et qui, de façon plus générale, ne contient aucun format d'édition spécial. En conséquence, les fichiers du DiCo ne contiennent aucun caractère formaté en gras, italique, indice, exposant, etc. Toute convention d'écriture du DEC mettant en jeu ce type de format doit se voir attribuer une convention équivalente qui puisse être encodée dans un fichier texte. Nous procédons ainsi afin d'assurer une plus grande « portabilité » de la base de données que constitue un DiCo : en principe, tout fichier créé pour le DiCo peut être LU, ANALYSÉ ET INTERPRÉTÉ par n'importe quel programme informatique sans qu'il faille jongler avec des problèmes de transcription de données<sup>4</sup>.

L'information décrivant un superarticle, c'est-à-dire un vocable, est structurée dans un fichier de DiCo de la façon suivante :

```

<Nombre entier> <- Statut
    « < Superarticle> »
    <Remarques diverses>
```

*Figure VI-1 Structure d'un fichier du DiCo*

Pour décrire formellement la syntaxe des articles de DiCo, nous utilisons les conventions d'écriture suivantes :

- Tout ce qui est indiqué entre chevrons (<...>) est, formellement, une variable. C'est-à-dire que c'est une « métacomposante » de la description, qui va être décrite plus en détail (expliquée et/ou formellement décomposée) plus loin.
- Tout ce qui n'est pas indiqué entre chevrons est une chaîne de caractères apparaissant comme telle dans l'article de DiCo.

Comme le montre la Fig. VI-1 ci-dessus, la première information que l'on trouve dans un fichier de DiCo est le « statut » du superarticle qu'il contient.

<sup>4</sup> Il serait bien sûr possible d'envisager une plus grande standardisation de l'encodage de l'information dans le DiCo, en faisant notamment usage des standards de type SGML [= *Standard Generalized Markup Language*]. Cet aspect très technique du problème de l'informatisation du DEC et du DiCo (de même que ce qui a été dit plus haut à propos de l'encodage de type hypertexte) est d'une grande importance pratique; il déborde néanmoins largement le cadre du présent chapitre.

<Nombre entier>, qui apparaît à gauche de l'expression <- Statut, est un entier indiquant l'état d'avancement de la rédaction du superarticle. Nous distinguons quatre statuts possibles, qui reflètent les quatre étapes successives de rédaction d'un superarticle de DiCo – du statut 3 pour arriver au statut final 0.

- Statut 3      Le superarticle a été créé mais ne possède pas nécessairement la structure canonique d'un superarticle de DiCo ; en fait, il peut tout à fait ne contenir aucune information, à part le statut, le nom d'un vocable et un squelette de superarticle. (C'est ce qui arrive quand on commence la rédaction d'une toute nouvelle description – voir un peu plus bas dans cette section, le squelette de superarticle créé par l'ÉDiCo pour le vocable ADHÉSION, p. 213.)
- Statut 2      Le superarticle a été ébauché; il contient au moins la structure canonique imposée par le DiCo – vérifiée au moyen du programme d'analyse des superarticles de l'ÉDiCo (voir 3).
- Statut 1      Le superarticle peut être considéré comme étant valide d'un point de vue formel puisqu'il suit la syntaxe du DiCo, et d'un point de vue linguistique puisqu'il contient toute l'information jugée nécessaire et suffisante pour décrire la cooccurrence des lexèmes du vocable. Pour passer du statut 2 au statut 1, il faut s'assurer que toutes les données présentes ont été vérifiées – avec des informateurs, des corpus informatisés, etc. Il faut aussi s'assurer qu'aucune lexie du vocable traité n'a été oubliée et qu'aucun cas important de cooccurrence lexicale restreinte ne manque dans les articles.
- Statut 0      Un superarticle de statut 0 peut être considéré comme terminé. Il peut être publié ou distribué comme tel. Avant d'attribuer le statut 0 à un superarticle, le lexicologue doit s'assurer que :
- le superarticle est consistant au niveau de la macrostructure du DiCo; c'est-à-dire que la description du vocable forme un tout homogène avec la description des autres vocables appartenant au même champ sémantique (suivant en cela, pour la rédaction du DiCo, le même principe de rédaction que pour le DEC – énoncé au Ch. IV, 2.4, p. 178);
  - les correspondances interarticles sont correctes; notamment, les renvois aux lexèmes déjà décrits sont faits en fonction de la bonne numérotation – par exemple, dans l'article pour ADHÉSIONI donné plus loin, la valeur adhérerI de la fonction lexicale {V0}<sup>5</sup> renvoie effectivement au bon lexème verbal correspondant décrit dans le superarticle de ADHÉRER;

5 Nous utilisons ici les conventions d'écriture du DiCo, qui vont être présentées en détail en 2.1.2, p. 214 *ssq.*

- les correspondances logiques entre les superarticles sont établies – par exemple, si un lexème  $L_1$  est décrit comme un  $\text{Conv}_{21}$  d'un lexème  $L_2$ ,  $L_2$  doit lui aussi être décrit comme étant un  $\text{Conv}_{21}$  de  $L_1$ .

Le superarticle comme tel – voir la variable <Superarticle> dans la Fig. VI-1 ci-dessus – apparaît immédiatement après la mention du statut, encadré par les symboles " « " et " » ". Il possède la structure formelle suivante :

```

<Nom du vocable>
Δ <Partie du discours>
• <Numéro du premier lexème>
  <Description du premier lexème>
...
• <Numéro du dernier lexème>
  <Description du dernier lexème>
    
```

*Figure VI-2 Structure d'un superarticle du DiCo*

Pour aider à la lecture de ce qui suit, rappelons que les articles de DiCo des vocables VENDRE et de ADHÉSION sont donnés plus bas, en 2.2.

Lorsqu'il utilise l'éditeur spécialisé ÉDiCo, le lexicologue n'a pas besoin d'introduire à la main les différents symboles qui structurent la description d'un vocable. Lors de la création d'un nouveau superarticle, le programme crée automatiquement son squelette (contenant la structure vide d'un premier article de lexème) : il ne reste plus au lexicologue qu'à « combler les vides ». On retrouve ici la même philosophie de traitement uniforme, que nous avons comparée aux « exercices à trous » dans le Chapitre II (1.3, p. 42). Ainsi, si le lexicologue demande à l'ÉDiCo de créer un nouveau superarticle pour le vocable ADHÉSION, celui-ci affichera automatiquement la structure de données suivante

```

3 <- Statut
« ADHÉSION
Δ <partie du discours>
• <lexie>
◊cg <carac. gram. & marques d'usage>
◊cs <caractéristiques sémantiques>
◊nb <N.B.>
◊tr <tableau de régime>
◊fl <fonctions lexicales>
◊ex <exemples>
◊ph <phraséologie> »
    
```

*Figure VI-3 Squelette de superarticle créé par l'ÉDiCo*

Comme on peut le voir dans la Fig. VI-3, chaque article de lexie contient exactement sept zones – cg, cs, nb, tr, fl, ex et ph – que nous allons décrire plus bas. Le début de chaque zone est indiqué par le symbole «  $\diamond$  », qui, avec le nom de la zone qu'il accompagne, informe l'ÉDiCo de la nature de l'information introduite.

Nous donnons ci-dessous les symboles qui, dans le DiCo, sont réservés à la structuration des superarticles (et articles) :

- « début d'un superarticle
- » fin d'un superarticle
- $\Delta$  introduction de la partie du discours pour un vocable
- début d'un article de lexie
- $\diamond$  début d'une zone à l'intérieur d'un article de lexie

Les symboles ci-dessus ne doivent JAMAIS apparaître dans un article de DiCo avec une autre signification que celle qui leur est attribuée ici.

Avant de passer à l'étude détaillée des différentes composantes d'un superarticle de DiCo, il faut noter que l'on peut écrire n'importe quoi – commentaires, données et remarques en attente de traitement, etc. – après le symbole " » ", qui termine le superarticle. Toute information apparaissant après ce symbole ne sera pas prise en compte par les différents utilitaires (au sens de 'logiciels de gestion des données') contenus dans l'éditeur de DiCo; c'est ainsi qu'il faut interpréter la variable <Remarques diverses>, apparaissant au bas de la Fig. VI-1.

### 2.1.2 Analyse détaillée des composantes d'un superarticle de DiCo

Nous passons maintenant en revue chacune des composantes d'un superarticle de DiCo, telles qu'elles apparaissent entre les symboles " « ... » " dans la Fig. VI-3 ci-dessus.

#### Nom de vocable – « ... :

Comme dans le DEC, les noms de vocables sont écrits en lettres majuscules. Ils ne sont normalement pas numérotés, à moins qu'il faille distinguer deux formes homographiques. Les fichiers informatiques qui constituent le DiCo étant en format texte (une seule fonte, « *courier* », et aucun format spécial du type italique, soulignement, indice, etc.), il a été nécessaire d'introduire certaines conventions d'écriture qui se démarquent des conventions normales du DEC. Elles sont ici de deux types :

- vocables homographiques

<b>DEC</b>	(un) MOULE <sup>1</sup>	vs	(une) MOULE <sup>2</sup>
<b>DiCo</b>	MOULE (1)	vs	MOULE (2)



- phrasèmes

**DEC**    [DANS LE BAIN]<sup>1</sup>

**DiCo**    ,DANS LE BAIN, <sup>6</sup>

**Partie du discours – Δ . . . :**

Comme dans le DEC, la partie du discours d'un vocable donné doit être indiquée après le nom du vocable, si toutes les lexies de ce vocable appartiennent à la même partie du discours (ce qui est le cas standard). Dans le DiCo, cette information est entrée dans une composante du superarticle spécialement prévue à cet effet – elle est identifiée par le symbole « Δ ». Cette composante reste vide si toutes les lexies n'appartiennent pas à la même partie du discours; la partie du discours de chaque lexie est alors indiquée dans la zone appropriée de son article (voir plus bas). Il existe bien entendu une liste de parties du discours à utiliser dans le DiCo, qui sont reconnues par l'éditeur (nom, verbe, adjectif, adverbe, préposition, conjonction, etc.).

Nous entrons maintenant dans la description des articles de DiCo correspondant à chaque lexie du vocable. Bien entendu, l'information décrite ci-dessous est répétée, dans un superarticle de DiCo, pour chaque lexie du vocable en question.

**Numéro de lexie – • . . . :**

Les lexies sont identifiées simplement par leur numéro lexical. Il est inutile d'entrer ici le nom complet de la lexie puisqu'il peut toujours être construit automatiquement par l'ÉDiCo, si besoin est, à partir du numéro de la lexie et du nom du vocable (donné au tout début du superarticle). Le DiCo utilise le même principe de numérotation des lexies que le DEC (voir Ch. IV, 1.4, p. 166) : les distinctions sémantiques majeures sont indiquées par des chiffres romains; les lexies plus proches sémantiquement sont distinguées par des chiffres arabes; finalement, les distinctions sémantiques plus fines sont indiquées par des lettres minuscules. Par exemple, les deux articles de DiCo donnés en 2.2 ci-dessous possèdent les structures en lexèmes suivantes :

VENDRE	•	I . 1	ADHÉSION	•	I
	•	I . 2		•	II . 1
	•	II		•	II . 2a
	•	III		•	II . 2b

Cela correspond très exactement à la structuration de ces vocables en lexèmes à l'intérieur du DEC.

6 Les symboles « , » et « . » sont des caractères standard dans l'environnement informatique pour lequel a été développé l'éditeur de DiCo (c'est-à-dire le micro-ordinateur Macintosh™). Il est clair que le choix de ces symboles, comme de toutes les autres conventions d'écriture du DiCo, est plus ou moins arbitraire. Peu importe quels symboles sont choisis, pourvu qu'ils soient réservés à une fin donnée et ne soient jamais utilisés dans un autre but dans les articles de DiCo.

**Caractéristiques grammaticales et marques d'usage –  $\diamond_{cg}$  . . . :**

Cette zone contient toute l'information concernant le comportement de la lexie qui n'est décrit ni dans la zone de combinatoire syntaxique (tableau de régime), ni dans la zone de combinatoire lexicale (fonctions lexicales). Il n'existe pas dans les articles de DEC de zone correspondant exactement à cette zone d'article de DiCo. (Nous reviendrons sur ce point particulier un peu plus bas.) On trouve dans la zone «  $\diamond_{cg}$  » deux types d'information.

Premièrement, cette zone présente les caractéristiques grammaticales de la lexie, c'est-à-dire (i) les contraintes sur les types de constructions syntaxiques dans lesquelles cette lexie peut apparaître comme dépendant d'une autre lexie et (ii) toute autre contrainte de type morphologique sur l'usage de la lexie : par exemple, si la lexie est un nom toujours employé au pluriel, la zone en question contiendra la mention pl.

Deuxièmement, cette zone contient aussi les marques d'usage telles que "litt" (pour les termes littéraires) ou "spéc" (pour les lexies relevant du langage spécialisé, par exemple technique). À noter que les marques d'usage sont toujours entrées entre guillemets, afin de les distinguer des caractéristiques grammaticales proprement dites.

Cette zone peut contenir plusieurs caractéristiques grammaticales et marques d'usage, séparées par des virgules. (On donne les caractéristiques grammaticales en premier, suivies des marques d'usage.) Par exemple, la formule suivante, apparaissant dans l'article de DiCo de AIMERIII, indique que ce lexème ne s'emploie pas au passif et est considéré comme littéraire. Nous donnons la zone d'exemples de l'article afin que le lecteur puisse se faire une idée de la lexie dont il est question ici :

$\diamond_{cg}$  pas de pass, "litt"

(...)

$\diamond_{ex}$  Ils s'aimèrent toute la nuit. Elle l'aima cette nuit-là de toute la force de ses vingt ans.

Toute information présentée dans cette zone peut être factorisée et « montée » au niveau du vocable, si toutes les lexies du vocable décrit ont la même caractérisation. Ainsi, si toutes les lexies d'un vocable nominal sont employées au pluriel, il n'est pas nécessaire de répéter cette information dans chaque zone «  $\diamond_{cg}$  » : elle est entrée une seule fois dans la composante «  $\Delta$  » du superarticle. Au contraire, quand la caractéristique n'appartient pas à toutes les lexies, il convient de l'entrer dans la zone «  $\diamond_{cg}$  » de chaque lexie concernée. On trouve une illustration d'un tel cas de figure dans le superarticle de ADHÉSION, donné en 2.2. On y voit que seuls les deux premiers lexèmes de ce vocable – ADHÉSIONI (l'adhésion de quelque chose à une surface) et ADHÉSIONII.1 (l'adhésion de quelqu'un à une idée) – ne connaissent pas d'emploi pluriel. Comme cette caractéristique ne s'applique pas aux deux autres lexèmes – ADHÉSIONII.2a

(l'adhésion de quelqu'un à une organisation) et ADHÉSIONII.2b (la carte d'adhésion) – le code pas de pl est entré dans la zone «  $\diamond_{CG}$  » des deux premiers lexèmes, puisqu'il ne peut être factorisé et monté au niveau de la composante «  $\Delta$  » du superarticle.

**NB :** Le lecteur attentif aura peut-être constaté que la composante «  $\Delta$  » d'un superarticle de DiCo et les zones «  $\diamond_{CG}$  » de chaque article contiennent potentiellement les mêmes types d'information : caractéristiques grammaticales et marques d'usage. Si nous identifions la composante «  $\Delta$  » comme le lieu où est entrée la partie du discours, c'est parce que cette caractéristique grammaticale particulière est dans la plupart des cas celle qui se trouve factorisée et montée au niveau du vocable.

DANS LE DEC, comme nous l'avons mentionné plus haut, toutes les informations qui viennent d'être décrites ne sont pas entrées dans une zone spéciale de l'article de dictionnaire. Elles apparaissent dans la zone sémantique, entre le numéro lexical et la définition proprement dite (lorsqu'elles ne sont pas factorisées au niveau du vocable). Ainsi, dans le superarticle de DEC pour VENDREL2, donné au Chapitre V (4, p. 198), on trouve :

I.2. pas de passif en *se. spéc.* X vend Y à Z pour W = Autorité juridique X (...).

On peut voir comment cette information est encodée dans le DiCo en comparant le superarticle de DEC avec celui de DiCo correspondant, donné en 2.2. Il est à noter que le DiCo offre une meilleure formalisation que le DEC, au niveau de la répartition de l'information dans des zones spécialisées des articles. Dans le cas précis que nous venons d'examiner, la façon d'entrer les caractéristiques grammaticales et marques d'usage dans le DEC a été héritée de la pratique courante des dictionnaires commerciaux. En principe, une zone spéciale pour ce type d'information devrait aussi être utilisée dans le DEC. Rappelons que ce problème a déjà été évoqué au Chapitre III (3.1, pp. 70-71). Il est normal que le travail sur le DiCo, et plus généralement sur l'informatisation du DEC, mette en évidence certaines lacunes formelles du DEC<sup>7</sup>.

### Caractéristiques sémantiques – $\diamond_{CS}$ . . . :

Le DiCo, et c'est là une de ses différences majeures avec un DEC, ne contient pas de véritables définitions lexicales. Il donne à la place une spécification approximative du sens général de chaque lexie ainsi que de sa structure actancielle. Les formules caractérisant sémantiquement les lexies dans le DiCo possèdent la structure donnée dans la Fig. VI-4 ci-dessous ; notons cependant que la zone

7 La description du vocable VENDEUR, toujours au Chapitre V, est un autre exemple de ce type de phénomène. Toutes les lexies de ce vocable n'appartenant pas à la même classe syntaxique, les caractéristiques grammaticales ne peuvent être factorisées et montées au niveau du vocable. Elles sont spécifiées pour chaque lexie : « nom, masc [fém -euse] » pour VENDEUR<sub>Ia</sub>, <sub>Ib</sub> et <sub>II</sub> ; « adjectif [fém -euse] » pour VENDEUR<sub>III</sub>.

(sémantique) «  $\diamond_{CS}$  » d'une lexie qui n'est pas un prédicat sémantique (c'est-à-dire une lexie qui est un nom d'objet sémantique) ne contient que la partie gauche – <Étiquette sémantique générale> – de cette structure.

<Étiquette sémantique générale>: <Forme propositionnelle>

*Figure VI-4 Caractérisation sémantique d'une lexie dans le DiCo*

Comme son nom le laisse entendre, l'<Étiquette sémantique générale> est une simple indication du sens général de la lexie. En quelque sorte, elle indique un champ sémantique auquel peut être identifiée la lexie. Nous n'utilisons pas ici le terme de *trait sémantique*, car il peut prêter à confusion en laissant penser à tort que la formule de la zone «  $\diamond_{CS}$  » présente une sorte d'analyse componentielle – du type [ $\pm$  humain], [ $\pm$  mâle], etc. Nous préférons utiliser le terme d'*étiquette sémantique* : il reflète assez précisément le rôle joué par cette partie de la description.

La <Forme propositionnelle> d'une zone sémantique du DiCo est très proche d'un défini d'une définition du DEC – les variables X, Y, Z, etc. sont utilisées pour indiquer les actants sémantiques et des mots grammaticaux (prépositions, notamment) sont utilisés pour aider à la lecture de la formule. Néanmoins, la forme propositionnelle du DiCo possède la propriété distinctive suivante : les actants y sont, dans la mesure du possible, caractérisés par des étiquettes sémantiques.

Ainsi, les formules suivantes en gras sont les caractérisations sémantiques des différents lexèmes du vocable ADHÉSION, telles qu'elles apparaissent dans le superarticle de DiCo donné en 2.2 ci-dessous (la lexie vedette est remplacée dans la formule par le symbole « ~ » et chaque caractérisation est accompagnée des exemples utilisés dans l'article correspondant) :

ADHÉSIONI

**événement physique: ~ de substance X à substance Y**  
 Cette surface lisse permet une très bonne adhésion.

ADHÉSIONII.1

**action mentale: ~ de personne X à idée Y**  
 Sa présentation a emporté l'adhésion spontanée de tous ses collègues.

ADHÉSIONII.2a

**action sociale: ~ de personne X à organisation Y**  
 L'adhésion du Portugal au Marché Commun...  
 Avez-vous reçu ma demande d'adhésion ?  
 En effet, l'adhésion au club ne revient qu'à 25 F.

ADHÉSIONII.2b

**document: ~ de personne X [à organisation]**  
 Je ne parviens pas à retrouver votre adhésion.

Le cas de la dernière lexie – ADHÉSIONII.2b – est une bonne illustration de l'intérêt que peut présenter l'étiquetage sémantique pour l'utilisateur du DiCo (que ce soit un utilisateur humain ou une machine). En effet, même en l'absence d'une définition véritable, on peut assez bien mesurer l'écart sémantique existant entre la lexie de base de ADHÉSION et cette dernière lexie. Le DiCo montre explicitement que la première lexie est un prédicat sémantique à deux actants, caractérisé comme événement physique; la dernière lexie est un quasi-prédicat sémantique à un actant, caractérisé comme document. Nous appelons ADHÉSIONII.2b un quasi-prédicat car, bien qu'il doive être décrit sémantiquement comme mettant en jeu un actant (l'adhésionII.2b DE QUELQU'UN), il est en même temps très proche d'un nom d'objet puisqu'il dénote un document physique (un morceau de papier, en quelque sorte). D'une certaine façon, ADHÉSIONII.2b est sémantiquement semblable aux noms de parties du corps (BRASI.1a, JAMBELI.1a, etc.), dont il a déjà été question – notamment au Ch. III, 3.3.1.2, p. 77. La numérotation de la lexie en question (II.2b) est elle aussi consistante avec l'approche du DEC. Même si une carte d'adhésion semble être une entité très différente des entités abstraites dénotées par les autres lexèmes de ADHÉSION, la distance sémantique entre ce lexème et ADHÉSIONII.2a est petite, si l'on prend en compte les critères proposés au Chap. IV, 1.4, p. 166 :

- ADHÉSIONII.2a et ADHÉSIONII.2b ont une partie sémantique commune importante, puisqu'une adhésionII.2b est en fait (comme il est indiqué dans la zone de fonctions lexicales de l'article de ce lexème) une *carte d'adhésionII.2a*, c'est-à-dire un document certifiant l'adhésionII.2a de quelqu'un à une organisation;
- la distinction entre ces deux lexèmes est régulière en français (par exemple, *Il m'a donné l'autorisation de...* [un fait]/*Montrez-moi votre autorisation* [un document], etc.).

Toutes les observations que nous venons de faire montrent bien que, même si le DiCo ne contient pas de véritables définitions lexicales, l'approche sémantique du lexique qu'il présuppose est bien la même que celle du DEC. La détermination des étiquettes sémantiques dont il est fait usage pose cependant de nombreux problèmes théoriques et pratiques. Il est nécessaire de donner un statut théorique précis au concept d'étiquette sémantique; il est aussi nécessaire d'utiliser une méthodologie claire pour construire la liste des étiquettes apparaissant dans le DiCo d'une langue donnée. Chaque étiquette doit être justifiée théoriquement, et sa valeur (au sens saussurien de *valeur d'un signe*) relativement aux autres étiquettes disponibles doit être clairement spécifiée. Ainsi, dans la hiérarchie provisoire d'étiquettes sémantiques que nous utilisons, personne peut renvoyer à la fois à des êtres humains et à des entités collectives comme des pays, des institutions, des syndicats, etc. L'étiquette plus spécifique que nous utilisons pour renvoyer à un être humain uniquement est individu; elle s'oppose aux trois étiquettes suivantes : communauté, établissement et organisation.

L'utilisation d'étiquettes sémantiques est un aspect du travail sur le DiCo qui en est encore au stade de la recherche, et nous nous excusons ici pour le caractère approximatif de notre présentation. Nous ne sommes ainsi pas en mesure pour l'instant de répondre de façon précise et définitive aux questions suivantes : Quels types d'étiquettes sémantiques faut-il utiliser ? Comment peuvent-elles être isolées ? Dans quelle mesure est-il possible de hiérarchiser ces étiquettes ? Correspondent-elles exactement à des lexies de la langue décrite, identifiant des champs sémantiques, ou devons-nous les considérer comme des outils techniques abstraits ? etc. Nous espérons que c'est par le biais du travail de construction d'un DiCo que nous pourrions étayer peu à peu la base théorique sur laquelle repose l'utilisation d'étiquettes sémantiques. Pour l'instant, comme hypothèse de travail, nous postulons que l'ensemble des étiquettes utilisées peut être strictement hiérarchisé, pourvu que l'on permette à une étiquette donnée de posséder dans la hiérarchie plus d'une étiquette « mère ». C'est en fonction de ce principe que nous avons commencé à développer graduellement une hiérarchie d'étiquettes sémantiques pour la description du lexique français. La Fig. VI-5 ci-dessous donne en exemple une portion de cette hiérarchie – celle dépendant de l'étiquette entité physique; on y voit notamment que l'étiquette individu dépend à la fois de l'étiquette être animé et de l'étiquette personne.

**NB :** Nous avons bien entendu des raisons précises pour structurer notre hiérarchie de cette façon – raisons basées sur des oppositions de lexies à l'intérieur d'un même vocable ou à l'intérieur d'un champ sémantique donné. Cet aspect du travail sur le DiCo dépasse cependant largement le cadre de l'ILEC.

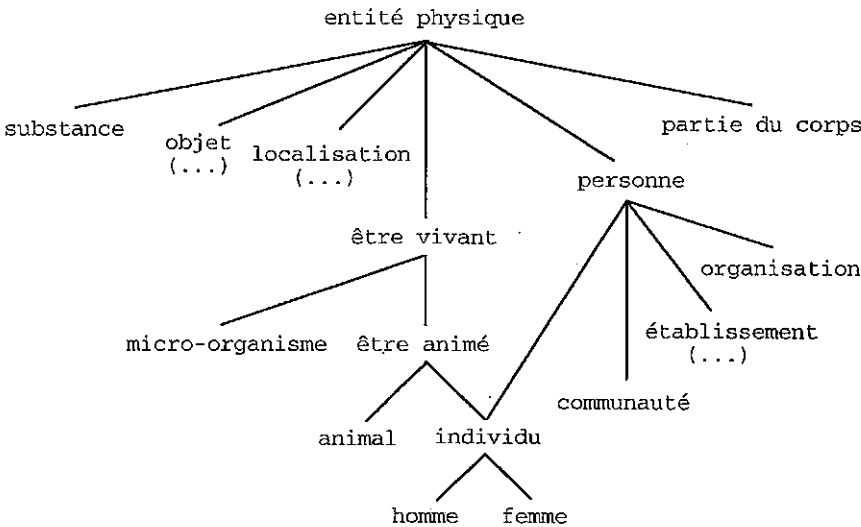


Figure VI-5 Fragment de la hiérarchie d'étiquettes sémantiques du DiCo

L'ÉDiCo contient une batterie d'outils grâce auxquels le lexicologue peut construire la hiérarchie d'étiquettes sémantiques au fur et à mesure que s'effectue la rédaction des articles de DiCo. Ces outils permettent (i) d'accéder à la hiérarchie provisoire déjà construite, et (ii) d'augmenter, de modifier et de raffiner cette hiérarchie (pour plus de détails, voir plus bas, 3.2.1, p. 230 *ssq.*).

**Nota bene** –  $\diamond_{nb}$  . . . :

Les articles de DiCo contiennent une zone de *nota bene* utilisée pour stocker des informations à propos de la lexie qui ne sont pas encore strictement formalisées : elles sont avant tout destinées au lecteur/rédacteur humain du DiCo, et non à des programmes informatiques. On peut trouver ici des restrictions sur le type de contextes pragmatiques dans lesquels la lexie peut être utilisée; le *nota bene* est aussi utilisé pour stocker des informations sur la connotation de la lexie. (Bien que les articles de DEC peuvent aussi contenir un *nota bene* NB, rappelons que la connotation est entrée dans une partie spéciale de la zone sémantique des articles de DEC – voir Ch. III, 3.3, p. 72.)

**Tableau de régime** –  $\diamond_{tr}$  . . . :

Cette zone contient la transcription linéaire d'un tableau de régime standard; elle est néanmoins potentiellement moins précise et complète que sa contrepartie du DEC. Dans l'état actuel de formalisation du DiCo, il est en effet impossible d'encoder toutes les contraintes sur la réalisation des actants syntaxiques profonds. Pour mieux comprendre le problème, on peut comparer ci-dessous le tableau de régime type DEC de VENDREI.1 (abrégé) et sa contrepartie en format DiCo :

Format DEC				Format DiCo	
X = I	Y = II	Z = III	W = IV	$\diamond_{tr}$ X = I = N	
1. N	1. N	1. à N	1. Num N	Y = II = N (oblig)	
			2. à Num N	Z = III = à N	
	obligatoire		3. pour N	W = IV = Num N, à Num N, pour N	

1) C<sub>II</sub> : si (... les droits de possession sexuelle temporaire de la personne Y),  
**alors** N = CORPS<sup>1</sup>II.1a, CHARMES,  
 FAVEURS, CARESSES, CUL, ...

(...)

On voit clairement que les contraintes sur la réalisation des actants de VENDREI.1, qui établissent une correspondance entre l'activation d'une composante sémantique de la définition et la présence de certains lexèmes réalisant l'actant Y, ne sont pas indiquées dans le tableau de régime du DiCo. En l'absence

d'une formalisation complète et rigoureuse, une solution pour stocker ce type d'information dans le DiCo est de l'entrer, de façon non formelle, dans la zone de *nota bene*. Une autre solution est celle adoptée pour les contraintes associées à l'expression de l'actant W dans tableau de régime de VENDREII (voir plus bas, 2.2.1, p. 225) : les contraintes sont directement entrées dans la ligne correspondante du régime.

### Fonctions lexicales – $\diamond f1$ . . . :

Le DiCo fait usage des mêmes fonctions lexicales que le DEC, en utilisant cependant certaines conventions d'écriture particulières. Par exemple, pour faciliter l'analyse automatique des articles de dictionnaire, les noms de fonctions lexicales doivent être entrés entre accolades, suivis par les valeurs elles-mêmes. Ainsi, on trouve les descriptions ci-dessous dans l'article de DiCo de VENDREI.1 :

```

 $\diamond f1$ {QSyn} écoulerII, débiter2; laisserIII.4; céderI.2
      {QConv3214} acheter1
      {S0} ventela
      {S1} vendeurI.a
      (...)
      {S4} prixI
      {AntiMagnS4} "fam" ,une bouchée de pain,,
                  "fam" ,des clopinettes,
      {Sloc} marché1
      {QSloc--usual} ,point de vente,; marché2; commerceI.3;
                  pharmacie2
      (...)

```

On constate ici l'application des conventions d'écriture suivantes :

1. On n'utilise pas dans le DiCo les indices sémantiques  $\underset{\sim}{\nu}$ ,  $\underset{\varepsilon}{\nu}$  et  $\underset{\circ}{\nu}$  (introduits au Ch. III, 3.5.2, pp. 129-130). Au lieu de cela, à chaque fois qu'il faut entrer une relation sémantique approximative, le DiCo utilise le préfixe Q-, pour Quasi-, sans préciser davantage s'il s'agit d'un sens plus riche, moins riche ou d'une intersection de sens. Ainsi, la description du DEC suivante :

Syn  $\underset{\sim}{\nu}$  : écoulerII, débiter2; laisserIII.4

Syn  $\underset{\circ}{\nu}$  : céderI.2

est en fait encodée dans le DiCo par :

```
{QSyn} écoulerII, débiter2; laisserIII.4; céderI.2
```

2. Le DiCo n'utilise ni indice ni exposant, puisque les superarticles de DiCo doivent être écrits sans aucun format d'édition. Pour remplacer ces formats, on utilise le trait d'union « - » pour indiquer un indice et le double trait d'union « -- » pour indiquer un exposant; les exposants apparaissent après les indices. De plus, les numéros et paramètres sémantiques (loc, instr, etc.) qui suivent immédiatement les noms de fonctions lexicales comme S, Oper, etc. sont toujours interprétés comme des indices. Par exemple, les données suivantes de DEC :



S<sub>4</sub> : prixI

S<sub>loc</sub><sup>usual</sup> : [point de vente]; marché2; commerceI.3; pharmacie2

correspondent dans le DiCo à :

{S4} prixI

{QSloc--usual} ,point de vente.; marché2; commerceI.3;  
pharmacie2

**Exemples** –  $\diamond$ ex... :

Il n'y a rien de spécial à dire à propos de cette zone. Comme dans le reste du DiCo, le texte est entré sans formatage; c'est-à-dire, entre autres, que nous n'utilisons pas d'italiques.

**Phraséologie** –  $\diamond$ ph... :

Dans la zone phraséologique d'un article de DiCo, on ne distingue pas les différents regroupements sémantiques des phrasèmes liés à la lexie vedette. Ainsi, la zone phraséologique structurée du DEC pour VENDREI.1, qui regroupe les phrasèmes en fonction de leur lien avec des connotations de ce lexème (voir Ch. V, 4, p. 197), correspond dans le DiCo à la description « plate » suivante :

$\diamond$ ph ,vendre chèrement [N],  
,vendre la peau de l'ours avant de l'avoir tué,  
,vendre la mèche,  
,vendre père et mère,  
,vendre son âme (au diable),  
,vendre son droit d'aînesse pour un plat de lentilles,  
,savoir se vendre,  
,vendre des frigidaires aux Esquimaux,

## 2.2 Exemples de deux superarticles de DiCo

À fin d'illustration, nous donnons maintenant, presque sans commentaires, deux exemples de superarticles de DiCo :

- en 2.2.1, un superarticle qui est la simple traduction en format DiCo de la description de type DEC de VENDRE (donnée au Ch. V, 4, p. 195 *ssq.*);
- en 2.2.2, un superarticle qui a directement été écrit pour le DiCo, pour le vocable nominal ADHÉSION.

### 2.2.1 Traduction en format DiCo d'un superarticle de DEC : VENDRE

2 <- Statut  
<< VENDRE  
Δ verbe

• I.1

◊cg se-pass  
 ◊cs action sociale: personne X ~ Y à personne Z pour argent W  
 [= QConv3214(acheter1)]  
 ◊nb  
 ◊tr X = I = N  
 Y = II = N (oblig)  
 Z = III = à N  
 W = IV = Num N, à Num N, pour N  
 ◊fl {QSyn} écoulerII, débiter2; laisserIII.4; céderI.2  
 {QConv3214} acheter1  
 {S0} ventela  
 {S1} vendeurI.a  
 {QS1} vendeurI.b; marchandI  
 {QS2} marchandisel  
 {S3} acheteur1  
 {S4} prixI  
 {AntiMagnS4} "fam" ,une bouchée de pain.,  
 , "fam" ,des clopinettes,  
 {Sloc} marché1  
 {QSloc--usual} ,point de vente.; marché2; commerceI.3;  
 pharmacie2  
 {QA2} en vente | Y est une marchandise régulière  
 {Able2} vendable  
 {AntiAble2} invendable  
 {PredAntiAble2} ,ne pas trouver acheteur <acquéreur>.  
 {Magn4} cher  
 {AntiMagn4} ,bon marché., à vil prixI //brader1  
 {A2nonPerf} invendu[Adj]  
 {QS2nonPerf} invendu[Nom]  
 {F1 = 'V. Y en grandes quantités  
 à un seul Z'} "spéc" ~ en gros > "spéc" ~ en demi-gros  
 {QS1F1} grossiste  
 {F2 = 'V. séparément les éléments  
 composant un lot de marchandises1'} "spéc" ~ au détail //  
 "spéc" faire le détail  
 {QS1F2} détaillant  
 {'V. à un prixI réduit'} ~ au rabais //solder3, liquider2  
 {'V. Y à un prixI qui est plus bas que  
 le prixI coûtant'} ~ à perte //sacrifierI.3  
 {'V. Y que X a acheté1 dans le but  
 de le V. avec un bénéfice'} //revendre  
 {'essayer de façon très active  
 de V. Y en faisant la réclame à Z'} //pousser à la ventela  
 {'V. dans le cadre d'enchères'} ~ aux enchères  
 {'V. en acceptantI.4a des  
 paiementsI.1a différés'} ~ à crédit <à tempérament>  
 {'V. en demandant d'abord un  
 engagement écrit de l'acheteur'} ~ par souscription

- { 'le propriétaire de l'objet en cause  
communique qu'il veut le V.' } à ~ [formule écrite]
- ◊ex On vendait des galettes anisées, des pains de seigle d'un  
sou, des pommes vertes. Les producteurs agréés sont ceux  
qui vendent au moins 10 000 litres de lait par année.
- ◊ph ,vendre chèrement [N],  
,vendre la peau de l'ours avant de l'avoir tué,  
,vendre la mère,  
,vendre père et mère,  
,vendre son âme (au diable),  
,vendre son droit d'aïnesse pour un plat de lentilles,  
,savoir se vendre,  
,vendre des frigidaires aux Esquimaux,
- I.2
- ◊cg pas de se-pass, "spéc"
- ◊cs action sociale: personne X ~ individu Y à personne Z  
pour argent W
- ◊nb
- ◊tr X = I = N  
Y = II = N (oblig)  
Z = III = à N  
W = IV = Num N, pour N
- ◊fl { 'officier ministériel dont la fonction  
est de saisir et V. les biens de Y' } //huissier
- ◊ex Les créanciers de Jean avaient envoyé un huissier pour le  
vendre. Si vous faites faillite, ils viendront vous vendre.
- ◊ph
- II
- ◊cg se-pass
- ◊cs action sociale: personne X ~ personne Y à personne Z pour  
W
- ◊nb
- ◊tr X = I = N  
Y = II = N (oblig)  
Z = III = à N  
W = IV = ,en échange, de N [≠ 'avantage matériel'], pour N  
[= 'avantage matériel']
- ◊fl {QSyn} trahir, donner; dénoncer  
{S1} vendu  
{QS1} traître; "fam" cafard, "fam" mouchard  
{AntiMagnS4} "fam" ,une bouchée de pain,, "fam"  
,des clopinettes,, rien, ,trois fois rien,
- ◊ex Judas a vendu Jésus aux prêtres pour trente deniers.  
A propos, je vais te raconter une histoire, tu ne  
me vendras pas [H. de Montherlant].
- ◊ph
- III
- ◊cg se-pass
- ◊cs action sociale: personne X ~ idée Y à personne Z
- ◊nb

Øtr X = I = N  
     Y = II = N (oblig)  
     Z = III = à N  
 Øfl {Conv321} "fam" acheterIII  
     {S1} vendeurII  
 Øex Pour bien vendre sa réforme électorale, le sénateur  
     McCloskey a dû faire des concessions à l'opposition. Depuis  
     trois mois, le parti essaie de vendre son programme à la  
     population.  
 Øph »

On peut remarquer que l'information suivante – extraite de l'article de VENDREI.1 – n'est pas formalisée à 100% si l'on considère ce qui apparaît entre '...'. Ces symboles servent justement à informer les programmes de traitement du DiCo que l'information qu'ils encadrent est une composante sémantique décrite de façon non formelle :

{F1 = 'V. Y en grandes quantités  
 à un seul Z'} "spéc" ~ en gros > "spéc" ~ en demi-gros

## 2.2.2 Un superarticle original de DiCo : ADHÉSION

2 <- Statut  
 « ADHÉSION  
 Δ nom, fém  
 •I  
 Øcg pas de pl  
 Øcs événement physique: ~ de substance X à substance Y  
 Ønb  
 Øtr X = I = de N, A-poss  
     Y = II = à N, sur N  
 Øfl {V0} adhérerI  
     {Magn} bon | prépos, fort  
 Øex Cette surface lisse permet une très bonne adhésion.  
 Øph  
 •II.1  
 Øcg pas de pl  
 Øcs action mentale: ~ de personne X à idée Y  
 Ønb  
 Øtr X = I = de N, A-poss  
     Y = II = à N  
 Øfl {V0} adhérerII.1  
     {Magn} complet, entier, total  
     {Ver} spontané  
     {Oper1} apporter [A-poss ~ à N], donner [A-poss ~]  
     {Oper2} emporter [ART ~]  
 Øex Sa présentation a emporté l'adhésion spontanée de tous  
     ses collègues.  
 Øph  
 •II.2a  
 Øcg

◊cs action sociale: ~ de personne X à organisation Y  
 ◊nb  
 ◊tr X = I = de N, A-poss  
 Y = II = à N  
 ◊fl {V0} adhérerII.2  
 {Magn1-quant} massif | pl  
 {Ver} spontané  
 {Oper2} enregistrer [ART ~]  
 ◊ex L'adhésion du Portugal au Marché Commun...  
 Avez-vous reçu ma demande d'adhésion ?  
 En effet, l'adhésion au club ne revient qu'à 25 F.  
 ◊ph  
 •II.2b  
 ◊cg  
 ◊cs document: ~ de personne X [à organisation]  
 ◊nb  
 ◊tr X = I = de N, A-poss  
 ◊fl {Syn} carte d'adhésionII.2a  
 ◊ex Je ne parviens pas à retrouver votre adhésion.  
 ◊ph »

Révision : 16 IV 95, APol.

### 3 Les éditeurs lexicologiques spécialisés

Un des avantages majeurs que présentent les dictionnaires formalisés comme le DEC ou le DiCo est que, du fait même de leur formalisation, il est possible de contrôler leur rédaction au moyen de programmes spécialement conçus à cet effet. En effet, plus un dictionnaire est formel, plus le traitement automatique de ses articles est aisé. Nous nous intéressons ici à un type de programme particulier, que nous appelons *éditeur de dictionnaire* et que nous présentons tout d'abord de façon générale (3.1). Puis, fidèles à la stratégie adoptée dans ce chapitre, nous nous concentrons sur le cas de l'éditeur du DiCo, l'ÉDiCo (3.2). Finalement, nous élargissons le champ de notre présentation en examinant la caractéristique majeure qui distinguerait un éditeur de DEC par rapport à l'ÉDiCo : le traitement automatique des définitions (3.3).

#### 3.1 Qu'est-ce qu'un éditeur spécialisé de dictionnaire ?

Pour répondre à cette question, il est peut-être plus commode de commencer par résumer brièvement ce qui se fait à l'heure actuelle au niveau de l'utilisation de l'informatique en lexicologie.

Pratiquement toutes les entreprises majeures de rédaction de dictionnaires font appel à l'outil informatique pour remplir au moins les deux fonctions suivantes :

- stocker des données sur la langue (corpus, fiches lexicologiques, etc.) – voir Ch. II, 3.2, p. 52 *ssq.*;
- contrôler la cohérence de la description des lexies proposée dans les dictionnaires en faisant des vérifications et des recherches automatiques sur les articles – voir, par exemple, ce qui est dit dans la table B15 du *Longman Dictionary of Contemporary English* (1987), à propos du vocabulaire restreint utilisé dans les définitions de ce dictionnaire : *We have also used a special computer program that checks every entry to ensure that words from outside the Defining Vocabulary do not appear in definitions.*

Cependant, le type de programme auquel font référence ici les auteurs du dictionnaire *Longman* n'est pas à proprement parler ce que nous entendons par *éditeur de dictionnaire*. En effet, un éditeur de dictionnaire n'est pas simplement un programme qui fait de la vérification sur les articles; c'est aussi tout un environnement informatique permettant la création, l'édition (au sens d'édition de fichiers informatiques<sup>s</sup>) et la consultation des articles d'un dictionnaire. On peut dire que les programmes utilisés aujourd'hui par de nombreuses équipes lexicographiques qui permettent de construire selon un format donné des fiches lexicologiques sur les vocables de la langue sont assez proches de ce que nous entendons par éditeur de dictionnaire; les programmes de vérification de la cohérence du dictionnaire sont des outils annexes qui peuvent, et sans doute doivent, faire partie des utilitaires informatiques d'un tel éditeur.

Il serait inutile de tenter de passer ici en revue tout ce qui se fait en lexicologie moderne au niveau de l'utilisation d'éditeurs de dictionnaires. Cela serait plutôt fastidieux et pour nous, et pour le lecteur, et présenterait surtout relativement peu d'intérêt par rapport au sujet même de l'ILEC – c'est-à-dire, la lexicologie explicative et combinatoire. Nous allons nous concentrer maintenant sur l'éditeur programmé pour le DiCo et sur les possibilités qu'il offre au lexicologue.

### **3.2 Possibilités offertes par l'éditeur du DiCo**

Même si certains programmes d'édition de dictionnaires formels existent déjà (voir par exemple Waters 1991), la construction de l'ÉDiCo, l'éditeur du DiCo, est une tâche nécessaire dans la mesure où il n'est pas possible d'envisager que l'on puisse disposer d'éditeurs tout à fait généraux, pouvant s'appliquer à de multiples approches en lexicologie. Quand on considère un dictionnaire aussi formalisé et avec une base théorique aussi précise que le DiCo (ou le DEC), on doit envisager la construction d'un éditeur spécialisé qui saura retirer tous les avantages de la modélisation offerte par l'approche linguistique Sens-Texte. Il n'y a pas de norme pour l'instant en lexicologie et, en fait, l'introduction d'une norme dans ce domaine est justement une des visées de notre travail.

Nous n'allons pas couvrir ici toutes les fonctions de l'ÉDiCo, et ce pour deux raisons. Premièrement, cela relève plus de la description d'un programme

informatique, description qui est justement faite en détail dans le manuel d'utilisation de l'ÉDiCo (Polguère 1995). Deuxièmement, les outils inclus dans l'ÉDiCo sont changés et améliorés au fur et à mesure que progresse le travail de rédaction du DiCo; il est donc plus intéressant de décrire les fonctions de l'ÉDiCo dans leurs grandes lignes, en mettant l'emphase sur le potentiel d'utilisation des formalismes du DiCo plutôt que sur l'utilisation qui en est faite au moment où nous rédigeons ce chapitre.

L'ÉDiCo permet d'effectuer trois types d'opérations majeurs, que nous allons maintenant présenter :

- l'édition d'articles de dictionnaire proprement dite (3.2.1);
- la recherche sur la base de données dictionnaires (3.2.2);
- la vérification de la cohérence du dictionnaire (3.2.3).

### 3.2.1 Opérations d'édition proprement dites

L'ÉDiCo permet bien entendu d'éditer (au sens informatique) les fichiers contenant les superarticles du DiCo – c'est-à-dire créer, modifier et, plus généralement, travailler sur les descriptions lexicales. Comme tout éditeur informatique récent, l'ÉDiCo fait usage du multi-fenêtrage et de menus. Il offre donc un environnement idéal pour travailler simultanément sur plusieurs articles de dictionnaires, ce qui est une condition nécessaire à l'approche de description lexicale que nous préconisons (par champs sémantiques et lexicaux). Nous avons déjà mentionné (voir Fig. VI-3, en 2.1.1, p. 213) que l'ÉDiCo pouvait aider le lexicologue par le fait même qu'il « connaît » la structure des articles de DiCo et les formalismes utilisés. L'ÉDiCo peut donc construire automatiquement des squelettes d'articles; il peut aussi afficher des informations sur les options offertes dans l'encodage des descriptions (liste des caractéristiques grammaticales, des marques d'usage, des étiquettes sémantiques, etc. qui font partie du formalisme « officiel ») – il possède ainsi potentiellement, en tant qu'éditeur, un ensemble ouvert d'utilitaires aide-mémoire.

À partir de cette connaissance que l'ÉDiCo possède des structures et du formalisme du DiCo, on peut envisager un nombre presque illimité d'utilitaires visant à aider le lexicologue dans son travail de rédaction. Nous ne citons ici que deux exemples, pris parmi les utilitaires déjà incorporés dans l'ÉDiCo : la construction automatique des tableaux synoptiques et l'environnement d'élaboration de la hiérarchie des étiquettes sémantiques.

#### Tableau synoptique

Lorsqu'il travaille avec l'ÉDiCo sur un superarticle, le lexicologue n'a pas à se préoccuper d'en construire le tableau synoptique (voir Ch. IV, 1.5, p. 171). Une fois qu'il a écrit la description d'un vocable donné, il peut demander à l'éditeur de produire directement le tableau synoptique correspondant. Par exemple, le résumé des lexèmes de ADHÉSION que nous avons donné plus haut en 2.1.2 –

lors de la présentation de la zone de caractéristiques sémantiques – est en fait presque exactement le tableau synoptique produit par l'ÉDiCo à partir des articles de ce vocable ! Ce qui est intéressant ici, ce n'est pas seulement que le travail fastidieux d'écriture du tableau synoptique est ainsi épargné au lexicologue; c'est aussi que ce dernier peut faire générer automatiquement un nouveau tableau, à chaque fois que des modifications importantes sont introduites dans la description du vocable. Il peut ainsi avoir toujours à sa disposition un tableau synoptique à jour, c'est-à-dire une table des matières valide pour le superarticle. Il est important de noter que le tableau synoptique est un outil très utile pour le lexicologue (et pas seulement pour l'utilisateur du DiCo), lorsque celui-ci doit se pencher sur une description de vocable déjà créée afin de la modifier ou d'établir des connexions avec d'autres descriptions.

Comme autre illustration de tableau synoptique produit par l'ÉDiCo, nous donnons maintenant celui du vocable VENDRE (que le lecteur peut comparer avec le tableau synoptique du superarticle de DEC, donné au Ch. V, 4, p. 195) :

VENDREI.1

action sociale: personne X ~ Y à personne Z pour argent W  
[= QConv3214(acheter1)]

On vendait des galettes anisées, des pains de seigle d'un sou, des pommes vertes. Les producteurs agréés sont ceux qui vendent au moins 10 000 litres de lait par année.

VENDREI.2

action sociale: personne X ~ individu Y à personne Z pour argent W

Les créanciers de Jean avaient envoyé un huissier pour le vendre. Si vous faites faillite, ils viendront vous vendre.

VENDREII

action sociale: personne X ~ personne Y à personne Z pour W  
Judas a vendu Jésus aux prêtres pour trente deniers.

A propos, je vais te raconter une histoire, tu ne me vendras pas [H. de Montherlant].

VENDREIII

action sociale: personne X ~ idée Y à personne Z

Pour bien vendre sa réforme électorale, le sénateur McCloskey a dû faire des concessions à l'opposition. Depuis trois mois, le parti essaie de vendre son programme à la population.

### Hierarchie des étiquettes sémantiques

Le second utilitaire de l'ÉDiCo dont nous allons parler est celui permettant la consultation et la mise à jour de la hiérarchie des étiquettes sémantiques utilisées dans le DiCo. Nous avons déjà souligné, en 2.1.2, à quel point ces étiquettes présentaient un réel intérêt à la fois théorique et pratique, tout en posant de nombreux problèmes au niveau de leur détermination. En gros, on ne peut pas POSTULER une hiérarchie d'étiquettes sémantiques valide : cette hiérarchie se construit au fur et à mesure de la progression du travail sur le DiCo. Il est donc nécessaire d'inclure dans l'ÉDiCo une série d'utilitaires permettant au lexicologue de construire cette hiérarchie EN MÊME TEMPS QU'IL RÉDIGE SES ARTICLES DE DICTION-



NAIRE. C'est à cette seule condition que l'on peut s'assurer d'obtenir, au bout du compte, un système d'étiquettes sémantiques cohérent, ayant une vraie valeur lexicologique. Les utilitaires correspondants de l'ÉDiCo donnent au lexicologue un accès direct à la hiérarchie existante. Il peut ainsi, par l'intermédiaire de menus, accéder à la description de n'importe quelle étiquette, et voir comment elle s'insère dans la hiérarchie. Reprenons l'exemple de l'étiquette *individu*, déjà utilisé en 2.1.2. L'ÉDiCo la décrira au lexicologue de la façon suivante :

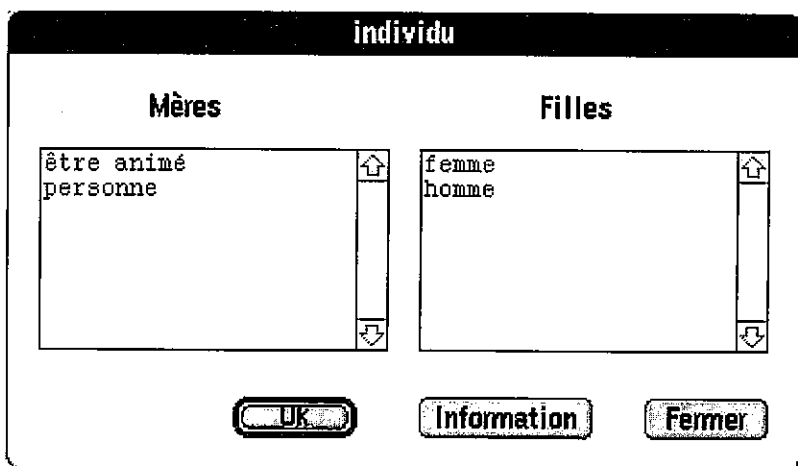


Figure VI-6 Affichage de l'étiquette sémantique *individu* par l'ÉDiCo

La description donnée dans cette figure est consistante avec l'extrait de hiérarchie que nous avons présenté dans la Fig. VI-5 (2.1.2) : *individu* dépend simultanément des deux étiquettes *être animé* et *personne*, et domine les deux étiquettes *femme* et *homme*. À partir de la fenêtre de la Fig. VI-6, le lexicologue peut – en cliquant simplement sur le bouton « information » – ouvrir le fichier contenant (l'encodage de) la base de données des étiquettes sémantiques. Dans le cas qui nous intéresse ici, le fichier sera automatiquement affiché au niveau des données suivantes :

Nom : "individu"

Sub : <femme> <homme>

Exemples : AMOURII.1a [Tu es son premier et unique amour.]

Commentaires : Nécessairement une personne physique. Fait partie  
 : d'un double système d'oppositions :  
 : • s'oppose à 'animal' sous 'être animé';  
 : • s'oppose à 'communauté', 'établissement' et  
 : 'organisation' sous 'personne'.  
 : Voir les actants X et Y dans les articles de AIMER.

Le lexicologue peut éditer et sauvegarder ces données, mettant ainsi à jour la hiérarchie d'étiquettes sémantiques. Par exemple, supposons qu'il ajoute

« <déité> » dans la liste « Sub » ci-dessus et ajoute les données suivantes au fichier en cours d'édition :

```
Nom : "déité"
Sub : --
Exemples : --
Commentaires : --
```

Il aura de ce fait (i) créé une nouvelle étiquette sémantique déité, et (ii) inclus cette étiquette dans la hiérarchie comme dépendant de l'étiquette individu.

On voit que chaque étiquette est spécifiée non seulement par son nom et le nom des étiquettes qu'elle domine, mais aussi par une série d'exemples et par des commentaires qui permettent au lexicologue de garder une trace des évidences linguistiques l'autorisant à postuler une telle étiquette. Nous pensons que cette façon de procéder peut seule permettre d'aboutir, en fin de compte, à une hiérarchie d'étiquettes sémantiques ayant une réelle portée LINGUISTIQUE théorique.

### 3.2.2 Opérations de recherche sur la base de données dictionnaires

Comme nous l'avons mentionné en 1, p. 209, un dictionnaire informatisé permet d'effectuer des opérations de recherche « intelligente » dans les articles. Par *recherche intelligente*, nous entendons essentiellement une recherche sélective qui soit basée sur une analyse de la structure des articles de dictionnaire. Par exemple, au moyen de l'ÉDiCo, on peut faire effectuer des recherches dans la base de données dictionnaires en limitant le domaine de la recherche à une zone particulière. Supposons ainsi que le lexicologue veuille obtenir la liste de toutes les lexies correspondant à des parties du corps – dans le cas, par exemple, où il doit travailler sur un vocable appartenant à ce champ lexical. Pour ce faire, il demande à l'ÉDiCo d'afficher la fenêtre de dialogue suivante, qu'il remplira de la façon indiquée :

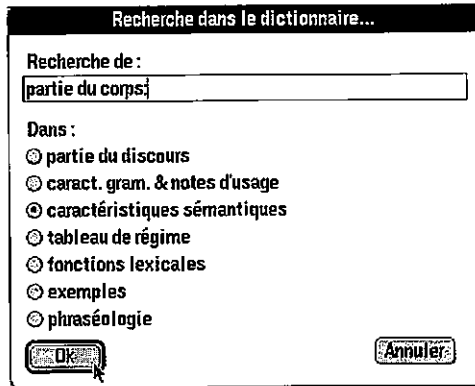


Figure VI-7 Recherche sélective par zones

À partir des spécifications données dans la Fig. VI-7, l'ÉDiCo fournira au lexicologue la liste de toutes les lexies décrites comme des parties du corps : c'est-à-dire toutes les lexies ayant l'étiquette *partie du corps* comme partie gauche de la formule contenue dans la zone de caractéristiques sémantiques. Il est à noter que l'ÉDiCo n'effectuera ses recherches que sur les descriptions de vocables ayant au moins atteint le statut 2. Il n'y a pas lieu en effet de tenter d'analyser la structure formelle des superarticles de statut 3, qui, précisément, ne sont pas valides de ce point de vue (voir la description des statuts en 2.1.1, p. 212).

L'utilitaire de recherche présenté ci-dessus est encore très limité. Cependant, le formalisme du DiCo est tel qu'il ouvre la voie à la programmation d'utilitaires vraiment sophistiqués. Il est ainsi très clair que le lexicologue devrait avoir à sa disposition un véritable langage formel de type logique, permettant d'effectuer des recherches complexes sur la base de données dictionnairiques.

### 3.2.3 Opérations de vérification automatique de la cohérence intra- et interarticles

En plus de la possibilité d'effectuer des recherches intelligentes, nous avons aussi mentionné (en 1) qu'un dictionnaire formel permet de mettre en place des procédures de vérification/validation automatiques. C'est un aspect du travail sur le DiCo qui n'a pratiquement pas été développé pour l'instant. La raison en est que les utilitaires de vérification formelle automatique doivent comprendre de véritables analyseurs syntaxiques complets des articles de dictionnaire. Pour que de tels analyseurs soient construits, il faut dans un premier temps stabiliser les formalismes utilisés dans le dictionnaire. C'est la première tâche que nous nous étions assignée et qui est pratiquement achevée maintenant : les formalismes développés pour le DiCo (et présentés en 2) sont suffisamment puissants pour encoder toutes les informations lexicales dont nous avons besoin dans un dictionnaire des cooccurrences. Il est donc possible de « figer » ces formalismes et d'envisager la construction d'utilitaires de vérification automatique<sup>8</sup>. Notons que, pour l'instant, le seul utilitaire de ce type que nous utilisons permet d'analyser la structure générale d'un superarticle. En gros, il peut déterminer si un superarticle est suffisamment bien structuré pour se voir attribuer le statut 2.

Il y aurait bien d'autres choses à dire à propos de l'éditeur de DiCo. Bien des outils qui viennent d'être présentés ne sont en fait que des ébauches de ce que le formalisme rigoureux du DiCo permet de faire. De plus, il existe de nombreuses autres pistes à suivre dans le cadre de l'édition du DiCo. Nous ne mentionnerons

8 Nous ne prétendons pas qu'il faille figer ces formalismes dès maintenant, une fois pour toute. Ce qui importe, c'est que la formalisation soit définie dans ses grandes lignes, quitte à introduire des modifications ponctuelles, au fur et à mesure qu'avance le travail d'écriture du DiCo – c'est-à-dire au fur et à mesure qu'apparaissent certains besoins d'encodage qui n'avaient pas été anticipés.

pour conclure que le cas fort intéressant de l'encodage des fonctions lexicales. En effet, lorsque l'on isole un phénomène de relation syntagmatique ou paradigmatique entre une lexie vedette et une autre lexie qui semble bien être un élément de la valeur d'une fonction lexicale, il n'est pas toujours aisé d'encoder cette relation au moyen du formalisme très riche et sophistiqué des fonctions lexicales. Il serait utile d'intégrer dans l'ÉDiCo une sorte de système expert de diagnostic des fonctions lexicales. Un tel utilitaire permettrait au lexicologue débutant, ou au lexicologue expérimenté mais confronté à un cas difficile de relation lexicale, de diagnostiquer la fonction lexicale en cause. Nous avons introduit dans l'ÉDiCo une première ébauche d'un tel système expert. Il est très rudimentaire : il permet à peine d'identifier les fonctions lexicales standard simples et ne peut faire aucune suggestion pour des fonctions lexicales complexes (qui sont précisément les cas véritablement problématiques). L'interaction avec cet utilitaire se fait de la façon suivante :

**Identification d'une fonction lexicale**

<b>Classe synt. de L :</b>	<b>Classe synt. de F(L) :</b>																	
<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black;">N</td><td style="border: 1px solid black;">☒</td></tr> <tr><td style="border: 1px solid black;">A</td><td style="border: 1px solid black;">☑</td></tr> <tr><td style="border: 1px solid black;">V</td><td style="border: 1px solid black;">☐</td></tr> <tr><td style="border: 1px solid black;">Adv</td><td style="border: 1px solid black;">☐</td></tr> </table>	N	☒	A	☑	V	☐	Adv	☐	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black;">N</td><td style="border: 1px solid black;">☒</td></tr> <tr><td style="border: 1px solid black;">A</td><td style="border: 1px solid black;">☑</td></tr> <tr><td style="border: 1px solid black;">V</td><td style="border: 1px solid black;">☐</td></tr> <tr><td style="border: 1px solid black;">Adv</td><td style="border: 1px solid black;">☐</td></tr> </table>	N	☒	A	☑	V	☐	Adv	☐	<input type="button" value="OK"/>  <input type="button" value="Annuler"/>
N	☒																	
A	☑																	
V	☐																	
Adv	☐																	
N	☒																	
A	☑																	
V	☐																	
Adv	☐																	
<b>Rôles syntaxiques :</b>																		
<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black;">Subs. L</td><td style="border: 1px solid black;">☒</td></tr> <tr><td style="border: 1px solid black;">L -ATTR-&gt; F(L)</td><td style="border: 1px solid black;">☑</td></tr> <tr><td style="border: 1px solid black;">F(L) -ATTR-&gt; L</td><td style="border: 1px solid black;">☐</td></tr> <tr><td style="border: 1px solid black;">F(L) -IT-&gt; L</td><td style="border: 1px solid black;">☐</td></tr> </table>			Subs. L	☒	L -ATTR-> F(L)	☑	F(L) -ATTR-> L	☐	F(L) -IT-> L	☐								
Subs. L	☒																	
L -ATTR-> F(L)	☑																	
F(L) -ATTR-> L	☐																	
F(L) -IT-> L	☐																	
<b>Relation sémantique entre L et F(L) :</b>																		
<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black;">'L' = 'F(L)' + 'Ω'</td><td style="border: 1px solid black;">☒</td></tr> <tr><td style="border: 1px solid black;">'F(L)' = 'L' + 'Ω'</td><td style="border: 1px solid black;">☑</td></tr> <tr><td style="border: 1px solid black;">intersection( 'L' , 'F(L)' ) ≠ ∅</td><td style="border: 1px solid black;">☐</td></tr> <tr><td style="border: 1px solid black;">intersection( 'L' , 'F(L)' ) = ∅</td><td style="border: 1px solid black;">☑</td></tr> </table>			'L' = 'F(L)' + 'Ω'	☒	'F(L)' = 'L' + 'Ω'	☑	intersection( 'L' , 'F(L)' ) ≠ ∅	☐	intersection( 'L' , 'F(L)' ) = ∅	☑								
'L' = 'F(L)' + 'Ω'	☒																	
'F(L)' = 'L' + 'Ω'	☑																	
intersection( 'L' , 'F(L)' ) ≠ ∅	☐																	
intersection( 'L' , 'F(L)' ) = ∅	☑																	

*Figure VI-8 Fenêtre d'identification d'une relation de fonction lexicale*

La fenêtre de dialogue de la Fig. VI-8 permet au lexicologue de faire une caractérisation approximative de la relation lexicale qu'il a à traiter. Ainsi, s'il cherche à décrire la collocation *demande légitime*, il pourra faire les sélections apparaissant dans la Fig. VI-8, que nous explicitons ci-dessous :

- la lexie vedette présente dans la collocation (= L) est nominale – « classe syntaxique de L : N » ;
- l'élément en cause de la valeur de la fonction lexicale (= F(L)) est adjectival – « classe syntaxique de F(L) : A » ;
- F(L) est un modificateur syntaxique de la lexie vedette dans la collocation – « rôles syntaxiques : L -ATTR-> F(L) » ;

- il n'y a pas d'intersection de sens entre L et F(L) – « relation sémantique :  $\text{intersection}('L', 'F(L)') = \emptyset$  ».

À partir de ces spécifications, l'ÉDiCo informe le lexicologue que les fonctions lexicales standard simples qui correspondent à la caractérisation ci-dessus, et qui donc pourraient être en jeu ici, sont **Magn**, **Plus**, **Minus**, **Ver**, **Bon** et **Pos<sub>i</sub>**. (Dans le cas de la collocation *demande légitime*, nous avons bien entendu affaire à un **Ver**.)

Comme on le voit, les possibilités offertes ici sont assez limitées, mais il y a tout lieu de penser qu'un véritable système expert de diagnostic de fonctions lexicales peut être construit et intégré à l'ÉDiCo.

### 3.3 Possibilités additionnelles qui seraient offertes par un éditeur de DEC véritable

Un éditeur de DEC se distinguerait de l'ÉDiCo avant tout par le fait qu'il devrait être capable de prendre en charge les définitions lexicographiques. Il devrait aider le lexicologue dans son travail de construction de ces définitions en contrôlant :

- la cohérence INTERNE de la définition – voir notamment le problème de la hiérarchisation des composantes, examiné au Ch. III, en 3.3.1.5.1, p. 102 *ssq.* ;
- la cohérence EXTERNE de la définition : Les éléments sémantiques définissants sont-ils bien numérotés ? Fonctionnent-ils dans la définition d'une façon qui est compatible avec leur propre description dans leur article de DEC (par exemple, un prédicat sémantique doit apparaître comme tel dans une définition, et il faut expliciter quels autres sens fonctionnent comme ses actants) ? etc.

Pour que toutes ces opérations soient possibles, il faut que la définition soit manipulée par l'éditeur de DEC sous sa véritable forme, c'est-à-dire en tant que réseau sémantique. Nous avons déjà noté (Ch. III, 3.3.1.1, p. 73) que, même si le DEC courant utilise pour des raisons pratiques une définition prenant la forme d'une expression linéaire en français « traité », la représentation d'un sens est formellement un réseau sémantique. Nous avons aussi mentionné le fait que, dans un cas idéal, le lexicologue devrait être capable de manipuler les deux formes (linéaire et en réseau) de représentation. Nous ne reviendrons pas sur les arguments déjà présentés ; rappelons ici simplement les avantages de chacune des deux représentations :

- Le formalisme de réseau sémantique aide le lexicologue à exercer ses CAPACITÉS LOGIQUES ; en effet, ce formalisme rend explicite (i) les propriétés formelles de chacun des éléments de la représentation – prédicat, opérateur sémantique, etc., (ii) les liens de dépendance sémantique entre les sens (c'est-à-dire les relations du type prédicat-argument) et (iii) l'organisation interne du sens défini (ce que nous avons appelé la hiérarchisation des composantes).

- La représentation linéaire dans une pseudo-langue (du français « traité ») permet au lexicologue de faire jouer son INTUITION LINGUISTIQUE afin de tester la valeur paraphrastique de la définition, ce qui ne serait pas possible ou très difficile en raisonnant directement sur un réseau sémantique.

Dans un véritable environnement d'édition de DEC, le lexicologue devrait donc être capable d'écrire directement ses définitions sous forme linéaire. L'éditeur, cependant, devrait automatiquement analyser ces définitions, déceler les problèmes formels potentiels (en demandant à ce moment au lexicologue de les résoudre), et finalement stocker la définition sous une double forme : linéaire et en réseau. Le lexicologue devrait aussi être capable d'afficher le réseau – c'est-à-dire la représentation graphique de la définition – et de l'éditer afin de corriger les problèmes qui apparaissent mieux dans la représentation graphique (notamment les problèmes de hiérarchisation des composantes). En retour, l'éditeur devrait être capable de générer une nouvelle version de la définition linéaire à partir d'un réseau sémantique ainsi édité par le lexicologue. On voit que les tâches accomplies par un tel éditeur sont assez semblables à celles effectuées, dans le cadre du traitement automatique du langage, en analyse de texte – construction d'un réseau représentant le sens à partir d'une expression linéaire (très proche de la vraie langue), et en synthèse de texte – production d'une expression linéaire linguistique à partir d'une représentation du sens par un réseau.

Il est clair que toutes les opérations de recherche et de vérification formelles que l'éditeur de DEC effectuerait sur les définitions devraient utiliser la représentation en réseau – qui est la seule véritablement « calculable » pour un programme informatique.

Nous touchons ici bien évidemment aux fameux aspects de l'informatisation du DEC, mentionnés au tout début de ce chapitre, qui rendent cette tâche très difficile et qui ont justifié le travail sur le DiCo (le DEC simplifié).

#### **4 DEC et DiCo en tant que bases de données lexicales pour les systèmes de traitement automatique du langage naturel**

Dans cette dernière section, nous établissons le pont entre l'informatisation de l'approche lexicologique THÉORIQUE qui est la nôtre, et ses APPLICATIONS potentielles dans le cadre du traitement automatique du langage naturel [= TALN]. Il existe une grande variété d'utilisations pratiques du TALN et nous ne pouvons pas entrer ici dans le détail de comment utiliser les articles de DEC ou de DiCo dans chaque type d'application. Mentionnons pour mémoire que par *utilisations pratiques du TALN* nous référons à des types de programmes aussi différents que :

- analyseurs de textes;
- générateurs de textes;
- systèmes de traduction automatique et d'aide à la traduction;
- logiciels d'enseignement assisté par ordinateur;
- etc.

Le point essentiel que nous voulons développer ici est que les entrées d'un dictionnaire formel complètement informatisé, DEC ou DiCo, peuvent DIRECTEMENT être utilisées pour construire les bases de données lexicales des systèmes de TALN. Ainsi, il est relativement simple d'écrire un programme de compilation qui traduirait les articles de DiCo en des bases de données Lisp ou Prolog (pour ne citer que deux des langages de programmation couramment utilisés en TALN). Le premier article du vocable ADHÉSION (voir 2.2.2), peut ainsi être compilé en Prolog sous la forme du segment de programme donné ci-dessous – les éléments apparaissant à droite du symbole « % » sont des commentaires, et ne font pas partie du programme exécutable comme tel.

**NB :** Dans un cas comme celui-ci, où nous illustrons nos propos au moyen d'un extrait de programme informatique, nous ne pouvons éviter le fait que seules les personnes possédant des rudiments de connaissances du langage de programmation utilisé peuvent saisir pleinement le principe de formalisation mis en jeu. Pour pallier ce problème, nous introduisons de nombreux commentaires dans ce qui suit, commentaires qui devraient aider le lecteur non spécialiste.

```
% Série de prédicats Prolog 'vocable/3',
% identifiant tous les superarticles de DiCo.
...
vocable('ADHÉSION',
        % • Nom du vocable
        ['I', 'II.1', 'II.2a', 'II.2b'],
        % • Liste des lexies contenues dans ADHÉSION
        [nom, fém]).
        % • Caractéristiques valides
        %   pour toutes les lexies
...
% Série de prédicats Prolog 'lexie/9',
% encodant tous les articles de DiCo.
...
lexie('ADHÉSION',
      % • Renvoi au superarticle correspondant
      'I',
      % • numéro lexical de la lexie
      ['pas de pl'],
      % • liste des caractéristiques grammaticales
      ['événement physique'('adhésionI'),
       substance(X),
       substance(Y)],
      % • caractéristiques sémantiques, où les
      %   étiquettes sémantiques sont encodées sous
      %   forme de prédicats Prolog
      [],
      % • NB (vide)
```

```

[('X', 'I', [de+'N', 'A-poss']),
 ('Y', 'II', [à+'N', sur+'N'])],
  % • Tableau de régime ("+" est défini
  %   comme un opérateur spécial)
[('V0', [adhérerI]),
 ('Magn', [bon:prépos, fort])],
  % • Liste des fonctions lexicales; pour chaque
  %   FL, la valeur est elle-même une liste
['Cette surface lisse permet une très bonne adhésion.'],
  % • Liste d'exemples
[]).
  % • Phraséologie (vide)

```

On peut noter que les parties non complètement formalisées (comme le **NB** ou les exemples) seront stockées comme de simples chaînes de caractères. Elles ne sont pas interprétées par le programme de TALN qui utiliserait le code informatique ci-dessus comme base de données lexicales; la seule chose qu'il peut faire est de les afficher.

Les entrées du type de celle donnée ci-dessus sont déjà relativement sophistiquées, si on les compare à ce qui se fait habituellement dans les applications en TALN. Le fait qu'elles peuvent être automatiquement construites à partir d'un DiCo montre l'importance d'un tel dictionnaire pour ce qui est des applications potentielles de la linguistique informatique. En fait, le mode d'encodage du dictionnaire que nous venons de présenter est sans doute trop primitif. On pourrait tout à fait utiliser les ressources offertes par les techniques de programmation orientée objet pour tirer au maximum parti de la structuration du DiCo. Ainsi, le DiCo pourrait être stocké comme un ensemble d'objets informatiques (objets vocables, lexies, etc.), ayant chacun des champs et des paramètres et pouvant exécuter des méthodes données. Les héritages de propriétés notamment (entre un vocable et ses lexies, entre lexies apparentées, etc.) seraient vraisemblablement mieux encodés si on pouvait faire usage des mécanismes d'héritage inhérents aux techniques de programmation orientée objet.

Ces considérations de programmation dépassent néanmoins le cadre de l'ILEC. Il suffira de retenir, pour ce qui concerne les rapports entre le DiCo et les utilisations pratiques en TALN, que :

1. le DiCo est directement utilisable comme lexique dans les systèmes de TALN;
2. en tant que lexique, il est dans l'ensemble plus riche que ce que l'on trouve habituellement dans les systèmes de TALN.

Cette dernière affirmation demanderait bien sûr à être étayée au moyen d'une étude détaillée de l'état de l'art en TALN. Nous pensons cependant pouvoir omettre une telle étude dans la mesure où il existe de nombreux ouvrages de linguistique informatique traitant précisément du problème des lexiques des systèmes de TALN – un des plus récents ouvrages sur la question est Saint-Dizier & Viegas 1995.



Pour conclure, nous voudrions maintenant revenir à un point essentiel mentionné au début de ce chapitre : le DEC est destiné à être informatisé; c'est sa finalité. Un DEC simplement sur papier, c'est en quelque sorte du gâchis car on ne peut retirer qu'une petite partie des avantages qu'une telle description offre. De plus, un DEC rédigé sans l'aide d'outils informatiques, du type des éditeurs présentés dans ce chapitre, n'aura vraisemblablement pas la cohérence voulue. La théorisation linguistique qui se trouve en arrière du DEC permet et même demande une informatisation. Bien sûr, une informatisation complète, comme nous l'avons démontré, n'est pas une tâche simple. Elle implique que soient faits des ajustements dans la modélisation du DEC (vers une plus grande maîtrise de la formalisation) et elle demande surtout que soient développés des outils informatiques de haut niveau. On peut noter à ce propos que lorsque l'approche du DEC a été élaborée, il y a plus de vingt-cinq ans, son informatisation représentait une tâche presque impossible – en regard du matériel informatique et des logiciels de programmation disponibles à cette époque. Aujourd'hui, compte tenu de l'évolution en informatique, cette tâche semble de plus en plus réaliste. Nous croyons d'ailleurs que ce phénomène est un exemple de la bonne façon de procéder en linguistique informatique et en linguistique formelle. Les linguistes doivent trouver la meilleure façon de modéliser la langue, et c'est à l'informatique de développer les outils adéquats de représentation et de manipulation des connaissances linguistiques. L'approche du DEC peut être considérée comme une approche « de pointe » justement du fait qu'aucun compromis de départ n'a été fait vis-à-vis de ce qu'était l'état de l'art en informatique à l'époque où cette approche a été conçue – même si le DEC avait explicitement pour finalité d'être informatisé. La formalisation et la structure logique du DEC sont les garants du fait qu'une informatisation doit être possible. Évidemment, il y a bien des problèmes qui peuvent échapper aux théoriciens quand ils élaborent une approche formelle, et il est clair que le processus d'informatisation permet, en retour, d'améliorer la théorie et les modèles qu'elle construit – notre conviction étant que ce ne sont là justement que des améliorations, et non des refontes complètes. Les contraintes essentielles sur la construction des principes de base d'une théorie linguistique doivent venir des caractéristiques mêmes de la langue, et non des caractéristiques des outils (informatiques ou non) utilisés pour la décrire. Ces outils, comme tous bons outils, doivent s'ajuster au travail qu'on leur demande de faire.



## RÉFÉRENCES

Apresjan Jurij

- 1969 O jazyke dlja opisanija značenij slov [Sur un langage pour la description des sens lexicaux]. *Izvestija Akademii Nauk SSSR, Serija lit. i jazyka*, 28 : 5, 415-428.
- 1974 *Leksičeskaja semantika. Sinonimičeskie sredstva jazyka*. Moscou : Nauka. 367 p. [Traduction anglaise : *Lexical Semantics*, 1992, Ann Arbor, MI : Karoma].
- 1979 Anglijskie sinonimy i sinonimičeskij slovar' [Les synonymes anglais et un dictionnaire synonymique]. In : A.I. Rozenman & Ju.D. Apresjan (réd.), *Anglo-russkij sinonimičeskij slovar'*, Moskva : Russkij jazyk, 500-543.
- 1986 Integral'noe opisanie jazyka i tolkovyj slovar' [Une description intégrale d'une langue et un dictionnaire explicatif], *Voprosy jazykoznanija*, n° 2, 57-70.
- 1988a Morfoložičeskaja informacija dlja tolkovogo slovarja [Informations morphologiques pour un dictionnaire explicatif]. In : Ju.N. Karaulov (red.), *Slovarnye kategorii*, Moskva : Nauka, 31-59.
- 1988b Tipy kommunikativnoj informacii dlja tolkovogo slovarja [Types d'informations communicatives pour un dictionnaire explicatif]. In : *Jazyk : sistema i funkcionirovanie*, Moskva : Nauka, 10-22.
- 1990a Tipy leksikografičeskoj informacii ob označajuščem leksemy [Types d'informations lexicographiques sur le signifiant d'un lexème]. In : *Tipologija i grammatika*, Moskva : Nauka, 91-108.
- 1990b Formal'naja model' jazyka i predstavlenie leksikografičeskix znanij [Un modèle formel de la langue et la représentation des connaissances lexicographiques]. *Voprosy jazykoznanija*, n° 6, 91-108.
- 1992 O novom slovare sinonimov russkogo jazyka [Sur un nouveau dictionnaire des synonymes russes]. *Izvestija Ross. AN, Serija lit-ry i jazyka*, 51 : 1, 18-39.
- 1995 Theoretical Linguistics, Formal Models of Language and Systemic Lexicography. In : I.-H. Lee (ed.), *Linguistics in the Morning Calm*, 3, Seoul : Hanshin, pp. 3-30.

Arbatchewsky-Jumarie, Nadia, & Lidija Iordanskaja

- 1988 Le champ lexical 'partie du corps' : description sémantique des lexèmes et structure des vocables. In : Mel'čuk *et al.* 1988 : 63-72.

Béjoint, Henri

- 1994 *Tradition and Innovation in Modern English Dictionaries*. Oxford : Clarendon. 276 p.

Bogusławski, Andrzej

- 1970 On Semantic Primitives and Meaningfulness. In : A. Greimas *et al.* (eds), *Sign, Language, Culture*, The Hague – Paris : Mouton, 143-152.

Chaurand, Jacques, & Francine Mazière (réd.)

- 1990 *La définition*. Paris : Larousse. 304 p.

Collignon, L., & M. Glatigny

- 1978 *Les Dictionnaires – Initiation à la lexicographie*. Paris : Cedic.

Deane, Paul D.

- 1988 Polysemy and Cognition. *Lingua*, 75, 325-361.

Décary, Michel

- 1986 *Un éditeur spécialisé pour le Dictionnaire Explicatif et Combinatoire du français contemporain*, mémoire de maîtrise, Département d'informatique et de recherche opérationnelle, Université de Montréal.

- Descamps, Jean-Luc, Gilberte Gagnon, M.-T. Gaultier, Danielle Lehmann, M. Ollivier, & M. Salomé  
1976 *Dictionnaire contextuel de français pour la géologie*. Paris : CREDIF – Didier. 1617 p.
- Di Sciullo, Anna Maria, & Edwin Williams  
1987 *On the Definition of Word*. Cambridge, MA : The MIT Press. 115 p.
- Dostie, Gaétane, Igor Mel'čuk & Alain Polguère  
1992 Le comment et le pourquoi dans l'élaboration des entrées du Dictionnaire explicatif et combinatoire du français contemporain : REPROCHER, REPROCHE et IRRÉPROCHABLE. *International Journal of Lexicography*, 5 : 3, 165-198.
- Dubois, Jean, & Claude Dubois  
1971 *Introduction à la lexicographie : le dictionnaire*. Paris : Larousse. 217 p.
- Goddard, Cliff, & Anna Wierzbicka (eds)  
1994 *Semantic and Lexical Universals*. Amsterdam / Philadelphia : John Benjamins. 510 p.
- Gougenheim, Georges  
1961 *Dictionnaire fondamental de la langue française*. Paris : Didier. 256 p.
- Green, Georgia  
1969 On the Notion 'Related Lexical Entry'. *Papers from the Fifth Regional Meeting, Chicago Linguistic Society*, 76-88.
- Gross, Gaston  
1989 *Les constructions converses du français*. Genève : Droz. 513 p.
- Gross, Maurice  
1981 Les bases empiriques de la notion de prédicat sémantique. *Langage* 63, 7-52.
- Hallig, R., & Walter von Wartburg  
1952 *Begriffssystem als Grundlage für die Lexicographie*. Berlin : Akademie [2<sup>e</sup> éd. : 1962].
- Hartmann, R.R.K. (ed.)  
1983 *Lexicography : Principles and Practice*. London : Academic Press.
- Hausmann, Franz Josef, & Herbert Ernst Wiegand  
1989 Component Parts and Structures of General Monolingual Dictionaries : A Survey. In : F.J. Hausmann *et al.* (eds) 1989 : 328-360.
- Hausmann, Franz Josef, Oskar Reichmann, Herbert Ernst Wiegand & Ladislav Zgusta (eds)  
1989 *Wörterbücher. Dictionaries. Dictionnaires*. Berlin – New York : W. de Gruyter.
- Householder, Fred W., & Sol Saporta (eds)  
1967 *Problems in Lexicography*. Bloomington : Indiana University.
- Hudson, Richard  
1988 The Linguistic Foundations for Lexical Research and Dictionary Design. *International Journal of Lexicography*, 1 : 4, 287-312.
- Iordanskaja, Lidija  
1986 Russian Expressions Denoting Physical Symptoms of Emotions: An Example of Two-argument Lexical Functions. *Lingua*, 69, 245-282.
- Iordanskaja, Lidija, & Igor Mel'čuk  
1990 Semantics of Two Emotion Verbs in Russian: BOJAT'SJA 'to be afraid' & NADEJAT'SJA 'to hope'. *Australian Journal of Linguistics*, 10 : 2, 307-357.
- Juilland, Alphonse, & Alexandra Roceric  
1972 *The Linguistic Concept of Word. Analytic Bibliography*. The Hague – Paris : Mouton  
118 p. [Janua Linguarum. Series Minor, 130].

- Kjellmer, Goran  
1994 *A Dictionary of English Collocations: Based on the Brown Corpus*. 3 volumes. Oxford : Clarendon Press. 2241 p.
- Krámský, Jiří  
1969 *The Word as a Linguistic Unit*. The Hague – Paris : Mouton. 82 p.
- Lakoff, George  
1970 A Note on Vagueness and Ambiguity. *Linguistic Inquiry*, 1, 357-359.
- Landau, Sydney  
1984 *Dictionaries : The Art and Craft of Lexicography*. New York : Charles Scribner's Sons.
- Lepage, François  
1991 *Éléments de logique contemporaine*. Montréal – Paris : Les Presses de l'Université de Montréal – Dunod. 198 p.
- Levin, Beth  
1993 *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago : University of Chicago Press. 345 p.
- Lexicologie  
1961 *Lexicologie et lexicographie françaises et romanes. Orientations et exigences actuelles. Actes du colloques de Strasbourg (12-16 novembre 1957)*. Paris : CNRS.
- Lipka, Leonhardt  
1990 *An Outline of English Lexicology*. Tübingen : Niemeyer. 212 p.
- Lyons, John  
1977 *Semantics*. Vol. I. Cambridge etc. : Cambridge University Press. 371 p.
- Mel'čuk, Igor  
1974a *Opyt teorii lingvističeskix modelej « Smysl ⇔ Tekst »* [Esquisse d'une théorie des modèles linguistiques du type « Sens ⇔ Texte »]. Moscou : Nauka. 314 p.  
1974b Esquisse d'un modèle linguistique du type « Sens ⇔ Texte ». In : *Problèmes actuels en psycholinguistique*, Colloques inter. du CNRS, n° 206, Paris : CNRS, 291-317.  
1978 Théorie du langage, théorie de la traduction. *Meta*, 23 : 4, 271-302.  
1981 Meaning-Text Models : A Recent Trend in Soviet Linguistics. *Annual Review of Anthropology*, 10, 27-62.  
1988 *Dependency Syntax : Theory and Practice*. Albany, N.Y. : The SUNY Press. 428 p.  
1989 Semantic Primitives from the Viewpoint of the Meaning-Text Linguistic Theory. *Cuaderni di semantica*, 10:1, 65-102.  
1993 *Cours de morphologie générale (théorique et descriptive)*. Vol. 1. *Introduction et Première partie : Le mot*. Montréal – Paris : Les Presses de l'Université de Montréal – CNRS. 412 p.
- Mel'čuk, Igor, & Leo Wanner  
1994 Lexical Co-occurrence and Lexical Inheritance. Emotion Lexemes in German : A Lexicographic Case Study. *Lexikos*, 4, 86-161.
- Mel'čuk, Igor, & Alexander Zholkovskiy  
1984 *Explanatory Combinatorial Dictionary of Modern Russian*. Vienna : Wiener Slawistischer Almanach. 992 p.
- Mel'čuk, Igor (avec N. Arbatchewsky-Jumarie, L. Elmitsky, L. Iordanskaja et A. Lessard)  
1984 *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexicosémantiques I*. Montréal : Les Presses de l'Université de Montréal. 172 p.
- Mel'čuk, Igor (avec N. Arbatchewsky-Jumarie, L. Dagenais, L. Elmitsky, L. Iordanskaja, M.-N. Lefebvre et S. Mantha)  
1988 *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexicosémantiques II*. Montréal : Les Presses de l'Université de Montréal. 332 p.

- Mel'čuk, Igor (avec N. Arbatchewsky-Jumarie, L. Iordanskaja et S. Mantha)  
 1992 *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexicosémantiques III*. Montréal : Les Presses de l'Université de Montréal. 323 p.
- Moon, Rosamund  
 1987 *The Analysis of Meaning*. In: Sinclair (ed.) 1987 : 86-103.
- Pergnier, Maurice  
 1986 *Le Mot*. Paris : Presses Universitaires de France. 127 p.
- Picoche, Jacqueline  
 1986 *Structures sémantiques du lexique français*. Paris : Fernand Nathan. 144 p.
- Polguère, Alain  
 1992 Remarque sur les réseaux sémantiques Sens  $\Leftrightarrow$  Texte. In : A. Clas (réd.), *Le mot, les mots, les bons mots*, Montréal : Les Presses de l'Univ. de Montréal, 109 -148.  
 1995 *The « DiCo » Editor User's Manual (ÉDiCo, version 2.0)*. National University of Singapore, Singapour.
- Pottier, Bernard (réd.)  
 1973 *Le langage*. Paris : CEPL. 544 p.
- Pottier, Bernard  
 1991 La lexie : une mise au point nécessaire. In : Zampolli 1991, 207-213.
- Quemada, Bernard  
 1968 *Les dictionnaires du français moderne, 1539-1863 : étude sur leur histoire, leurs types et leurs méthodes*. Paris : Didier.  
 1987 Notes sur lexicographie et dictionnaire. *Cahiers de lexicologie*, 51 : 2, 229-242.
- Rey, Alain  
 1965 A propos de la définition lexicographique. *Cahiers de lexicologie*, 6 : 1, 67-80.  
 1970 *La Lexicologie*. Paris : Klincksieck.  
 1973-76 *Théories du signe et du sens*. Vol. 1-2. Paris : Klincksieck.  
 1977 *Le Lexique : Images et modèles. Du dictionnaire à la lexicologie*. Paris : Armand Colin.  
 1982 *Encyclopédies et dictionnaires*. Paris : Presses Universitaires de France.  
 1990 *Definitional Semantics: Its Evolution in French Lexicography*. In : J. Tomaszczyk and B. Lewandowska-Tomaszczyk (eds), *Meaning and Lexicography*, Amsterdam/Philadelphia : John Benjamins, 43-55.
- Rey-Debove, Josette  
 1966 La définition lexicographique : recherches sur l'équation sémique. *Cahiers de lexicologie*, 8 : 1, 71-94.  
 1971 *Etude linguistique et sémiotique des dictionnaires français contemporains*. La Haye – Paris : Mouton.  
 1978 *Le Métalangage*. Paris : Le Robert.
- Robinson, J.  
 1954 *Definition*. Oxford : Oxford University Press.
- Saint-Dizier, Patrick, & Evelyne Viegas (eds)  
 1995 *Computational Lexical Semantics*. Cambridge, U.K. : Cambridge University Press.
- Sinclair, John M. (ed.)  
 1987 *Looking Up. An Account of the COBUILD Project in Lexical Computing*. London – Glasgow : Collins. 182 p.
- Ščerba, Lev V.  
 1940 [1958] Opyt obščej teorii leksikografii [Essai d'une théorie générale de lexicographie]. In : L.V. Ščerba, *Izbrannye raboty po jazykoznaniju i fonetike*, Leningrad : Izd-vo LGU, 54-91.

- Svenson, Bo  
1994 *Practical Lexicography. Principles and Methods of Dictionary Making*. Oxford : Clarendon.
- Tesnière, Lucien  
1969 *Éléments de syntaxe structurale*. Paris : Klincksieck. 670 p.
- Tournier, Jean  
1985 *Introduction descriptive à la lexicogénétiq ue de l'anglais contemporain*. Paris : Champion – Slatkine.  
1988 *Précis de lexicologie anglaise*. Paris : Nathan.
- Tuggy, David  
1993 Ambiguity, Polysemy, and Vagueness. *Cognitive Linguistics*, 4 : 3, 273-290.
- van der Eijk, Pim, Olga Alejandro & Maria Florenza  
1995 Lexical Semantics and Lexicographic Sense Distinction. *IJL*, 8 : 1, 1-27.
- Wagner, Robert-Léon  
1967-70 *Les Vocabulaires français. Vol. I : Définitions. Les dictionnaires. Vol. II. Les tâches de la lexicologie synchronique. Glossaires et dépouillements. Analyse lexicale*. Paris : Didier.
- Waters, Bruce  
1991 *MacLex for Lexicon Management and Reversal*. SIL, Papua New Guinea.
- Wierzbicka, Anna  
1972 *Semantic Primitives*. Frankfurt am Main : Athenäum. 235 p.  
1980 *Lingua Mentalis. The Semantics of Natural Language*. Sydney etc. : Academic Press. 367 p.  
1987 *English Speech Act Verbs. A Semantic Dictionary*. Sydney etc. : Academic Press. 397 p.  
1992 Back to Definitions : Cognition, Semantics, and Lexicography. *Lexicographica*, 8 : 146-174.  
1992/93 What are the Uses of Theoretical Lexicography? *Dictionaries*, 14 : 46-78.
- Zaliznjak, Anna A.  
1987 O tipax vzaimodejstvija semantičeskix priznakov [Sur les types d'interactions entre traits sémantiques]. In : *Èksperimental'nye metody v psixolingvistike*, Moskva : Institut Jazykoznanija AN SSSR, 136-145.
- Zampolli, Antonio (ed.)  
1991 *Computational Lexicology and Lexicography. Special Issue Dedicated to Bernard Quemada*, Pisa : Giardini [= *Linguistica Computazionale*, vol. VII].
- Zgusta, Ladislav  
1971 *Manual of Lexicography*. The Hague : Mouton.
- Žolkovskij, Aleksandr  
1964a Pravila semantičeskogo analizy [Règles d'analyse sémantique]. *Mašinnyj perevod i prikladnaja lingvistika*, vol. 8, 17-32.  
1964b Leksika celesoobraznoj dejatel'nosti [Lexique de l'activité intentionnelle]. *Mašinnyj perevod i prikladnaja lingvistika*, vol. 8, 67-103.  
[Voir une traduction anglaise des deux articles dans : V. Ju. Rozencvejg (ed.), *Essays in Lexical Semantics*, vol. 1, 1974, Stockholm : Skriptor.]





# INDEX

## INDEX DES DÉFINITIONS

Définition I.1	:	lexique .....	18
Définition I.2	:	vocabulaire .....	18
Définition I.3	:	dictionnaire .....	19
Définition I.4	:	glossaire .....	21
Définition III.1	:	lexème .....	56
Définition III.2	:	phrasème .....	57
Définition III.3	:	lexie .....	57
Définition III.4	:	ambigu .....	60
Définition III.5	:	vague .....	60
Définition III.6	:	vedette .....	70
Définition III.7	:	actant sémantique .....	76
Définition III.8	:	connotation lexicographique .....	112
Définition III.9	:	actant syntaxique profond .....	117
Définition III.10	:	fonction lexicale standard .....	127
Définition III.11	:	fonction lexicale complexe .....	148
Définition III.12	:	configuration de fonctions lexicales .....	149
Définition IV.1	:	pont sémantique .....	157
Définition IV.2/3	:	lien direct/indirect .....	158
Définition IV.4	:	vocable .....	159
Définition IV.5	:	lexie de base d'un vocable .....	159
Définition IV.6	:	champ sémantique .....	173
Définition IV.7	:	champ lexical .....	176

## INDEX DES TERMES ET DES CONCEPTS

## A

actant sémantique .....	36, 39, 72, 75, 76, 93, 109, 143
actant syntaxique [profond/de surface].....	36, 117, 118, 119
ambigu .....	59-60, 61, 86-90
antonyme .....	114, 131
archilexie .....	193
argument de prédicat sémantique .....	75-76
article de dictionnaire .....	29, 69-70
attestation littéraire [des unités lexicales].....	44
autocausation .....	30
axe chronolectal .....	44
axe sociolectal .....	43
axe topolectal .....	44

## B

base de données textuelles .....	50-52
bloc maximal [dans une définition] .....	83-86

## C

calembour [en tant que moyen de distinction des acceptions] .....	64, 65
cercles vicieux [dans les définitions] .....	79, 82
champ lexical .....	172, 176, 181, 232
champ sémantique .....	40-41, 173
choix paradigmatiques .....	126
choix syntagmatiques .....	127
collocation .....	46, 150, 153
composante sémantique commune caractéristique .....	157
composante sémantique faible .....	95, 104
composante sémantique générique .....	94, 104, 109, 188
composante sémantique identificatoire de champ .....	173-175
composante sémantique optionnelle .....	74, 91, 104, 108
composante sémantique présuppositionnelle [= présupposé].....	106
composante sémantique sous-jacente au transfert métaphorique .....	165
condition de nécessité [des composantes d'une définition] .....	92
condition de suffisance [des composantes d'une définition] .....	93
configuration de fonctions lexicales .....	149
connotation lexicographique .....	72, 112, 117, 190, 197
contrainte additionnelle [au régime] .....	119, 122-125
contrainte sur une variable [dans une définition] .....	102, 105, 109
conversif .....	131

cooccurrence différentielle .....	66-67
cooccurrence lexicale restreinte .....	125 <i>ssq.</i>
cooccurents lexicalement contraints .....	126
critère de construction avec « aussi » .....	65
critère de cooccurrence avec modificateurs qualificatif .....	98
critère de cooccurrence avec négation .....	101
critère de cooccurrence avec quantificateurs .....	99
critère de cooccurrence compatible [= de Green-Appresjan] .....	64
critère de cooccurrence différentielle .....	66
critère de dérivation différentielle .....	68
critère de différence sémantique locale/globale .....	63
critère de « mauvais état » .....	115
critère de pertinence linguistique .....	96
critère de sens antonymique .....	115
critère d'interprétation multiple .....	61
critères lexicologiques .....	58

## D

décomposition sémantique .....	79, 82
défini [en tant que partie d'une définition] .....	75, 218
définissant [en tant que partie d'une définition] .....	75
définition lexicographique .....	72-73, 75
définition lexicographique disjonctive .....	62
délimitation des lexies .....	58, 59
dérivation différentielle .....	68
dictionnaire .....	18-19
Dictionnaire de Cooccurrences [= DiCo] .....	207
dictionnaire de langue .....	20
Dictionnaire explicatif et combinatoire [= DEC] .....	10, 25, 29
dictionnaire simplifié .....	207
dictionnairique .....	23
différence sémantique(s) globale(s) .....	63, 186
différence sémantique(s) locale(s) .....	63, 185
disjonction [dans la définition] .....	64-65, 108
distance(s) sémantique(s) entre les lexies .....	87, 167, 170, 187
division des acceptions d'un mot [= délimitation d'une lexie] .....	59, 184
données prosodiques .....	71

## E

éditeur de DEC .....	235
éditeur de dictionnaire .....	207, 227
élément fusionné [de la valeur d'une fonction lexicale] .....	134, 151
élémentarité [des unités de description] .....	56

étiquettes sémantiques .....	218-220, 230, 232
exemples [dans un dictionnaire] .....	121, 124
exemples négatifs .....	124

**F**

---

factif, sens .....	110, 120
faible, composante sémantique .....	94
falsifiabilité des affirmations scientifiques .....	58
fichier texte .....	211
FL = fonction lexicale .....	127, 192, 222
FL complexe .....	143, 144, 148
FL non standard .....	38, 150
FL paradigmatique .....	10, 130
FL standard simple .....	39, 128
FL syntagmatique .....	10
forme propositionnelle .....	75, 188, 218
français « traité » = métrafrançais .....	73
fusion [dans les valeurs des FL] .....	134, 151

**G**

---

glossaire de X .....	18, 21
grammaire du dictionnaire .....	40

**H**

---

héritage lexical .....	175
hiérarchie des étiquettes sémantiques .....	220, 230
homonymes .....	87
hypertexte .....	210

**I**

---

identificatrice de champ, composante sémantique .....	173
indicateur de champs .....	180
indice actanciel .....	135
industries de la langue .....	205
informations lexicographiques .....	23
informatisation des descriptions lexicologiques .....	12
informatisation du DEC .....	206
introspection .....	50
intuition(s) linguistique(s) .....	50, 86

**L**

---

langage définitoire = langue de la définition .....	73, 78, 89
lexème .....	16, 56
lexicographie .....	26

lexicologie .....	26
lexie .....	16, 22, 56, 57
lexie de base [d'un vocable] .....	159, 163
lexie vedette [d'un article de dictionnaire] .....	33, 70
lexies figurées .....	164
lexique .....	18
lien sémantique direct .....	158
lien sémantique indirect .....	158
linguistique informatique .....	205
liste de mots .....	20
liste de vérification lexicographique .....	108
locution .....	45
locution semi-figée .....	150

## M

---

macrostructure [du dictionnaire] .....	12, 32
«métafrançais» .....	73
métalangage formel .....	34
métaphore [dans le dictionnaire] .....	163-164, 186
métonymie [dans le dictionnaire] .....	163-164
microstructure [du dictionnaire] .....	12, 32
mot .....	15, 55
mot-clé [d'une FL] .....	127

## N

---

nœud dominant .....	75
numéro lexical distinctif .....	86-88, 166, 215

## O

---

obligatoire, actant .....	190, 191
ordonnancement des articles de dictionnaire (= des lexies) au sein d'un vocable .....	162, 187

## P

---

« portabilité » de la base de données .....	211
paraphrasage .....	130
performatif .....	111
perspective onomasiologique .....	48
perspective sémasiologique .....	48
pertinence linguistique .....	96
pertinence sémantique .....	107
phrase à cooccurrence compatible .....	64
phrasème .....	16, 56-57, 97, 114, 194, 247
phrasème complet .....	46, 153

polysème .....	59
polysémie .....	87, 97, 155
pont sémantique .....	97, 114, 117, 157-158, 160-161
prédicat sémantique .....	76, 109
présupposé .....	102, 110
primitifs sémantiques .....	82-85
principe de cohérence interne .....	36
principe de consultation obligatoire des bases de données textuelles .....	50-51
principe de disponibilité .....	43
principe de formalité .....	34
principe de la primauté du locuteur .....	49
principe de lexicalité .....	45
principe de traitement par champs sémantiques et lexicaux .....	178
principe de traitement uniforme .....	40
principe d'exhaustivité .....	42
programmation orientée objet .....	238
prosodie [dans le dictionnaire] .....	71-72

**Q**

---

quasi-phrasème .....	46, 153
quasi-prédicat .....	77, 219

**R**

---

rasoir d'Occam .....	185
réflexivation .....	30
régime .....	36, 119-120, 190, 221
règle de caractère typique de la lexie .....	166
règle de décomposition .....	80
règle de hiérarchie des composantes sémantiques .....	165
règle de la forme propositionnelle .....	79
règle de priorité métonymique .....	164
règle de proximité sémantique .....	163
règle de standardisation .....	86
règle de substituabilité .....	91
règle d'inclusion sémantique .....	163
règle d'ordonnancement par pertinence sémantique .....	107
règle du bloc maximal .....	84
regroupement des acceptions .....	186
régularité de la distinction sémantique .....	87, 168
représentation sémantique .....	73
réseau sémantique .....	73, 235

**S**

schéma généralisé [de la définition/du vocable] .....	42, 175
section publique [dans un article de dictionnaire] .....	175
sémantiquement plus simple .....	30, 80
semi-phrasème .....	46, 153
sens dénotatif .....	73
sens figuré .....	164
sens prédicatif .....	76
substituabilité du défini et du définissant .....	93, 189
supplément sémantique .....	63
superarticle [de dictionnaire = vocable] .....	156, 167
synonyme(s) .....	82, 86, 130
système expert de diagnostic des fonctions lexicales .....	234

**T**

tableau de régime .....	119, 221
tableau synoptique [du vocable] .....	171, 229
théorie linguistique Sens-Texte .....	17
trait sémantique .....	218
Traitement Automatique du Langage Naturel .....	208
typique pour le champ lexical [lexie] .....	166

**U**

unité lexicale [= lexie] .....	16
utilitaire [programme informatique] .....	214, 228, 229, 231, 233

**V**

vague .....	60, 62, 64
valeur [d'une fonction lexicale] .....	127
variable [dans la définition] .....	76, 78, 79, 218
variable scindée .....	199
variation chronologique/chronolectale .....	24, 44
variation géographique/topolectale .....	24, 44
variation sociale/sociolectale .....	24, 43
vedette [= lexie vedette) .....	70
verbe causatif .....	143
verbe support .....	126, 138
vocable .....	156, 159-160, 163, 167, 184, 214
vocabulaire .....	18

**Z**

zeugme .....	64-65
zone phonologique [de l'article de dictionnaire] .....	71

## INDEX DES VOCABLES TRAITÉS

ACCEPTER .....	66-67
ADHÉSION .....	218-219
AIDER .....	106, 120-125
AIL .....	100
ALLEMAND (nom) .....	40-41
AMERLOQUE .....	109
ÂNE .....	45-46
ANNONCER .....	110
APPLAUDISSEMENT .....	98
ASTRONOME .....	81
ASTRONOMIE .....	81
AUSSI .....	65
BAISSER .....	90
BARRIR .....	105-106
BATTRE .....	99
BELLE-MÈRE .....	64-65
BLESSÉ (nom) .....	38-39
BLESSURE .....	38-39
BOMBARDER .....	65
BRAS .....	165, 176
BRETONS (nom) .....	41
BRÛLER .....	163-165
BUREAU .....	158-160
CAROTTE .....	100-101
CÉLIBATAIRE (nom) .....	37-38
「C'EST DU VENT」 .....	111, 114
CHAISE .....	103-104,
CHANGER .....	50, 58-59, 160, 163, 168
CHINOIS (nom) .....	40
CHOU .....	100-101
CŒUR .....	115, 175
COMBLER .....	87
CONTRÔLE [au sens de 'maîtrise [d'un véhicule]'] .....	23
COUCHER ~ SE COUCHER .....	30
COUDE .....	166
「COUPER LES PONTS」 .....	16
COUTEAU .....	88-89
CUILLÈRE .....	88-89
DOIGT .....	107, 164, 166
DOS .....	163, 168
ÉCRIRE .....	39



ÉLEVER .....	68
「EN COUP DE VENT」 .....	112-113
ENDURCI .....	37-39
ENSEIGNER .....	74, 85
ÉPAULE .....	115
ÉPOUX .....	76, 77, 132
ESPAGNOL (nom) .....	40
ÉTUDIANT .....	93-95
ÉTUDIANTE .....	94
FACTURE .....	110
FÂCHÉ .....	24
FAIRE SAVOIR .....	110
FEMME .....	80
FLAMBER .....	65, 171-172
FRANÇAIS (nom) .....	40
FRÈRE .....	95
GUÉRIDON .....	80
HAUTEUR .....	87
HOMME .....	80
INTERDIRE .....	109
JAMBE .....	176
LAPIN .....	155-156
LAPINE .....	156
LETTRE .....	109
MAGAZINE .....	35-37, 40
MARTEAU .....	88-89
MERCI (BIEN) .....	71
MER .....	169-170, 179
MINE .....	158, 160
MONTRE .....	88-89
NEIGE .....	97-98
NEZ .....	168
NUAGE .....	97
NIVEAU .....	104
OCCUPER, S' .....	57, 62, 65, 67
OIGNON .....	24, 100-101
「ON NE PEUT PAS SE PLAINDRE」 .....	72
ORAGE .....	97
ORAGEUX .....	97
PAYER .....	103, 105, 126, 183
PEINDRE .....	60-61, 173
PERMETTRE .....	102
PIED <sup>1</sup> 'partie du corps' .....	157, 161

PIED <sup>2</sup> ('unité minimale rythmique d'un vers') .....	161
PLEUVOIR .....	174
PLUIE .....	128, 173-175
PONT .....	16
['PONT AÉRIEN'] .....	16
['PONT AUX ÂNES'] .....	16, 46
['PONTS ET CHAUSSÉES'] .....	16
PRENDRE .....	64, 69
POUR .....	88
PRODUCTION .....	125
PROFESSEUR .....	74, 80, 85, 91-94, 104-105
PROPOSER .....	43, 61
RASER ~ SE RASER .....	30
['RAT DE CAVE'] .....	46-47, 153
RÉALISER [au sens de 'se rendre compte'] .....	23
REFROIDIR, SE .....	62, 64, 66
REMERCIER .....	110
REEMPLIR ~ SE REMPLIR .....	30
REPROCHER .....	78, 109
RIPOSTER .....	82
RIZ .....	96-97
RUSSE(S) (nom) .....	40-41
SEL .....	96-97
SKIS .....	40, 111
SUCRE .....	96-97
['SUR PIED'] .....	157
TABOURET .....	103-104, 177
TANTE .....	60-61, 65
TÊTE .....	76, 114-115, 166, 176
VENDRE .....	63-64, 77-78, 181 <i>ssq.</i> , 230
VENDRE, SE .....	200-201
VENDEUR .....	202-203
VENT .....	111-114
VEUVE (nom) .....	101-102
VIEUX .....	37-39
VITESSE .....	168
VRAI .....	31, 71
YEUX .....	40, 111, 177



Cet ouvrage se veut une initiation systématique à la lexicologie moderne. La théorie lexicologique présentée s'inscrit dans le cadre de l'approche Sens-Texte et met l'accent tout particulièrement sur l'explication sémantique et la combinatoire des unités lexicales. On propose un modèle de dictionnaire. Après avoir défini les caractéristiques du domaine lexicologique, l'ouvrage décrit en détails les traits spécifiques d'un **Dictionnaire explicatif et combinatoire** et en formule les principes de rédaction. Les divers chapitres abordent les problèmes spécifiques à la microstructure, à la macrostructure et à l'informatisation du dictionnaire.

L'ouvrage donne de très nombreux exemples bien expliqués. L'ensemble est une généralisation et une systématisation d'un modèle de dictionnaire à discuter, à développer et à imiter. Cet ouvrage est destiné à tous ceux qui s'intéressent aux questions de linguistique théorique et appliquée.

---

Igor A. Mel'čuk

*Après des études à Moscou, il a été chercheur à l'Académie des sciences de Moscou. Depuis 1978, il est professeur à l'Université de Montréal. Il est le fondateur de l'approche linguistique Sens-Texte et l'auteur de très nombreuses publications en linguistique théorique et descriptive. Il est membre de la Société royale du Canada et vient d'être nommé professeur au Collège de France.*

André Clas

*Après des études en France, en Allemagne et au Canada, il est professeur à l'Université de Montréal. Il est directeur de la revue Meta – Journal des traducteurs/Translators' Journal, et coordonnateur du Réseau Lexicologie, Terminologie, Traduction de l'Agence francophone pour l'enseignement supérieur et la recherche (AUPELF-UREF). Il est l'auteur de nombreuses publications en linguistique appliquée.*

Alain Polguère

*A fait ses études en France et au Canada. Il s'intéresse tout particulièrement au traitement automatique de la langue. Après avoir travaillé à Montréal sur les systèmes de génération de textes multilingues (Recherche & Développement), il a été professeur à l'Université Nationale de Singapour. Il vient d'être nommé professeur à l'Université de Montréal.*

150 FRF - 820 BEF - 40 CDN - Prix préférentiel UREF 60 FRF.

ISBN 2-8011-1106-6



9 782801 111062

LEEXCO

C200

Diffusion Edicel ou Ellipses selon pays

1995

59.