
*Notions de base
en lexicologie*

ALAIN POLGUÈRE

(Version préliminaire septembre 2002, pour LNG 1080)

**Observatoire de Linguistique Sens-Texte (OLST)
Département de Linguistique et traduction
Université de Montréal
Montréal (Québec) — Canada**



© Alain Polguère, 2000, 2001, 2002
Observatoire de Linguistique Sens-Texte
<http://www.fas.umontreal.ca/ling/olst>



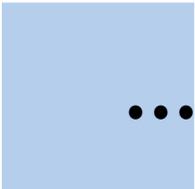
Table des matières

Table des matières	i
Avant-propos	1
Chapitre 1 Quelques notions préliminaires	3
La langue : objet d'étude de la linguistique	4
Définition de la notion de langue	4
La parole : actualisation de la langue	6
Langue vs langage	8
De quoi sont constituées les langues ?	9
Lexique et grammaire	9
Niveaux de fonctionnement des langues	10
Difficulté de l'étude linguistique	11
Limitations du présent ouvrage	11
Lectures complémentaires	12
Exercices	12
Chapitre 2 Le signe linguistique	15
Signe et sémiotique	16
Définition du signe	16
La science des signes	17
Types de rapports contenu-forme dans les signes	18
Le signe linguistique selon Ferdinand de Saussure	21
Les deux composantes du signe linguistique	21
Caractère arbitraire du signe linguistique	22
Caractère figé du signe linguistique	22
Caractère évolutif du signe linguistique	23
Caractère linéaire du signe linguistique	23

Types de signes linguistiques	24
Signe lexical vs grammatical	24
Signe élémentaire vs complexe	25
Propriétés de combinatoire des signes linguistiques	25
Lectures complémentaires	27
Exercices	28
Chapitre 3 L'unité lexicale ou <i>lexie</i>	31
Mot, mot-forme et lexème	33
Le mot n'est pas une notion linguistique	33
Le mot-forme	34
Le lexème	36
Les locutions	38
Définition de la notion de locution	38
(Non-)compositionnalité sémantique	39
Définition de la notion de lexie	41
Regroupement des lexies en vocables	41
Étude des lexies et étude de la sémantique des langues	43
Remarques sur la terminologie et les conventions d'écriture	44
Lectures complémentaires	45
Exercices	46
Chapitre 4 Éléments de morphologie	47
Signe morphologique élémentaire	48
Le morphe	48
Le morphème	50
Radical et affixe	52
Flexion	53
Définition de la flexion	53
Remarque sur les signes zéro	54
Dérivation	55
Définition de la dérivation	55
Types de liens dérivationnels	55
Dérivation synchronique vs dérivation diachronique	57
Composition	58
Lectures complémentaires	60
Exercices	60

Relations sémantiques fondamentales	120
Hyperonymie et hyponymie	120
Synonymie	122
Antonymie	124
Conversivité	126
Homonymie	127
Polysémie	128
La modélisation formelle des relations lexicales	130
La notion de fonction lexicale	131
Fonctions lexicales paradigmatiques	132
Les collocations	135
Fonctions lexicales syntagmatiques	137
En guise de transition...	144
Lectures complémentaires	146
Exercices	147
Chapitre 8 L'analyse du sens	151
La définition lexicale : outil d'analyse du sens	152
Définition par genre prochain et différences spécifiques	152
Méthodologie d'élaboration d'une définition	154
Le problème des cercles vicieux	158
Analyse par champs sémantiques	159
Remarque sur les notions d'ambiguïté et de vague	160
Analyse componentielle	164
Structure sémantique des vocables	165
Exercice pratique sur la métaphore	169
Présentation de l'exercice	169
La métaphore est-elle lexicalisée ?	169
La métaphore est-elle une valeur de fonction lexicale ?	171
Lectures complémentaires	172
Exercices	173
Chapitre 9 La lexicographie	175
Dictionnaires et lexicographie	176
Qu'est-ce qu'un dictionnaire ?	176
Types de dictionnaires	178
Macrostructure et microstructure des dictionnaires	179
Les définitions lexicographiques	181

Lectures complémentaires	183
Exercices	183
Chapitre 10 La pragmatique	185
La pragmatique	186
Nature des échanges linguistiques	187
La théorie des actes de parole	190
Interférences pragmatiques dans le lexique	193
Lectures complémentaires	195
Exercices	195
Conclusion	197
Annexe : Correction des exercices	199
Index des notions	203
Index des auteurs	209



Avant-propos

Ce livre est un cours d'introduction à l'étude du lexique. Il est constitué de dix chapitres, qui doivent être étudiés de façon linéaire. En effet, chaque chapitre correspond à une leçon introduisant un ensemble de notions de base, dont la plupart sont réutilisées dans la suite du texte.

Deux outils permettent au lecteur de naviguer à travers le réseau notionnel du cours :

- 1 au tout début de chaque chapitre, une liste des notions introduites (selon leur ordre d'apparition dans le texte) ;
- 2 à la fin de l'ouvrage, un index alphabétique (voir *Index des notions*, page 203 et suivantes).

Les notions importantes apparaissent dans une **typographie spéciale** lorsqu'elles sont introduites pour la première fois ou font l'objet de remarques importantes.



La paire de lunettes apparaissant dans la marge signale au lecteur une partie de texte qui doit être lue avec une attention spéciale.



La main tenant un crayon est utilisée pour marquer l'introduction d'un symbole ou d'une convention d'écriture dont il sera fait usage dans la suite de l'ouvrage.

Chaque chapitre se termine par une liste de lectures complémentaires permettant de consolider les notions introduites et par des exercices pratiques mettant en jeu ces notions. On trouvera en fin d'ouvrage (*Annexe : Correction des exercices*, page 199 et suivantes) de brefs corrigés pour la plupart de ces exercices.

Il existe de nombreux livres d'introduction à la lexicologie et à la sémantique. Certains peuvent compléter avantageusement le présent

ouvrage, soit parce qu'ils proposent une perspective différente sur la question, soit parce qu'ils ne couvrent pas exactement les mêmes domaines d'étude (plus d'importance donnée à la morphologie, à l'évolution du langage, etc.). Aucun manuel n'est parfait, ce qui vaut très certainement pour celui que le lecteur a maintenant entre les mains. Il est donc toujours utile de consulter différentes sources, par exemple :

Eluerd, Roland (2000) *La lexicologie*. Collection « Que sais-je ? », n° 3548, Paris : Presses Universitaires de France.

Lehmann, Alise et Françoise Martin-Berthet (1998) *Introduction à la lexicologie : Sémantique et morphologie*. Collection « Lettres Sup », Paris : Dunod.

Niklas-Salminen, Aïno (1997) *La lexicologie*. Collection « Coursus », Paris : Armand Colin.

Picoche, Jacqueline (1977) *Précis de lexicologie française*. Collection « Nathan-Université », Paris : Nathan.

Touratier, Christian (2000) *La sémantique*. Collection « Coursus », Paris : Armand Colin.

Pour conclure, je tiens à remercier du fond de mon cœur Lidija Iordanskaja, François Lareau, Igor Mel'čuk, Jasmina Milićević et Ophélie Tremblay pour leur lecture des versions préliminaires de cet ouvrage. Ils ont parsemé mes manuscrits de commentaires du type , ,  et autres ¹, m'aidant par là-même à m'approcher de la ou d'une vérité, dont il faut croire à toute force qu'elle existe.

1. Avec quelques ☺, tout de même !



Chapitre 1

Quelques notions préliminaires

Lexicologie, linguistique, langue, locuteur, destinataire, parole, langue parlée vs écrite, langage, lexique, grammaire, sémantique, syntaxe, morphologie, phonétique, diachronie, synchronie.

L'un des convives amena vers lui les cartes éparses, débarrassant ainsi une bonne partie de la table ; mais il ne les rassembla pas en un seul paquet ni ne les battit ; il prit une carte, et la posa devant lui. Nous notâmes tous la ressemblance de son visage avec celui de la figure peinte : il nous parut qu'avec cette carte il voulait dire « je » et qu'il s'apprêtait à nous raconter son histoire.

Italo Calvino, *Le château des destins croisés*

Ce livre vise à familiariser le lecteur avec les notions de base nécessaires à l'étude du lexique et de son utilisation. Les mots sont au cœur de la connaissance linguistique puisque parler une langue consiste avant tout à combiner des mots au sein de phrases en vue de communiquer. Il serait donc légitime de considérer la **lexicologie**, la discipline qui étudie les phénomènes lexicaux, comme étant la branche maîtresse de la linguistique.

La spécificité de la lexicologie va se préciser au fur et à mesure que nous progresserons dans le présent ouvrage. Il faut cependant dès maintenant définir certaines notions de base sur lesquelles s'appuie la lexicologie, notions qui relèvent toutes de la discipline mère qu'est la linguistique.

Nous allons donc commencer par préciser ce qu'est la *linguistique* en tant que science, c'est-à-dire que nous allons définir son objet d'étude : la langue.

La langue : objet d'étude de la linguistique

Définition de la notion de langue

Adoptons, pour commencer, une définition très approximative de la notion de langue :

La *langue* est notre « outil » de communication privilégié. Chaque langue est un système de signes conventionnels et de règles de combinaison de ces signes, qui forment un tout complexe et structuré.

Cette définition de la langue met en évidence deux points importants. Tout d'abord, comme le présuppose l'expression *outil privilégié*, la langue n'est pas le seul outil que nous utilisons pour communiquer.

En effet, pour nous exprimer, pour transmettre de l'information, nous employons de nombreuses autres ressources que la langue. En voici quatre exemples :

- 1 gestes de la main — agiter la main pour dire au revoir, mettre l'index devant la bouche pour demander le silence, ... ;
- 2 expressions faciales, qui sont des formes de gestes faits avec le visage — tirer la langue, sourire, faire la moue, ... ;
- 3 gestes faits avec l'ensemble du corps — tourner le dos à quelqu'un, croiser les bras et baisser la tête pour bouder, ... ;
- 4 gestes basés sur un contact physique avec une autre personne — serrer la main, donner une tape sur l'épaule, embrasser, ...

Qui plus est, nous sommes entourés de machines ou d'objets qui ont été construits, programmés et installés dans notre environnement quotidien pour nous transmettre de l'information et, donc, nous faire récepteurs d'une forme de communication : guichets automatiques, panneaux de signalisation, feux de circulation, horloges, sonneries, etc.

Nous pouvons aussi, par jeu ou par nécessité, créer de toutes pièces de nouveaux systèmes de communication, un peu comme dans le court texte donné en exergue au début de ce chapitre.

La deuxième caractéristique importante que la définition ci-dessus met en évidence est qu'une langue est un **système** de signes¹ et de règles. Il faut entendre par là que ce n'est pas un simple répertoire d'éléments indécomposables et autonomes servant à communiquer. Les éléments constitutifs de chaque langue sont liés ; ils sont fait pour interagir et se combiner. C'est cette organisation interne d'une langue qui en fait un outil de communication particulièrement puissant permettant de produire un nombre infini de messages différents.

Si j'ai indiqué que les signes et les règles linguistiques sont conventionnels, c'est pour mettre en évidence le fait qu'ils fonctionnent comme une sorte de norme, d'ensemble de lois régissant la façon dont nous communiquons. Ces lois, nous ne les possédons pas de façon innée. Nous devons les apprendre de façon progressive. Le fait qu'il ne soit pas nécessaire d'avoir suivi un enseignement « scolaire » pour parler une langue ne doit pas nous faire oublier que la maîtrise d'une langue est le résultat d'un apprentissage.

Les deux caractéristiques de la notion de langue mentionnées dans la définition ci-dessus sont bien entendu liées. Malgré la grande variété des modes de communication auxquels nous avons recours, les langues restent, du fait de leur grande puissance expressive, nos outils privilégiés pour échanger de l'information, organiser notre pensée et, en fait, exister en tant qu'êtres humains. Pour s'en convaincre, il suffit de penser aux deux étapes du développement d'un très jeune enfant qui vont toujours être mentionnées par ses parents :

- le moment où il a commencé à marcher — mode de déplacement caractéristique de l'humain ;
- le moment où il a commencé à parler (et quels étaient ses premiers mots) — mode de communication caractéristique de l'humain.

1. On pourrait dire de façon très grossière qu'il s'agit des mots de la langue. La notion de signe linguistique est cependant très complexe et c'est pourquoi le Chapitre 2 lui est entièrement consacré.

Noter, pour terminer, que les langues ne forment pas un système inerte, figé pour toujours dans le temps. Les langues naissent, évoluent dans le temps et meurent. Bien entendu, elles n'ont pas de dates de naissance et de décès précises et leur évolution est extrêmement graduelle. Il n'en demeure pas moins qu'on ne parlait pas le français il y a deux mille ans, que le français que l'on parle aujourd'hui à Montréal ou à Paris n'est pas le même que celui qu'on y parlait il y a trois cents ans et qu'il est tout à fait possible que plus personne ne parle le français sur terre d'ici deux mille ans. Du fait de sa nature sociale, des liens étroits qu'elle entretient avec la société humaine qui l'utilise, chaque langue est destinée à évoluer, à se transformer et, éventuellement, à disparaître suivant en cela l'évolution des sociétés qui en font usage.

La parole : actualisation de la langue

En tant qu'ensemble de signes et de règles de combinaison de ces signes, la langue peut être considérée comme étant une entité « abstraite », comme l'est un code civil, ou plus généralement, un ensemble de conventions sociales.

Tout le monde n'est sans doute pas habitué à considérer la langue ainsi et il est utile d'illustrer ce qui vient d'être dit. Soit la phrase suivante :

(1) *Est-ce que tu peux me passer le sel ?*

Cette phrase contient des mots du français, qui sont « assemblés » selon certaines règles grammaticales de cette langue. Elle est construite à partir d'un tout petit sous-ensemble du système de la langue française, mais, cependant, cette phrase elle-même n'est pas une partie de la langue : c'est un exemple, parmi une infinité d'autres exemples possibles, d'une actualisation de la langue. C'est un produit de la langue en quelque sorte.

On remarquera la tournure bien particulière dont je viens de faire usage pour introduire (1) : *Soit la phrase suivante...* Cette façon de faire, typique d'un ouvrage de linguistique, a quelque chose d'un peu étrange, si l'on y réfléchit bien. Elle présente un énoncé comme s'il sortait de nulle part. Elle présuppose que nous allons être capable de discuter d'un énoncé en l'isolant totalement de la situation dans laquelle il a pu être produit. Or, pour que (1) « soit », il nous faut considérer au moins deux

individus. L'un, appelé *locuteur*, a voulu communiquer un message au second individu, le *destinataire*. Pour ce faire, il a utilisé un sous-ensemble des signes et des règles de la langue, sous-ensemble qui lui a permis de réaliser linguistiquement le message qu'il voulait communiquer, sous la forme de l'énoncé (1).

Nous savons donc que les signes et les règles linguistiques (la langue elle-même) existent parce qu'ils s'actualisent dans des comportements particuliers et donnent lieu à des « événements linguistiques » dans lesquels un locuteur communique une information à un destinataire. L'actualisation de la langue n'est pas la langue elle-même et il nous faut donc disposer d'un terme particulier pour la désigner : il s'agit de la parole.

La langue trouve son actualisation dans la *parole*, c'est-à-dire dans des instances d'échanges langagiers entre au moins deux individus : le *locuteur* et le *destinataire*.

L'usage de termes tels que *parole* et *locuteur* ne doit pas faire penser que la linguistique, telle que définie ici, ne s'intéresse qu'à la *langue parlée* et ignore la *langue écrite*. Cette terminologie est simplement le reflet du fait que la langue est par nature avant tout orale, l'écrit étant originellement une transcription plus ou moins fidèle (selon les langues et les systèmes d'écriture) de la chaîne linguistique parlée. On appelle *parole* l'actualisation de la langue parce que l'oral est la forme première d'actualisation de la langue. Mais on sait fort bien que l'on ne s'exprime pas de la même façon à l'écrit qu'à l'oral, que les choix lexicaux et grammaticaux peuvent considérablement varier selon le médium linguistique utilisé et que la linguistique doit donc prendre en considération aussi bien la langue parlée que la langue écrite. Cela est d'autant plus nécessaire que le code écrit a pris de plus en plus d'importance au cours des siècles. Il y a tout d'abord eu l'invention et la généralisation de l'imprimerie, qui a en quelque sorte « industrialisé » la diffusion des textes. Puis, beaucoup plus récemment mais avec des conséquences tout aussi considérables, l'apparition des ordinateurs, du traitement de texte et d'Internet, qui a fait croître de façon exponentielle la production et la diffusion linguistique écrite. Cette nouvelle forme de manipulation physique de la langue a, de plus, donné naissance à un phénomène d'hybridation des codes ; en effet, la frontière entre langue parlée et

langue écrite devient presque impossible à cerner dans certains contextes, comme dans le cas du courrier électronique (voir l'Exercice 3 à la fin de ce chapitre).

Langue vs langage

Pourquoi parlons-nous ? Pourquoi communiquons-nous au moyen de langues ?

Les langues que nous maîtrisons, nous les avons apprises et nous avons pu les apprendre pour plusieurs raisons :

- 1 ce sont des outils de communication que la vie en société nous **impose** d'acquérir ;
- 2 ce sont des systèmes de signes et de règles de combinaison de ces signes que notre cerveau a la capacité de mémoriser et de manipuler ;
- 3 elles se manifestent physiquement par des sons que notre constitution biologique nous permet de produire (appareil phonatoire) et de percevoir (appareil auditif).

Les langues sont donc liées directement à des prédispositions sociales, psychiques et physiologiques des êtres humains.

On appellera **langage** la **faculté** humaine de communiquer des idées au moyen de la langue.

Le langage, par opposition à la langue, est donc intimement lié aux aspects sociologiques, psychologiques, physiologiques et même physiques de l'utilisation de la langue.

Il faut se rappeler qu'il existe de nombreuses langues, aux alentours de 6 000 (français, anglais, allemand, russe, espagnol, mandarin, japonais, swahili, etc.), mais que l'on va parler du langage comme d'une faculté générale possédée par les humains : la faculté d'apprendre et d'utiliser des langues données.

Pour résumer, nous voyons que la notion de langue, le code linguistique lui-même, nous amène à considérer les notions

- de parole — l'actualisation de la langue dans des actes de communication impliquant un locuteur et un destinataire,

que les signes constituant la langue sont *grosso modo* les mots de la langue et, de façon provisoire, on appellera **lexique** d'une langue donnée l'ensemble des mots de cette langue.

Les règles générales qui permettent de combiner les mots de la langue pour former des phrases constituent ce que l'on appelle la **grammaire** de la langue.

Chaque langue est donc avant tout constituée d'un lexique et d'une grammaire. Apprendre une langue consiste à assimiler ces deux ensembles de connaissances et à développer les automatismes permettant de les utiliser de façon spontanée.

Niveaux de fonctionnement des langues

On reconnaît habituellement au moins quatre niveaux principaux de fonctionnement dans toutes les langues :

- 1 la **sémantique**, qui concerne les sens et leur organisation au sein des messages que l'on peut exprimer dans cette langue ;
- 2 la **syntaxe**, qui concerne la structure des phrases ;
- 3 la **morphologie**, qui concerne la structure des mots ;
- 4 la **phonétique**, qui concerne les éléments sonores qui sont la forme même des énoncés.

À chacun de ces niveaux de fonctionnement, correspond une sous-discipline de la linguistique, qui s'attache plus particulièrement à l'étude et à la description du niveau en question. Ainsi, la **sémantique** linguistique est l'étude et la description de la **sémantique des langues**². Comme nous le verrons en progressant à travers les chapitres de ce livre, l'ensemble des sens véhiculés par une langue donnée est en grande partie en correspondance avec l'ensemble des mots de cette langue. Il existe donc un lien privilégié entre l'étude **sémantique** et l'étude du **lexique** —la **lexicologie**. C'est la raison pour laquelle ce livre introduit simultanément les notions de base de ces deux sous-disciplines de la linguistique, même s'il est conçu avant tout comme un cours de lexicologie.

2. Je reviendrai sur cette distinction à la fin de ce livre, dans le Chapitre 10.

Difficulté de l'étude linguistique

Tout lecteur de cet ouvrage peut légitimement penser qu'il sait ce qu'est la langue puisque il parle, écrit et lit au moins l'une d'elles : le français. Cependant, une chose est de maîtriser un ensemble de connaissances, une autre est de comprendre de façon consciente comment cet ensemble de connaissances est organisé et comment il fonctionne. Certaines tâches, comme l'enseignement de la langue, la traduction, l'élaboration et la correction de documents, demandent que l'on soit capable non seulement de parler une ou plusieurs langues, mais aussi de raisonner sur elles. C'est ce qui donne à la linguistique, en tant que science visant l'étude des langues, son intérêt et sa raison d'être.

L'étude linguistique est une activité scientifique particulièrement délicate car il n'existe pas d'autre façon de décrire les langues, de parler d'elles, que de le faire au moyen d'une langue. On est ainsi confronté à une dangereuse circularité : on se sert de notre objet d'étude pour parler de celui-ci. Il faut donc en linguistique, plus que dans toute autre sciences, se doter d'une terminologie et de conventions d'écriture très bien définies, et les utiliser de façon rigoureuse. Nous reviendrons sur ce point dans le Chapitre 3 (Section *Remarques sur la terminologie et les conventions d'écriture*, page 44).

Limitations du présent ouvrage

Ce livre se situe dans le cadre d'une linguistique qui se limite à l'étude de la langue, par opposition à une linguistique du langage en général.

J'ai aussi mentionné plus haut le fait que la langue est en constante transformation : elle évolue dans le temps. On peut donc l'étudier de deux façons :

- 1 dans le contexte de son évolution — ce que l'on appelle l'***étude diachronique*** ;
- 2 à un moment donné de son évolution, notamment telle qu'elle est utilisée actuellement — ce que l'on appelle l'***étude synchronique***.

Dans ce livre, nous nous situerons dans le cadre d'une étude synchronique de la langue. Ainsi, mes exemples étant la plupart du temps

empruntés au français, nous nous concentrerons sur le français contemporain et nous parlerons très peu des problèmes liés à l'évolution du lexique des langues (apparition ou disparition de mots, évolution des significations exprimées par les mots de la langue, etc.).

Lectures complémentaires



Les lectures données à la fin de chaque chapitre sont toujours présentées par ordre de priorité décroissante. Il s'agit bien entendu d'une priorité établie vis-à-vis du seul contenu du chapitre. Il n'est pas question ici d'un jugement de valeur sur l'importance ou la qualité intrinsèque de ces textes.

Saussure, Ferdinand de (1972) Introduction, Chapitres III, IV et V. In : *Cours de linguistique générale*, Paris : Payot, pp. 23-43.

Le *Cours de linguistique générale* de F. de Saussure est un des ouvrages fondateurs de la linguistique moderne. Il date un peu, par certains aspects, mais reste un texte de référence incontournable à cause des notions fondamentales qu'il introduit.

Rey-Debove, Josette (1998) 1. À la recherche de la distinction oral/écrit. In : *La linguistique du signe : Une approche sémiotique du langage*, Chapitre I, Collection U, série « Linguistique », Paris : S.E.S.J.M./Armand Colin, pp. 10-19.

Ce texte est particulièrement intéressant pour nous puisqu'il présente une réflexion sur la distinction langue parlée vs langue écrite selon un point de vue lexicologique. J. Rey-Debove est notamment la co-directrice de l'équipe de rédaction du dictionnaire *Petit Robert*.

Exercices

Je ne propose que trois exercices de « mise en jambes » pour ce premier chapitre. Le travail sur les données linguistiques va devenir de plus en plus important au fur et à mesure que nous progresserons dans l'ouvrage



Chapitre 2 Le signe linguistique

Signe, signe intentionnel vs non intentionnel, sémiotique / sémiologie, icône, symbole, indice, signe linguistique, signifié, signifiant, mot, signe lexical vs grammatical, signe élémentaire vs complexe, (propriété de) combinatoire, expression grammaticale vs agrammaticale.

“All right – can you explain, in 200 words or less, the difference between a sign and a symbol?”

Sumire lifted the napkin from her lap, lightly dabbed at her mouth, and put it back. What was the woman driving at? “A sign and a symbol?”

“No special significance. It’s just an example.”

Again Sumire shook her head. “I have no idea.”

Haruki Murakami, *Sputnik Sweetheart*

Il a été mentionné dans le chapitre précédent qu’une langue était un système de signes et de règles (voir notamment page 4, la définition de *langue*). La notion de signe joue un rôle central dans la façon dont nous aborderons la lexicologie et la sémantique. Nous allons donc voir, sans plus attendre, comment se caractérisent les signes linguistiques, en commençant par examiner ce qu’il faut entendre par *signe* au sens large.

Signe et sémiotique

Définition du signe

Je propose d'adopter la définition suivante de la notion de signe :

Un **signe**, au sens large, est une association entre une idée (le contenu du signe) et une forme — les trois termes *association*, *idée* (ou *contenu*) et *forme* étant pris dans leur acception la plus générale possible.

Par exemple, un clin d'œil est un signe dans la mesure où il sert à véhiculer une idée donnée, que l'on pourrait définir de façon très vague comme étant la manifestation d'une forme de connivence entre celui qui fait le clin d'œil et celui à qui il est destiné. Pour coller de plus près à notre définition, nous dirons que le signe clin d'œil est une association entre une idée — l'expression de la connivence — et une forme — la déformation complexe du visage que l'on appelle *clin d'œil* en français. Un signe de ce type peut être appelé **signe intentionnel**, puisqu'il est utilisé consciemment par un individu pour communiquer quelque chose.

La définition ci-dessus nous autorise aussi à appeler *signe* des associations idée-forme qui ne sont pas des outils de communication mais des phénomènes se manifestant naturellement. Par exemple, un certain type de vague sur la mer pourra être interprété par un marin expérimenté comme l'indication qu'un vent violent va bientôt se lever. Cette association entre une forme (type particulier de vague) et une idée (risque de vent violent) est un **signe non intentionnel**.

Noter que, dans ce dernier cas, il semble plus naturel de parler d'une association forme-idée, plutôt qu'idée-forme, dans la mesure où un type de vague n'est pas fait pour exprimer une idée donnée. C'est nous, en tant qu'observateurs, qui l'interprétons ainsi. Il semble donc toujours plus naturel de considérer les signes non intentionnels comme des associations forme-idée et les signes intentionnels — qui sont des outils de communication — comme des associations idée-forme.



La remarque ci-dessus met bien en évidence le fait que les signes intentionnels et les signes non intentionnels sont de nature très différente

puisque les premiers sont de véritables outils de communication alors que les seconds se manifestent naturellement et ne sont qu'interprétés. La raison pour laquelle ces deux types d'entités sont fréquemment regroupés est qu'à un niveau très profond, celui du fonctionnement de notre système cognitif, ils se rejoignent en tant qu'associations binaires entre idée et forme.

Il n'est pas exagéré de dire que notre existence, en tant qu'êtres humains, est presque entièrement consacrée à produire, recevoir et analyser des signes, ainsi qu'à réagir en conséquence. D'où l'importance de définir une discipline dédiée à l'étude des signes.

La science des signes

La **sémiotique** est la science qui étudie les différents systèmes de signes — la linguistique étant une branche de la sémiotique générale qui étudie les systèmes de signes linguistiques que sont les langues.

On emploie aussi le terme de **sémiologie** pour désigner cette science. Pour certaines personnes, le terme *sémiotique* est un anglicisme qu'il faut absolument remplacer par *sémiologie*. D'autres affirmeront qu'il existe une nuance entre la sémiotique, discipline d'origine nord-américaine qui s'attache à l'étude de tous les types de signes, et la sémiologie, d'origine européenne, qui ne reconnaît comme objet d'étude que les signes intentionnels. D'autres enfin diront que ce sont deux façons acceptables de désigner exactement la même chose. Étant par nature enclin à choisir la solution la plus économique à moins de voir clairement la nécessité du contraire, j'adopterai la troisième voie et utiliserai ici le terme *sémiotique* : il apparaît au moins aussi fréquemment que *sémiologie* dans les textes de référence en français et... il présente l'avantage d'avoir une syllabe de moins que son concurrent¹. Le texte de J.-M. Klinkenberg proposé comme lecture à la fin du présent chapitre fait le point la question.

1. Si l'on se fait reprocher d'utiliser un anglicisme avec le terme *sémiotique*, on peut sournoisement répondre qu'il est attesté dans le *Petit Robert*. Il ne restera plus alors à notre interlocuteur que la solution de critiquer le dictionnaire en question, ce qui s'avère toujours être un exercice douloureux pour un puriste.

Comme il a été dit plus haut, toutes les approches sémiotiques ne se donnent pas les mêmes objets d'étude. Ainsi, certains sémioticiens excluent du champ d'étude de la sémiotique les signes non intentionnels. Pour les linguistes, cependant, le problème ne se pose pas puisqu'ils n'ont à étudier que des signes intentionnels : les signes de la langue ne se manifestent pas naturellement mais sont émis par un locuteur afin de communiquer.

Types de rapports contenu-forme dans les signes

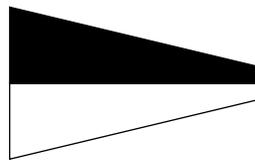
Outre la distinction entre signes intentionnels et non intentionnels, on peut classer les signes en fonction du type de rapport existant entre leur contenu et leur forme. Cela nous donne trois grandes familles de signes :

- 1 Un *icône*² est un signe tel qu'il existe un « lien de ressemblance » entre l'idée qu'il véhicule et la forme qui lui est associée. Ainsi, le dessin (1a) ci-dessous, apposé sur la porte des toilettes d'un restaurant, est iconique dans la mesure où le contenu qu'il véhicule alors ('toilettes pour femmes') est évoqué dans sa forme (dessin d'une silhouette féminine). Au contraire, le pavillon maritime (1b), qui signifie 'six', ne possède rien dans sa forme qui puisse évoquer son contenu et ne sera donc pas considéré comme étant un icône :

(1) a.



b.



2. Ce terme technique est souvent employé au masculin (*un icône*) pour le distinguer de l'icône qui est une image religieuse (*une icône*, nécessairement au féminin).

Pourtant, elle est mise en scène volontairement par quelqu'un pour transmettre une information (même si cette dernière est fausse). La distinction entre signe intentionnel et non intentionnel ne recoupe donc qu'en partie une opposition icône-symbole vs indice.

Distinguer entre icônes, symboles et indices n'est qu'une des classifications et des terminologies possibles. Le texte d'O. Ducrot et J.-M. Schaeffer suggéré comme lecture à la fin de ce chapitre permet d'élargir la perspective à ce propos.

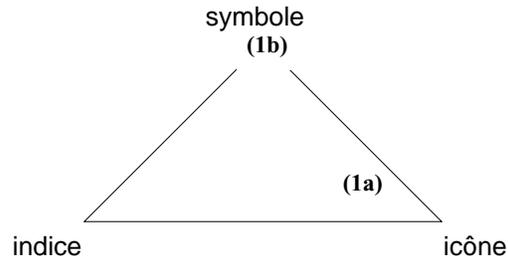


Il faut interpréter la classification des signes donnée ci-dessus comme identifiant des **tendances**. Un signe est rarement purement iconique, symbolique ou indiciel. Par exemple, serrer la main, pour « officialiser » une rencontre, est iconique dans la mesure où il y a un lien évident entre le fait de saisir la main de quelqu'un — de se lier physiquement à lui — et celui d'établir une rencontre. Mais ce signe est en même temps symbolique, car en partie arbitraire. On pourrait aussi bien se frotter le nez, se mordre l'oreille ou faire des choses encore plus étonnantes encore. Même le signe (1a), qui indique des toilettes pour femmes, n'est pas purement iconique. Il est aussi symbolique car c'est par convention que l'on considère que la silhouette en question identifie la femme par opposition à l'homme. Il est aussi en partie indiciel dans la mesure où il ne prend sa signification véritable ('réservé aux femmes') qu'une fois apposé sur une porte de toilettes⁴.

Au lieu de classer de façon rigide les signes en tant que symbole, indice ou icône, on pourrait par exemple visualiser la distinction qui a été opérée au moyen d'un triangle dont chaque angle représenterait un type « pur » de signe, chaque autre point géométrique indiquant un type

4. Cette caractérisation du signe (1a) comme partiellement indiciel pourrait être débattue. Ce que je veux mettre en évidence ici, c'est que ce signe pris en isolation n'est que la représentation schématique d'une personne de sexe féminin. Il doit être **physiquement associé** à un lieu déterminé pour indiquer que ce lieu est réservé aux femmes (et donc interdit aux hommes).

intermédiaire. Par exemple, la différence de nature entre (1a) et (1b) pourrait être visualisée de la façon suivante :



Nous verrons plus bas que cette remarque sur la nature hybride des signes peut s'appliquer aussi, dans une certaine mesure, aux signes linguistiques.

Le signe linguistique selon Ferdinand de Saussure

Le *signe linguistique*, tel que défini par F. de Saussure, se distingue des autres signes par au moins cinq propriétés, que nous allons maintenant examiner à tour de rôle.

Les deux composantes du signe linguistique

Le signe linguistique est fait de l'association entre un *signifié* — une « idée » — et un *signifiant*, qui est une « image acoustique ».

Le terme *image acoustique* peut poser problème. Ce que Saussure veut mettre en évidence ici c'est que le signifiant du signe n'est pas une suite de sons, puisqu'on ne retrouve les sons que dans la parole (la manifestation concrète de la langue). Le signifiant est plutôt un patron sonore abstrait, qui est stocké dans la mémoire du locuteur, et que ce dernier pourra utiliser soit pour émettre (concrétiser) le signe en question, soit pour identifier un signe dont il est le récepteur.

Il est très important de bien comprendre que le signe, en tant qu'élément de la langue, est une entité entièrement psychique, qui réside dans notre cerveau. Deux personnes distinctes ne vont jamais prononcer le même mot de la même façon, mais on saura que c'est le même mot qu'elles ont utilisé parce que dans chacun des cas la suite de sons produite pourra

être associée à un même patron sonore abstrait, à une même image acoustique.



On utilisera la convention d'écriture suivante pour bien distinguer les deux composantes du signe linguistique :

- 1 guillemets simples pour le signifié du signe, par ex. 'cheval' ;
- 2 italique pour le signifiant du signe, par ex. : *cheval*.

Attention ! Pour être parfaitement rigoureux, il nous faudrait aussi utiliser une convention d'écriture particulière pour nommer le signe dans sa totalité, en tant qu'entité linguistique. On se contentera cependant de désigner le signe de la même façon que l'on désigne son signifiant (en italique), par ex. : le signe *cheval*.

Caractère arbitraire du signe linguistique

Le signe linguistique est arbitraire, en ce sens que l'association entre le signifié et le signifiant n'est pas logiquement motivée. Cela revient à caractériser le signe linguistique en tant que symbole.

On peut noter cependant que certains mots de la langue sont en partie iconiques. Il s'agit ici des onomatopées, comme le nom français *cocorico*. Il existe un lien de ressemblance évident entre la forme linguistique /kɔkɔriko/ et le contenu qu'elle exprime. Malgré cela, *cocorico* reste un mot du français, en partie symbolique. En anglais, par exemple, on désigne le cri du coq par *cock-a-doodle-do* et il est fort douteux que les coqs des pays francophones poussent un cri différent de celui des coqs des pays anglophones. Le fait qu'un mot, même onomatopéique, doive être traduit d'une langue à l'autre est une preuve de son caractère partiellement arbitraire (voir ci-dessus, page 21, à la toute fin de la section sur les différents types de signes).

Caractère figé du signe linguistique

Le signe nous est imposé par le code social qu'est la langue et est donc figé. C'est ce que Saussure appelle l'*immutabilité du signe*. Le nom français *maison* existait déjà en français il y a cent ans et existera encore vraisemblablement dans cent ans. C'est grâce à cette stabilité du système linguistique que l'on peut apprendre les langues, les utiliser tout

au cours de notre existence et transmettre des informations à travers le temps.

Si chaque individu avait la liberté de créer de nouveaux signes linguistiques, ou de faire évoluer les signes de la langue selon son bon vouloir, il n'y aurait plus de fonctionnement social possible pour la langue. Une langue, c'est un peu comme un code de la route que tout le monde adopte et suit afin que la plus grande harmonie règne entre les humains quand ils communiquent, comme elle règne sur les routes quand ils conduisent.

Caractère évolutif du signe linguistique

Oui, il y a une « légère » nuance d'ironie dans ce que je viens de dire, car on sait bien que les règles sont sans cesse transgressées, sur la route comme dans la communication linguistique, et que, de plus, elles sont faites pour être changées.

Les signes d'une langue, comme les lois ou les règlements, sont ainsi sujets à deux types de variations :

- 1 des variations individuelles, dans la mesure où tout le monde n'applique pas nécessairement ces lois de la même façon, ou, même, n'a pas nécessairement appris exactement le même système de lois ;
- 2 des variations dans le temps, dans la mesure où les lois, les systèmes de signes linguistiques, évoluent dans le temps.

Bien entendu, les variations linguistiques se manifestent la plupart du temps sans que les locuteurs le désirent de façon consciente. Un mot de la langue va généralement mourir très progressivement de sa belle mort, tout simplement parce que de moins en moins de personnes vont l'utiliser. Tout se passe de façon graduelle, insensible et non volontaire, mais le fait demeure : les signes de la langue évoluent, ce qui est en apparence contradiction avec la caractéristique précédemment mentionnée. C'est ce que Saussure appelle la *mutabilité du signe*.

Caractère linéaire du signe linguistique

Le signifiant du signe linguistique est linéaire, du fait de la nature orale de la langue et de la physiologie humaine. En effet, nous ne sommes capables de produire aisément avec notre système phonatoire qu'un son

à la fois : la réalisation d'un message linguistique est donc une suite linéaire de sons. Au niveau du signe, les signifiants sont des patrons sonores linéaires. Bien entendu, la contrainte de linéarité ne s'appliquerait pas avec un système sémiotique gestuel, puisque l'on peut tout à fait produire et identifier plusieurs gestes de façon simultanée.

Pour conclure, on peut noter une propriété importante de la **langue**, en tant que système de signes : l'ensemble des signes de chaque langue et l'ensemble des règles de combinaison de ces signes sont synchroniquement finis. C'est ce qui rend possible une description relativement complète du lexique et de la grammaire des langues (dictionnaires et traités de grammaire).

Types de signes linguistiques

Signe lexical vs grammatical

Les signes linguistiques sont-ils uniquement des *mots* ? J'ai souvent employé, dans ce qui précède, le terme *mot* en parlant de signes linguistiques. Le problème est que ce terme est très ambigu et, comme nous le verrons dans le chapitre suivant, nous éviterons de l'employer pour lui préférer des termes techniques plus précis, désignant chacun une entité linguistique clairement identifiée.

On peut cependant se permettre de dire pour l'instant que des mots comme *boire*, *dormir*, *chemin*, *maison*, etc. sont des signes linguistiques. Cependant, si tous les mots sont des signes linguistiques, tous les signes linguistiques ne sont pas des mots. On doit ainsi au moins distinguer les **signes lexicaux**, comme ceux qui viennent d'être mentionnés, et les **signes grammaticaux**, dont voici deux exemples :

- 1 Le suffixe du pluriel des noms français *-s* est un signe linguistique :
 - il a un signifiant — qui ne se manifeste cependant à l'oral que dans le phénomène de liaison, *des amis épatants* se prononçant /dezamizepatã/ ;
 - il a un signifié — le sens de pluralité.
- 2 Le préfixe français *re-*, qui se combine aux verbes pour exprimer le sens de répétition : *refaire*, *rediscuter*, *revisiter*, etc.

Nous aurons l'occasion de revenir sur les signes linguistiques non lexicaux dans le Chapitre 4, qui introduit les notions de base de morphologie dont la connaissance est nécessaire dans le cadre de ce cours.

Signe élémentaire vs complexe

La distinction entre signe lexical et signe grammatical nous permet de mettre en évidence une autre opposition :

- 1 Les **signes élémentaires** sont des signes qui ne peuvent être analysés en terme de signes plus simples dont ils seraient constitués. Par exemple, la préposition *avec* est un signe linguistique élémentaire.
- 2 À l'opposé, les **signes complexes** sont analysables en terme d'autres signes. Par exemple, *refaire* peut être analysé comme la combinaison des deux signes *re-* + *faire*.

Il faut souligner qu'un signe linguistique complexe n'est pas nécessairement constitué d'un signe lexical et de un ou plusieurs signes grammaticaux. *Fruit de mer*, par exemple, est formellement analysable en terme de trois signes lexicaux : *fruit* + *de* + *mer*. On remarquera que je présuppose ici que *fruit de mer* est un signe linguistique. En fait, toute expression linguistique de ce type — *pomme de discorde*, *conter fleurette*, *à propos de*, etc. — est un signe linguistique, puisqu'elle possède les caractéristiques du signe identifiées plus haut.

Ce qui peut poser problème dans le cas de *fruit de mer*, c'est la nature **lexicale** de ce signe. Ce point sera examiné en détail dans le prochain chapitre.

Propriétés de combinatoire des signes linguistiques

Le signe linguistique peut être conceptualisé comme comportant une composante additionnelle : ses propriétés de combinatoire.

Attention ! Cet aspect essentiel de la caractérisation du signe linguistique n'est pas présenté dans la plupart des textes d'introduction, mais c'est une extension du signe saussurien (proposée dans le cadre de la théorie linguistique Sens-Texte⁵ sous le nom de *syntactique* du signe

linguistique) qui va nous permettre de rendre compte de façon propre de nombreux phénomènes linguistiques.

Les **propriétés de combinatoire** d'un signe linguistique sont les contraintes propres à ce signe qui limitent sa capacité de se combiner avec d'autres signes linguistiques et qui ne peuvent être déduites ni de son signifié ni de son signifiant.

Par exemple, le signe français *sommeil* est l'association entre un signifié et un signifiant, mais il se caractérise aussi par de multiples propriétés de combinatoire. Je n'en citerai ici que quatre :

- 1 C'est un nom commun et, en conséquence, il doit normalement s'employer avec un déterminant (article, pronom possessif, etc.) :
(3) *un/le/son/... sommeil*
- 2 Il est masculin et conditionne donc la forme masculine de ses déterminants et des adjectifs qui le modifient :
(4) *un sommeil étonnant* vs **une sommeil étonnante*



Usage du symbole « * » : En linguistique, on est fréquemment obligé d'établir des comparaisons entre des expressions bien formées, qui respectent les règles linguistiques, et des expressions qui transgressent certaines de ces règles. Les premières sont appelées **expressions grammaticales** et les secondes **expressions agrammaticales**. On utilise couramment l'astérisque (*) devant une expression pour indiquer que celle-ci est agrammaticale. Cela nous permet de donner des exemples de constructions incorrectes pour illustrer notre propos, tout en s'assurant que le lecteur interprétera de la bonne façon la portée de nos exemples.

- 3 Ce signe est la forme du singulier d'un nom français, le signe pluriel correspondant étant *sommeils*.

5. Nous aurons l'occasion de revenir assez fréquemment sur certaines notions importantes introduites dans le cadre de la théorie Sens-Texte (voir *théorie Sens-Texte* dans l'Index des notions, à la fin de l'ouvrage).

- 4 On peut le combiner avec les adjectifs *lourd* et *profond* pour exprimer l'intensification de son sens et avec le verbe *tomber* [*dans*] pour exprimer le sens 'commencer à être dans un état de sommeil', d'autres combinaisons « logiquement » possibles étant ici exclues :

- (5) a. *un lourd/profond sommeil*
b. **un pesant/grand sommeil*
c. *Léo tomba dans un profond sommeil.*
d. **Léo dégringola dans un profond sommeil.*

Toutes ces caractéristiques de combinatoire sembleront peut-être évidentes au lecteur et il se demandera peut-être en quoi elles sont véritablement des caractéristiques du signe *sommeil*. L'Exercice 7 ci-dessous éclaircira ce point.

Ici se conclut cet important chapitre sur le signe linguistique. La maîtrise de toutes les notions qui viennent d'être introduites est absolument essentielle pour bien appréhender le contenu des chapitres qui s'en viennent. Équipés de ces notions, nous pouvons maintenant aborder la caractérisation de l'objet d'étude de la lexicologie : la lexie.

Lectures complémentaires

Saussure, Ferdinand de (1972) Introduction, Chapitres III, IV et V. In : *Cours de linguistique générale*, Paris : Payot, pp. 32-35.

Il est utile de revenir sur ces quelques pages de la lecture donnée au Chapitre 1, maintenant que la science des signes (la sémiotique, ou sémiologie) a été introduite. On remarquera que Saussure parle de la sémiotique comme d'une discipline à construire, mais la situation a bien évolué depuis lors.

Saussure, Ferdinand de (1972) Première Partie, Chapitres I et II. In : *Cours de linguistique générale*. Paris : Payot, pp. 97-113.

Cette lecture contient la célèbre caractérisation du signe linguistique proposée par Saussure dans son *Cours de linguistique générale*.

Klinkenberg, Jean-Marie (1996) Chapitre I, Section 2. Sémiologie ou sémiotique ? In : *Précis de sémiotique générale*. Collection

« Points Essais », n° 411, Bruxelles : De Boeck Université, pp. 21-27.

À lire pour une discussion claire et condensée de l'usage des deux termes *sémiologie* et *sémiotique*. Le point de vue terminologique adopté par l'auteur est en accord avec mon propre usage du terme *sémiotique* (voir Section *La science des signes* page 17).

Ducrot, Oswald et Jean-Marie Schaeffer (1995) *Signe*. In : *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris : Éditions du Seuil, pp. 213-223.

On trouvera dans ce texte une classification des types de signes plus élaborée que celle introduite dans le présent chapitre. Ce texte situe la notion de signe dans une perspective plus large : l'étude des systèmes sémiotiques, dont la langue est un cas particulier. Il n'est pas nécessaire que le lecteur comprenne ou parvienne à assimiler tout ce qui est dit dans ce texte, dans la mesure où il contient de nombreuses références à des notions liées à celle de signe. Les termes apparaissant en gras dans ce texte possèdent eux-mêmes une entrée l'ouvrage de Ducrot et Schaeffer. Je recommande l'achat de ce dictionnaire aux étudiants en linguistique car c'est un ouvrage très utile et assez bon marché. Attention : il s'agit ici du **Nouveau dictionnaire encyclopédique des sciences du langage** (paru en 1995).

Benveniste, Émile (1966) Chapitre V. Communication animale et langage humain. In : *Problèmes de linguistique générale*, Paris : Gallimard, pp. 56-62.

Ce texte est intéressant à lire pour mieux saisir la notion de système sémiotique en général.

Exercices

- 1 En quoi les mains qui tremblent quand on a froid, les bandes blanches que l'on trouve à un passage pour piétons et le « V » de la victoire sont-ils des signes de nature différente ?

Chapitre 3

L'unité lexicale ou *lexie*

Lexicologie, lexie (= unité lexicale), mot, mot-forme, lexème, flexion, locution nominale / verbale / adjectivale / adverbiale / prépositionnelle, (principe de) compositionnalité sémantique, vocable, acception, polysémie, homonymie.

Il est utile d'avoir examiné le nombre des acceptions d'un terme, tant pour la clarté de la discussion (car on peut mieux connaître ce qu'on soutient, une fois qu'a été mise en lumière la diversité de ses significations), qu'en vue de nous assurer que nos raisonnements s'appliquent à la chose elle-même et non pas seulement à son nom. Faute, en effet, de voir clairement en combien de sens un terme se prend, il peut se faire que celui qui répond, comme celui qui interroge, ne dirigent pas leur esprit vers la même chose. Au contraire, une fois qu'on a mis en lumière les différents sens d'un terme et qu'on sait sur lequel d'entre eux l'interlocuteur dirige son esprit en posant son assertion, celui qui interroge paraîtrait ridicule de ne pas appliquer son argument à ce sens-là.

Aristote, *Les Topiques*, I, 18.

Il a été mentionné au début du Chapitre 1 (page 3) que, cet ouvrage traitant de l'étude du lexique, la discipline centrale dont il relève est la lexicologie. Voici maintenant comment nous définirons cette discipline :

La **lexicologie** est une branche de la linguistique qui étudie les propriétés des unités lexicales de la langue, appelées **lexies**.

J'attire l'attention du lecteur sur le fait que je n'ai pas employé *mot* pour désigner l'objet d'étude de la lexicologie, préférant faire usage d'un terme aux consonances plus « techniques » : *lexie*. Nous verrons les raisons de ce choix terminologique dans la section suivante.

La *lexie* fait l'objet de tout un chapitre, comme cela a été le cas pour le signe linguistique. En effet, la notion de *lexie* ne se laisse pas facilement isoler et, pour la définir, il faut en fait mettre à jour et expliciter tout un réseau de notions intimement liées. Un peu comme si l'on tirait sur le fil d'une pelote de laine qui va finir par se dérouler entièrement, la nécessité de définir une notion implique presque nécessairement, dans le cas d'une science descriptive comme la linguistique, celle de définir un champ notionnel où tout se tient. Il est même difficile, en linguistique, de s'enfermer complètement à l'intérieur d'une branche donnée de cette discipline. La lexicologie, en particulier, n'est pas un domaine d'étude qu'on peut véritablement circonscrire. Pour bien faire de la lexicologie, pour bien aborder l'étude du lexique, il faut définir des notions de base de sémantique, bien entendu, mais aussi des notions de morphologie, de syntaxe et de phonologie. Les incursions dans le domaine de la phonologie seront extrêmement rares et ponctuelles dans les chapitres qui viennent. De plus, les notions phonologiques de base tendent à être admises, à un niveau élémentaire, de façon relativement consensuelle ; je vais donc les présupposer connues et ne pas m'attarder à les définir. Il en va tout autrement des notions de sémantique, de morphologie et (plus marginalement) de syntaxe, qu'il nous faudra souvent prendre à bras le corps en les définissant explicitement.

Ce chapitre commence par une justification du rejet du terme *mot*, en tant que notion lexicologique, et par l'introduction des deux notions centrales de mot-forme et lexème. Nous examinerons ensuite le cas des locutions. C'est alors seulement que nous serons en mesure de définir la *lexie* de façon relativement précise. Les autres notions qu'il nous faudra ensuite introduire avant de terminer ce chapitre sur la *lexie* sont celles de vocable, de polysémie et d'homonymie. Bref, je n'ai pas menti au lecteur en parlant de dérouler toute une pelote de laine...

Mot, mot-forme et lexème

Le mot n'est pas une notion linguistique

Le terme *mot* est ambigu, comme le montrent les exemples ci-dessous :

- (1) a. *Sa réponse tient en deux mots* : « *Sûrement pas !* ».
b. « *Suis* », « *es* », « *est* », « *sommes* », etc. sont toutes des formes du même mot.

De façon très naïve et approximative on pourrait dire que, dans la phrase (1a), *mot* est employé pour désigner des formes linguistiques qui sont séparées à l'écrit par des espaces ou des marques de ponctuation. Dans (1b), en revanche, il est explicitement dit qu'un mot est quelque chose de plus « abstrait », plus général qu'une forme linguistique. On retrouve la même distinction lorsque l'on dit par exemple

- *parce que* s'écrit en **deux** mots — sens de (1a) ;

vs

- *parce que* est **un** mot, qui se traduit en anglais par *because* — sens de (1b).

Tout cela démontre que le terme *mot* est utilisé pour désigner deux notions distinctes liées à la langue. De nombreux textes de linguistique entretiennent malheureusement la confusion, en employant ce terme indifféremment dans l'un ou l'autre sens. À cela s'ajoute le fait que *mot* peut signifier encore bien d'autres choses, dans la langue de tous les jours ou en tant que terme technique :

- (2) a. *Je vais te rapporter quelques bons mots* [= 'réflexion ou remarque intéressante'] *qu'il nous a sortis*.
b. *Il a glissé un mot* [= 'court message écrit'] *sous la porte de Germaine*.
c. *En informatique, un octet est un mot* [= 'séquence d'informations élémentaires'] *de huit bits*.



Pour éviter toute confusion, je n'utiliserai jamais *mot* en tant que terme linguistique technique. Je vais introduire deux termes bien distincts, un

pour chacune des deux notions dont il est question dans les phrases (1a) et (1b) ci-dessus.

Le mot-forme

La notion correspondant à l'exemple (1a) sera désignée par le terme *mot-forme*¹ ; elle peut se définir de la façon suivante :

Un ***mot-forme*** est un signe linguistique

- qui possède une certaine autonomie de fonctionnement
- et qui possède une certaine cohésion interne.

Pour que la définition ci-dessus soit compréhensible, il faut préciser ce que l'on entend par *autonomie de fonctionnement* et *cohésion interne*. Je le ferai en examinant la phrase ci-dessous, qui contient quatre mots-formes :

(3) *Le chemin est encombré.*

L'autonomie de fonctionnement peut être testée de multiples façons ; en voici trois.

- 1 Il est possible de remplacer chacun des quatre mots-formes de (3) par d'autres mots-formes pouvant avoir la même fonction grammaticale dans la phrase, comme l'illustre le tableau ci-dessous :

Position 1	Position 2	Position 3	Position 4
<i>Le</i>	<i>chemin</i>	<i>est</i>	<i>encombré</i>
<i>Ce</i>	<i>passage</i>	<i>sera</i>	<i>libre</i>
<i>Un</i>	<i>couloir</i>	<i>devenait</i>	<i>bizarre</i>

Patron de phrase à quatre positions avec trois alternatives pour chaque position

1. À propos de ce terme, voir le commentaire qui accompagne le texte Mel'čuk (1993) dans la liste des lectures complémentaires pour ce chapitre, page 45.

Ce tableau contient quatre colonnes de mots-formes, correspondant chacune à la position linéaire d'un des mots-formes de (3). Chaque colonne contient trois rangées de mots-formes ayant une même « valeur fonctionnelle », c'est-à-dire pouvant avoir la même fonction grammaticale dans la phrase (déterminant de nom, sujet du verbe, verbe principal ou adjectif attribut du sujet). Il est possible de faire une phrase correcte en employant n'importe lequel des mots-formes de la première colonne suivi de n'importe lequel des mots-formes de chacune des colonnes consécutives. En voici trois exemples choisis au hasard parmi les 81 possibles (= 3^4) :

- (4) a. *Ce chemin sera encombré.*
 b. *Un passage devenait libre.*
 c. *Ce couloir sera bizarre.*

On voit que chacun des mots-formes énumérés dans le tableau ci-dessus, et par voie de conséquence chacun des mots-formes de (3), est relativement autonome puisqu'il n'a besoin d'aucun autre mot-forme **en particulier** pour fonctionner dans une phrase. Ainsi, le mot-forme *ce* n'a pas besoin d'être suivi par *chemin* dans *Ce _ est encombré* pour que la phrase soit acceptable. Il suffit que la position en question soit occupée par un mot-forme ayant les mêmes caractéristiques fonctionnelles que *chemin* (*passage, couloir, etc.*).

- 2 On peut employer chaque mot-forme de (3) dans d'autres contextes que ceux de la phrase initiale :

- (5) a. *Il regarde **le** chien.*
 b. *C'est un **chemin** ombragé.*
 c. *Je pense qu'il **est** fragile.*
 d. *Je trouve ton bureau bien **encombré** de choses inutiles.*

- 3 Les mots-formes apparaissant dans une phrase sont séparables les uns des autres, par insertion d'autres mots-formes. Ainsi, on peut insérer des mots-formes entre chaque paire de mots-formes de (3) — comme le montre (6) ci-dessous :

- (6) *Le petit chemin ombragé est bien encombré aujourd'hui.*

La cohésion interne des mots-formes de (3), quant à elle, se manifeste justement dans le fait qu'une insertion de nouveaux mots-formes à **l'intérieur des mots-formes eux-mêmes** est impossible. On ne peut construire de phrases comme :

(7) **Le cheombragémin est encombienbré.*²

Finalement, les critères d'identification des mots-formes peuvent varier d'une langue à l'autre. Il est notamment possible pour certaines langues de faire usage de critères phoniques. Par exemple, en mandarin, la grande majorité des mots-formes sont formés de deux syllabes. Ces critères ne sont cependant jamais suffisants pour identifier les mots-formes d'une langue et ils sont pratiquement inexistantes dans le cas du français.

Le lexème

Examinons maintenant l'emploi qui était fait de *mot* dans la phrase (1b), que je répète ci-dessous en (8a), avec un autre exemple mettant en jeu la même utilisation de ce terme :

- (8) a. « *Suis* », « *es* », « *est* », « *sommes* », *etc.* sont toutes des formes du même **mot**.
- b. *Ce matin, José a appris deux nouveaux mots anglais.*

Il n'est pas question dans ces exemples de signes linguistiques du type mots-formes, mais bien plutôt d'unités lexicales. Une unité lexicale de ce type, que nous appellerons **lexème**, est un élément de base de la connaissance lexicale. Lorsque l'on parle d'apprendre un « nouveau mot » dans une langue étrangère, on réfère en fait à un lexème de cette langue : une entité générale qui se « matérialise » dans les phrases par un ou plusieurs mots-formes. Ainsi, DOG est un lexème de l'anglais, qui est associé aux deux mots-formes *dog* (singulier) et *dogs* (pluriel).



Les lexèmes sont toujours mentionnés dans ces notes en petites majuscules, pour bien les distinguer des mots-formes. Ces derniers sont, eux, écrits en italique, puisque ce sont des signes linguistiques³.

2. Rappel (Chapitre 2) : l'astérisque (*) est utilisé en linguistique pour indiquer l'agrammaticalité d'une expression.

Plusieurs mots-formes sont associés au même lexème s'ils ne se distinguent que par ce que l'on appelle, en morphologie, la **flexion**. Nous reviendrons sur cette notion dans le chapitre suivant. Cependant, le lecteur a sûrement déjà une bonne idée intuitive de ce que recouvre la flexion ; c'est ce qui oppose les mots-formes à l'intérieur de chacune des séries suivantes :

- (9) a. *route ~ routes*
b. *canal ~ canaux*
c. *lent ~ lente ~ lents ~ lentes*
d. *avoir ~ ai ~ as ~ a ~ avons ...*



Usage du symbole « ~ » : Le tilde (~) est fréquemment utilisé en linguistique pour séparer des expressions que l'on veut contraster ou énumérer.

Voici, pour conclure sur le lexème, une définition synthétisant ce qui vient d'être dit à propos de cette notion :

Un **lexème** est une lexie regroupant des mots-formes ne se distinguant que par la flexion.

Nous pouvons maintenant résumer brièvement le contenu de cette première section :

- 1 *Mot* est ambigu et son usage en linguistique peut prêter à confusion.
- 2 On considérera qu'il ne fait pas partie de notre terminologie linguistique.
- 3 On utilisera à la place soit le terme *mot-forme* soit le terme *lexème*, selon la notion dont il sera question.
- 4 La notion de mot-forme a reçu une définition qui sera satisfaisante pour nos présents besoins.

3. Cf. les conventions d'écriture introduites dans le Chapitre 2, page 22. Voir aussi, ci-dessous, la Section *Remarques sur la terminologie et les conventions d'écriture*, page 44.

- 5 La notion de lexème a été caractérisée comme étant un regroupement de mots-formes qui ne se distinguent que par la flexion. Un lexème est une lexie (une unité lexicale) de la langue.

Nous allons maintenant voir qu'il existe aussi des lexies qui ne sont pas des lexèmes : les locutions.

Les locutions

Définition de la notion de locution

J'ai pris soin dans la section précédente de ne pas établir une équivalence stricte entre la notion de lexème et celle d'unité lexicale, c'est-à-dire de lexie : tous les lexèmes sont des lexies, mais toutes les lexies ne sont pas des lexèmes. Faisons un petit exercice simple pour démontrer ce fait.

Il semble légitime d'admettre qu'il existe en anglais une unité lexicale POTATO ; celle-ci est bien un lexème, associé aux deux mots-formes *potato* et *potatoes*. Maintenant, quelle lexie française est la traduction de POTATO ? Il s'agit de POMME DE TERRE, bien entendu. Donc, POMME DE TERRE est une lexie. Cependant, ce n'est pas un lexème puisqu'elle est un regroupement de deux expressions linguistiques complexes — *pomme de terre* et *pommes de terre* — et non un regroupement de mots-formes. Nous sommes ici en présence d'une locution :

Une **locution** est une lexie regroupant des expressions linguistiques complexes ne se distinguant que par la flexion.

Bien entendu, les expressions associées à une locution sont formellement constituées de mots-formes : *pomme de terre* et *pommes de terre* mettent en jeu les quatre mots-formes *pomme*, *pommes*, *de* et *terre*. Cependant, aucun de ces mots-formes **pris individuellement** n'est une des réalisations de la locution POMME DE TERRE. Il est essentiel de bien garder à l'esprit que la locution se conceptualise comme un regroupement d'expressions. C'est ce qui la distingue du lexème.

Il existe plusieurs types de locutions, notamment :

- 1 les **locutions nominales** : FRUIT DE MER, NID DE POULE, ... ;

- 2 les **locutions verbales** : PASSER À TABAC, ROULER SA BOSSE, ... ;
- 3 les **locutions adjectivales** : D'ACCORD, EN PANNE, ... ;
- 4 les **locutions adverbiales** : EN VITESSE, AU HASARD, ... ;
- 5 les **locutions prépositionnelles** : À PROPOS DE, EN REGARD DE, ...

Parce que les locutions sont des lexies, des tous lexicaux, elles tendent à faire perdre aux éléments dont elles sont formellement constituées leur autonomie de fonctionnement dans la phrase. Il est ainsi souvent difficile, voire impossible, d'insérer des éléments dans une expression lorsque celle-ci correspond à une locution. Comparons les trois phrases ci-dessous :

- (10) a. *Il a mangé une salade **pourrie** de la ferme.*
- b. **Il a mangé un fruit **pourri** de mer.*
- c. *Il a mangé un fruit de mer **pourri**.*

La phrase (10a) est parfaite alors que (10b) est agrammaticale. L'adjectif *pourri* ne peut être inséré dans l'expression *fruit de mer* ; pour qualifier l'aliment en question, il faut utiliser (10c). Cela est dû au fait que *fruit de mer* correspond à une locution du français alors que l'expression *salade de la ferme*, quant à elle, n'est que la résultante de la combinaison régulière des quatre lexèmes distincts SALADE, DE, LE (au féminin singulier) et FERME.

Toutes les locutions n'ont pas la même rigidité d'emploi que FRUIT DE MER. Comparons le comportement de cette dernière avec celui de la locution verbale CASSER LES PIEDS, dont les éléments constitutifs sont plus facilement séparables :

- (11) *Il nous casse **souvent** les pieds, ce type-là.*

Malgré ces divergences, on peut dire que toutes les locutions sont des expressions qui, un peu comme des mots-formes, manifestent une autonomie de fonctionnement et un certain degré de cohésion (ce dernier variant d'une locution à l'autre).

(Non-)compositionalité sémantique

Les exemples qui viennent d'être examinés peuvent laisser penser que les mots-formes apparaissant dans des expressions comme *fruit de mer*

(ou dans le pluriel *fruits de mer*) ont perdu une partie de leurs propriétés de combinatoire (cf. dans le chapitre précédent, la combinatoire du signe linguistique). Cela est en réalité la conséquence d'un fait plus profond : ces mots-formes ont perdu une partie de leur **nature** de signe linguistique, et notamment leur sens. En effet, alors que le sens d'une expression libre comme *salade de la ferme* est la résultante de la composition des sens de chacun de ses constituants, un fruit de mer n'est pas un fruit qui pousse dans la mer. Bien entendu, on peut comprendre la métaphore mise en jeu dans *fruit de mer* : cet aliment est un peu comme un fruit (on le « récolte » pour le manger) qui pousserait dans la mer. Mais la tentative d'explication s'arrête là et nous avons ici une métaphore que la langue a figée.

On dira que les locutions transgressent, au moins en partie, le principe de **compositionalité sémantique**, qui veut qu'une expression linguistique soit directement calculable — dans sa composition lexicale et sa structure syntaxique — à partir du sens de chacun de ses constituants. Nous reviendrons sur ce point dans le Chapitre 7 (voir Section *Les collocations*, page 135).

Il est parfois difficile de percevoir la non-compositionalité sémantique des locutions appartenant à notre langue maternelle, locutions qui nous sont devenues au cours des ans tellement familières. Il est alors utile de se tourner vers des cas empruntés à d'autres langues. Voici trois exemples anglais que, volontairement, je ne traduirai pas :

- (12) a. *Let's go Dutch.*
b. *You jumped the gun, once again.*
c. *He did it with flying colors.*

Nous n'en avons pas fini avec les locutions et, de façon plus générale, avec la non-compositionalité sémantique. Loin de là ! Nous reviendrons plusieurs fois dans l'ouvrage sur cette question essentielle.

(13) *Est-ce que tu aurais un verre en verre, pas en plastique ?*

Les deux emplois de *verre* dans cette phrase ne servent pas à exprimer la même signification ; il est tout d'abord question d'un type de contenant, puis d'un type de matériau. On doit donc considérer que l'on est ici en présence de deux lexies distinctes.



Dans un cas comme celui-ci, on est amené à utiliser des numéros distinctifs, identifiant chacune des lexies en cause. On peut ainsi dire que l'on a dans (13) les deux lexies :

- VERRE1, qui désigne un matériau transparent cassable ;
- VERRE2, qui désigne un type de contenant servant à boire et généralement fait en verre1.

L'exemple qui vient d'être examiné montre que certaines lexies entretiennent entre elles une relation formelle et sémantique privilégiée. On dira qu'elles appartiennent au même vocable :

Un **vocable** est un regroupement de lexies

- qui sont associées aux mêmes signifiants
- et qui ont un lien sémantique évident.

Par exemple, si l'on consulte le *Nouveau Petit Robert*, on verra que ce dictionnaire considère que le vocable français PORC contient les quatre lexies : PORC1 ('animal domestique...'), PORC2 ('individu sale...'), PORC3 ('viande de porc1...') et PORC4 ('peau du porc1...').

Les lexies d'un vocable sont souvent appelées **acceptations** de ce vocable. La **polysémie** est la propriété d'un vocable donné de contenir plus d'une lexie. Ainsi, le vocable PORC est polysémique alors que le vocable PHACOCHÈRE (pour rester dans le même domaine) est monosémique. Nous reviendrons sur ces importantes notions plus loin dans le cours.



Ce que l'on appelle habituellement une « entrée » de dictionnaire correspond en fait à la description d'un vocable. Il est important de se souvenir que c'est par simple convention que, dans les dictionnaires français, les vocables verbaux sont nommés et ordonnés alphabéti-

quement selon leur forme de l'infinif, les noms selon leur forme du singulier et les adjectifs selon leur forme du masculin singulier. Pour des raisons pratiques, on doit utiliser une forme particulière pour référer à un vocable et en stocker la description dans le dictionnaire ; cependant, dans les faits, une lexie est un regroupement de signes et le vocable un regroupement de lexies.

Il peut arriver que deux lexies distinctes soient associées aux mêmes signifiants alors qu'elles n'entretiennent aucune relation de sens ; il s'agit d'un cas d'*homonymie*.



Nous distinguerons les lexies homonymes par des numéros mis en exposant. Par exemple :

- ADRESSE¹ [*Il fait preuve de beaucoup d'adresse.*]
~
ADRESSE² [*Est-ce que tu connais son adresse en France ?*]
- PAVILLON¹ [*Il vit dans un pavillon de banlieue.*]
~
PAVILLON² [*Ils ont hissé le pavillon noir.*]



En théorie, notre approche étant synchronique, nous devons nous baser sur la présence ou l'absence d'un lien sémantique en français contemporain pour décider si nous sommes en présence d'homonymes ou de lexies d'un même vocable. S'il existe un lien étymologique mais que ce lien n'est plus concrétisé par une relation de sens couramment perçue par les locuteurs, nous devons l'ignorer. Nous reviendrons plus loin dans l'ouvrage sur la notion d'homonymie (Chapitre 7).

Étude des lexies et étude de la sémantique des langues

En tant que signe linguistique ou, plus précisément, en tant que regroupement de signes linguistiques (ses mots-formes ou expressions associés), la lexie peut être décrite selon trois axes :

- 1 son sens (signifié) ;
- 2 sa forme (signifiant) ;
- 3 ses propriétés de combinatoire.

Le sens est mentionné ici en premier car c'est bien évidemment la caractéristique de la lexie qui nous intéresse le plus en lexicologie. Pour des raisons pédagogiques, je vais cependant introduire dans le prochain chapitre certaines notions de morphologie (qui sont donc plutôt liées aux signifiants lexicaux) avant de passer au plat de résistance : l'étude des sens lexicaux.

Remarques sur la terminologie et les conventions d'écriture

Comme le met en évidence la citation d'Aristote donnée en exergue à ce chapitre, la construction et la manipulation d'une terminologie est un aspect essentiel de toute activité scientifique et a donc une importance considérable en linguistique.

De plus, les conventions d'écriture utilisées dans les textes scientifiques sont un reflet de l'activité de formalisation inhérente à tout travail scientifique et vont de pair avec la construction d'une terminologie. Toute personne voulant se former à l'étude du lexique (et, plus généralement, à la linguistique), devra apprendre de telles conventions et, ensuite, devra les respecter dans ses écrits.



Je récapitule ici trois conventions d'écriture que j'utilise systématiquement dans ce livre et qui ont déjà été introduites précédemment :

- | | | | |
|----------|---|---|---|
| 1 | lexie (unité lexicale) | : | HUÎTRE, CASSER1, FRUIT DE MER, ... |
| 2 | signe linguistique ou sa forme (signifiant écrit) | : | <i>huître</i> , <i>casser1</i> , <i>fruit de mer</i> , <i>-s</i> [pluriel des noms], <i>re-</i> , ...
ou
<u>huître</u> , <u>casser1</u> , <u>fruit de mer</u> , <u>-s</u> , <u>re-</u> , ... ⁴ |
| 3 | sens (signifié) | : | ('huître'), ('casser1'), ('fruit de mer'), ('plusieurs'), ('encore'), ... |

4. Lorsqu'on écrit à la main, il est normal d'utiliser le soulignement pour remplacer l'italique. Le lecteur se rappellera de plus (voir page 22) qu'il faudrait en théorie utiliser deux conventions différentes : une pour les signes linguistiques et une pour les signifiants de signes.

À noter que l’italique (ou le soulignement) a souvent d’autres emplois ; il sert notamment pour les titres d’ouvrages. De plus, on peut aussi mentionner un signifiant en utilisant une transcription phonologique.

À fin d’illustration, voici un exemple de texte où ces conventions d’écriture ont été rigoureusement respectées :

En français, on doit considérer que FRUIT DE MER est une unité lexicale à part entière puisque son sens ne peut être compris comme résultant de la composition régulière de ‘fruit’, ‘de’ et ‘mer’. Elle est d’ailleurs décrite comme telle dans pratiquement tous les dictionnaires du français, comme le *Petit Robert*. Ses formes singulier et pluriel sont, respectivement, *fruit de mer* et *fruits de mer*.

Lectures complémentaires

Mel’čuk, Igor (1993) Chapitre I. Mot-forme et lexème : étude préliminaire. In : *Cours de morphologie générale*, Volume 1, Montréal/Paris : Les Presses de l’Université de Montréal/CNRS ÉDITIONS, pp. 97-107.

Le terme *mot-forme* est assez peu répandu dans la littérature linguistique francophone, beaucoup moins que son équivalent anglais *wordform* (parfois écrit *word-form* ou *word form*). Il est avant tout associé à l’approche morphologique d’Igor Mel’čuk. Que l’on utilise ce terme ou un autre, on ne peut cependant pas faire l’économie de la notion correspondante, qui est au centre de la caractérisation du lexème et, par voie de conséquence, de la lexicologie elle-même. La lecture suggérée ci-dessus fait une première présentation assez développée des notions liées de mot-forme et lexème. Si le lecteur éprouve des difficultés à bien en saisir le contenu, il sera peut-être préférable d’attendre d’avoir complété l’étude du prochain chapitre (Chapitre 4, *Éléments de morphologie*) pour revenir sur ce texte.

Eluerd, Roland (2000) Chapitre I. Situation de la lexicologie, Sections I et II. In : *La lexicologie*, Collection « Que sais-je ? », n° 3548, Paris : Presses Universitaires de France, pp. 6-22.

Ce texte peut être lu pour les nombreux pointeurs vers d’autres ouvrages sur la lexicologie qu’il contient. Il propose aussi une perspective un peu plus « philosophique » sur la discipline que ce que j’ai présenté dans ce chapitre.

Exercices

- 1 Soit la phrase suivante :
(14) *La grève des pilotes devrait prendre fin cet après-midi.*
Analyser (14) en terme (i) de mots-formes et (ii) de lexies.
- 2 Démontrer qu'il existe plusieurs lexies CERCLE en français.
- 3 Démontrer que COUP DE MAIN [*Il m'a donné un coup de main pour refaire mon mur.*] est une lexie.
- 4 Trouver le plus d'arguments possible pour démontrer que les deux expressions en gras dans la phrase ci-dessous correspondent à deux lexies distinctes du français :
(15) *Pour sortir, il **pousse** brutalement la branche qui **pousse** près de la porte d'entrée.*
- 5 Même question pour la phrase ci-dessous, mais avec trois expressions correspondant à trois lexies :
(16) — *Je n'**arrive** pas à comprendre qu'est-ce qui t'**arrive** d'**arriver** à cinq heures ?*
- 6 En quoi les deux expressions en gras dans les phrases ci-dessous sont-elles de natures différentes ?
(17) a. *Il s'**est cassé la jambe** en tombant.*
b. *Il s'**est cassé la tête** pour résoudre ce problème.*
- 7 Expliquer pourquoi *casser un jugement* [*La Cour d'Appel a cassé le jugement condamnant Jules à quinze ans de prison.*] n'est pas une locution verbale au même titre que *casser du sucre sur le dos de quelqu'un* (qui signifie 'dire du mal de quelqu'un') [*C'est pas sympa de casser du sucre sur le dos de ta collègue.*].



Chapitre 4 Éléments de morphologie

Morphologie, morphe, signe segmental vs suprasegmental, morphème, allomorphe, radical, affixe, suffixe, préfixe, flexion, affixe flexionnel, catégorie flexionnelle, forme fléchie, signe zéro, dérivation, affixe dérivationnel, partie du discours, composition, abréviation, siglaison, acronyme.

*Je ne communiquerai plus par écrit avec mon voisin.
Tu ne communiqueras plus par écrit avec ton voisin.
Il ne communiquera plus par écrit avec son voisin,
etc., etc.*

Patrick Cauvin, *Tout ce que Joseph écrivit cette année-là*

Nous avons vu dans le chapitre précédent qu'il était impossible de définir la lexie avec un minimum de rigueur sans entrer pour cela dans le domaine de la morphologie (cf. les notions de mot-forme et de flexion, qu'il nous a fallu aborder). Il est utile de consacrer maintenant un chapitre entier à l'introduction des notions élémentaires de morphologie dont l'usage est incontournable en lexicologie.

L'examen de ces notions est nécessaire pour au moins deux raisons. Tout d'abord, l'analyse morphologique permet de mieux comprendre la notion de lexie au niveau de la structure du signifiant. Ensuite, elle permet de modéliser certaines relations formelles et sémantiques existant entre les lexies de la langue.

Commençons par définir la discipline qui fait l'objet du présent chapitre :

La **morphologie** est la branche de la linguistique qui étudie la structure des mots-formes.

La morphologie visant la description de la structure des mots-formes¹, il va nous falloir introduire :

- les éléments constitutifs de cette structure ;
- les différents « mécanismes d'assemblage » de ces éléments.

Commençons par l'examen des éléments constitutifs des mots-formes et, plus spécifiquement, par les signes morphologiques élémentaires.

Signe morphologique élémentaire

Le morphe

Alors que certains mots-formes ne semblent pas, à première vue, analysables en terme d'autres signes :

(1) *chien, assez, vite, ...*

d'autres se décrivent clairement comme résultant de la combinaison de plusieurs signes :

(2) a. *chiens = chien + -s*

b. *tristement = triste + -ment*

Lorsque l'on décompose complètement un mot-forme en signes plus simples le constituant, on arrive à des signes élémentaires (voir Chapitre

1. Rappel : cette notion a été définie au chapitre précédent.

2, page 25), qui ne sont pas eux-mêmes analysables. C'est ce qui se produit en (2a-b) avec les signes élémentaires *chien*, *-s*, *triste* et *-ment*.²

On appelle **morphe** un signe linguistique

- dont le signifiant est un segment de la chaîne parlée
- et qui est un signe élémentaire, c'est-à-dire qui ne peut être représenté en terme d'autres signes de la langue.

Les morphes sont donc les signes élémentaires mis à jour par l'analyse morphologique. Mais pourquoi est-il spécifié, dans la définition ci-dessus, que le signifiant du morphe est un **segment** de la chaîne parlée (*grosso modo*, une séquence phonique) ?

Cette précision est nécessaire parce qu'il existe des signes linguistiques dont le signifiant **se superpose** à la chaîne parlée, comme par exemple les intonations. Comparons les deux phrases ci-dessous :

- (3) a. *Il dort.*
b. *Il dort ?*

Ces phrases sont formellement identiques à l'oral si l'on ne considère que la séquence de sons élémentaires (plus précisément, de phonèmes) dont elles sont constituées. Cependant, une différence très importante existe, différence indiquée à l'écrit par les deux marques de ponctuation finales : le point indique que (3a) est une affirmation, qui doit donc se lire avec une intonation légèrement descendante, et le point d'interrogation indique que (3b) est une interrogation, qui doit être lue avec une intonation montante. Pour mieux comprendre de quoi il s'agit, raisonnons directement sur les signifiants oraux plutôt qu'écrits, en examinant les transcriptions phonémiques de nos deux exemples, transcriptions agrémentées de schémas intonatifs approximatifs :

2. Ce que je dis ici est en partie inexact. Je reviendrai sur l'analyse de *chien* un peu plus bas (Section *Remarque sur les signes zéro*, page 54).

- (4) a. 
/ildɔ̃r̃/
- b. 
/ildɔ̃r̃/

Les schémas intonatifs apparaissant en (4a) et (4b) peuvent être analysés comme deux associations signifié ~ signifiant :

- affirmation ~ intonation légèrement descendante
vs
- interrogation ~ intonation montante à la fin de la suite sonore.

Si l'on donne une interprétation un peu large au terme d'*image acoustique* utilisé par Saussure pour parler du signifiant (voir Chapitre 2, page 21), on est ici en présence de deux signes linguistiques. Il sont cependant très spéciaux dans la mesure où leur signifiant n'est justement pas un segment de la chaîne parlée, mais se superpose à elle.

On appelle **signe segmental** un signe linguistique dont le signifiant est un segment de la chaîne parlée et **signe suprasegmental** un signe linguistique dont le signifiant se superpose à la chaîne parlée. Bien entendu, les signes segmentaux sont les signes « standard », ceux auxquels on pense de prime abord lorsque l'on parle de signes linguistiques.

On voit qu'il était très important, dans la définition de la page 49, de caractériser les morphes en tant que signes segmentaux élémentaires pour ne pas inclure les phénomènes tels que les contours intonatifs dans le champ d'étude de la morphologie. Leur analyse ne relève en effet pas de l'étude de la structure interne des signes lexicaux.

Le morphème

La description du morphe en tant que signe linguistique implique que le morphe est une association entre **un** signifié et **un** signifiant uniques. Or, il est fréquent de rencontrer des morphes qui semblent « changer de

signifiant » dans certains contextes. Par exemple, considérons les données suivants :

- (5) a. *stupide* ~ *stupidité*
b. *majeur* ~ *majorité*

Il est clair que les deux paires de mots-formes ci-dessus sont naturellement en relation au niveau sémantique **et** formel :

- le rapport de sens existant entre *stupide* et *stupidité* est identique au rapport de sens existant entre *majeur* et *majorité* ;
- l'écart formel existant entre *stupide* et *stupidité* est **presque** identique à celui existant entre *majeur* et *majorité*.

En d'autres termes, on voudrait associer à (5a) et (5b) deux analyses morphologiques que l'on pourrait **très grossièrement** représenter de la façon suivante :

- (6) a. *stupidité* = *stupide* + *-ité*
b. *majorité* = *majeur* + *-ité*

Cependant, si l'on ne dispose que de la notion de morphe pour décrire la structure formelle des mots-formes, on va être obligé de considérer un morphe « à signifiant variable » *majeur* qui devient *majeur-* en combinaison avec *-ité*. C'est une stratégie très dangereuse, qui comporte le risque de rendre vague et fuyante la notion de signe linguistique. Il vaut mieux considérer que la structure des mots-formes doit se décrire à partir d'une entité de plus haut niveau que le morphe : il s'agit du **morphème**, qui est un regroupement de morphes « alternatifs » ayant le même signifié. Ainsi, les deux morphes *majeur-* et *major-* sont regroupés sous le même morphème {MAJEUR}. (J'utilise les accolades pour indiquer qu'un morphème est un ensemble de morphes.) Selon les morphèmes avec lesquels {MAJEUR} doit se combiner, il trouvera son expression soit sous la forme du morphe *majeur-* soit sous celle du morphe *major-*.

Les morphes regroupés sous un même morphème sont appelés **allomorphes** de ce morphème. On voit que cette terminologie est parallèle à celle utilisée en phonologie : phonème *vs* allophones de ce phonème. Comme en phonologie pour les allophones, il est nécessaire d'avoir recours à des critères bien spécifiques pour déterminer sous

quelles conditions plusieurs morphes peuvent être considérés comme allomorphes d'un même morphème. L'étude de ces critères nous ferait entrer trop profondément dans le domaine de la morphologie. Je me contente donc ici d'une ébauche de présentation de la notion de morphème, évitant soigneusement d'en proposer une définition véritable.

Nous sommes de toute façon arrivés à la limite de ce qu'il est possible de dire sur le sujet de l'analyse morphologique sans introduire d'autres notions centrales en morphologie : celles de radical et d'affixe.

Radical et affixe

Toute lexie possède un radical :

Le **radical** d'une lexie est son support morphologique. C'est l'élément morphologique qui porte le signifié associé en propre à cette lexie.

Dans les cas standard, on retrouve le radical dans toutes les « manifestations morphologiques » de la lexie. Par exemple, le radical de CHANTER est *chant-* — que l'on retrouve dans *chanter*, *chante*, *chantes*, ... —, celui de RECONSIDÉRER est *reconsidér-*, etc.

Ce dernier cas illustre le fait que le radical d'une lexie n'est pas nécessairement **un** morphe : *reconsidér-* peut s'analyser en tant que *re-* + *considér-* (le radical de CONSIDÉRER).

Bien entendu, la notion de radical ne concerne que les lexèmes (c'est-à-dire, les lexies qui ne sont pas des locutions).

La présentation qui vient d'être faite du radical présuppose une autre notion, puisque le radical d'une lexie — par exemple, *chant-* (pour la lexie CHANTER) — est un signe segmental auquel vient s'ajouter d'autres signes morphologiques d'une nature bien particulière — *-er*, *-e*, *-es*, etc.

On appelle **affixe** un morphe non autonome, qui est destiné à se combiner avec d'autres signes morphologiques au sein d'un mot-forme.

- 3 sa combinaison avec le radical d'une lexie donne un mot-forme qui est associé à la même lexie.

Par exemple, le radical verbal *chant-* peut être combiné avec l'affixe flexionnel de l'infinitif des verbes du premier groupe *-er* pour produire le mot-forme *chanter* ; il peut être combiné avec l'affixe flexionnel de première personne du singulier de l'indicatif présent des verbes du premier groupe *-e* pour former *chante* ; etc. On dira que *chanter*, *chante*, etc. sont des **formes fléchies** de la lexie CHANTER.

Toutes les langues n'ont pas la même richesse flexionnelle. Ainsi, le français est flexionnellement beaucoup plus riche que l'anglais. De plus, les affixes flexionnels du français sont tous des suffixes : ils se combinent à droite du radical.

Remarque sur les signes zéro

Les notions de radical et de flexion nous permettent de revenir sur le cas de *chien* mentionné ci-dessus (page 48). J'ai dit que *chien*, par opposition à *chiens*, était un morphe dans la mesure où il n'était pas analysable en signes plus simples. Cela doit être considéré comme inexact, maintenant que les notions de radical et de flexion (catégorie flexionnelle) ont été introduites. Dans une phrase comme :

(7) *Le chien de Charles cherche un chat.*

le signe *chien* est en fait la combinaison du radical *chien-* (radical de la lexie CHIEN) et d'un suffixe flexionnel très particulier exprimant le singulier des noms. En effet, on sait que dans (7) le nom *chien* est au singulier du fait de l'absence du signifiant de l'affixe flexionnel de pluriel des noms *-s*. Le singulier des noms est donc exprimé en français par un signe dont le signifiant est une absence physique de signifiant, un « trou » morphologique en quelque sorte. De tels signes sont appelés **signes zéro**.



Pour résumer, la forme *chien* apparaissant dans l'exemple (7) devrait être analysée comme un signe complexe formé de deux morphes : *chien-* + \emptyset_{sing} (radical de CHIEN + suffixe zéro du singulier des noms).

Dérivation

Définition de la dérivation

La **dérivation** est, dans le cas le plus standard, un mécanisme morphologique consistant en la combinaison d'un radical et d'un affixe — appelé **affixe dérivationnel** — ayant les trois propriétés suivantes :

- 1 son signifié est moins général et moins abstrait que celui d'un affixe flexionnel — il se rapproche du signifié d'une lexie ;
- 2 l'expression de son signifié correspond normalement à un choix libre du locuteur, qui **décide** de communiquer le signifié en question ;
- 3 sa combinaison avec le radical d'une lexie donne un mot-forme qui est associé à une autre lexie.

Ainsi, le radical verbal *chant-* (de la lexie CHANTER) peut être combiné avec l'affixe dérivationnel *-eur* (dont le signifié est *grosso modo* 'personne qui fait...'), pour produire le radical nominal *chanteur* ('personne qui chante').

La dérivation est donc un mécanisme morphologique renvoyant à une relation entre lexies. Par exemple, CHANTEUR est un dérivé nominal de CHANTER.

Les affixes dérivationnels du français sont soit des suffixes (cf. *-eur*) soit des préfixes (cf. *re-* dans *reconsidérer*). Attention cependant : il existe d'autres types de dérivations ! Voir à ce propos l'Exercice 6, à la fin du chapitre.

Types de liens dérivationnels

La dérivation permet de « passer » d'une lexie à une autre. On peut caractériser les différents types de dérivations en fonction des écarts sémantiques et grammaticaux existant entre les lexies que ces dérivations mettent en relation. Ainsi, on peut regarder si les lexies en question sont synonymes ou non et si elles appartiennent ou non à la même **partie du discours** (nom, verbe, adjectif ou adverbe)⁴.

4. La notion de partie du discours sera examinée en détail dans le chapitre suivant.

Lorsqu'il y a combinaison d'un affixe dérivationnel avec le radical d'une lexie L_1 pour donner le radical d'une seconde lexie L_2 , on peut être en présence d'au moins un des quatre cas de figure suivants⁵ :

Cas de figure	Exemples : $L_1 \rightarrow L_2$
<ul style="list-style-type: none"> • sens 'L_1' fait partie du sens 'L_2' • partie du discours de L_1 = partie du discours de L_2 	LOUER _V → RELOUER _V
<ul style="list-style-type: none"> • sens 'L_1' fait partie du sens 'L_2' • partie du discours de L_1 ≠ partie du discours de L_2 	CHANTER _V → CHANTEUR _N
<ul style="list-style-type: none"> • 'L_1' et 'L_2' sont équivalents • partie du discours de L_1 = partie du discours de L_2 	CON _N → CONARD _N FURIBOND _{Adj} → FURIBARD _{Adj}
<ul style="list-style-type: none"> • 'L_1' et 'L_2' sont équivalents • partie du discours de L_1 ≠ partie du discours de L_2 	DÉCIDER _V → DÉCISION _N

Quatre types de liens dérivationnels

Il est évident que, pour bien rendre compte des exemples présentés dans la colonne de droite du tableau ci-dessus, il faudrait faire une analyse plus fine des relations de sens existant entre les lexies en cause. Il nous suffit pour l'instant de noter que les dérivations peuvent

- ajouter ou non un sens au sens de départ,
- associer à la lexie d'origine une lexie appartenant ou non à la même partie du discours.

À ce stade, il peut sembler étrange de dire que des lexies comme, par exemple, DÉCIDER et DÉCISION ont en gros le même sens. Nous verrons

5. Le troisième cas de figure apparaissant dans ce tableau est très rare en français et je n'ai pu en trouver des exemples que dans le langage familier ou vulgaire. On remarquera que les deux noms en question sont sémantiquement presque impossibles à distinguer. Leur différence réside dans le fait que le nom dérivé est plus vulgaire et d'un emploi encore plus injurieux que le nom dont il est morphologiquement issu.

cependant qu'il existe des raisons théoriques et pratiques pour procéder de la sorte.

Dérivation synchronique vs dérivation diachronique

J'ai présenté plus haut la dérivation sur le même plan que la flexion, c'est-à-dire comme étant un mécanisme morphologique. Or, dans le cas du français, cela est en grande partie un abus de langage : il conviendrait de distinguer la dérivation synchronique et la dérivation diachronique.

D'un point de vue synchronique, la véritable dérivation est rare en français. On peut citer comme exemple la dérivation qui se fait par ajout du préfixe *re-* et qui produit un verbe avec ajout de la signification 'de nouveau' :

- (8) a. *manger + re- → remanger*
b. *lire + re- → relire*
c. *dessiner + re- → redessiner*

On voit que cette dérivation est très productive car le préfixe *re-* peut être combiné avec la plupart des radicaux verbaux pour former un autre verbe. On ne va donc pas entrer dans un dictionnaire la description des lexies REMANGER, RELIRE, REDESSINER, etc. Celles-ci sont construites à volonté par le locuteur à partir de sa connaissance des radicaux correspondants et de la règle de dérivation en *re-*.

Mais les règles de dérivation de ce type sont rares en français. Dans notre langue, on trouve surtout des cas de dérivations diachroniques, c'est-à-dire des dérivations non productives au niveau du locuteur : c'est la langue elle-même, dans son évolution, qui présente un tel cas de dérivation. Par exemple, les liens dérivationnels suivants n'existent que de façon diachronique :

- (9) a. *consomm- + -ation → consommation*
b. *communiqu- + -ation → communication*

En effet, un locuteur va employer CONSOMMATION parce que cette lexie existe en français, et non en la construisant lui-même à partir du radical *consomm-* et d'une règle morphologique de dérivation. Sinon, comment

expliquer que l'on ait *consommation* mais pas **mangeation*, *communication* mais pas **parlation*, etc. ?



L'étude de la dérivation en français va donc surtout porter sur des dérivations lexicalisées (qui n'apparaissent comme telles que dans une perspective diachronique). On peut considérer que, dans de tels cas, ce sont les **signifiants** des mots-formes que l'on décrit au moyen d'un mécanisme de dérivation et non les mots-formes (les signes linguistiques) eux-mêmes.

Composition

La composition est, comme la dérivation, un mécanisme morphologique « construisant » de nouveaux radicaux : un mot-forme est formé par **composition** lorsqu'il résulte de la concaténation de plusieurs radicaux. Par exemple : *bon* + *homme* ⇒ *bonhomme*, *porte_V* + *manteau* ⇒ *portemanteau*, etc.

Bien entendu, la composition n'est pas une concaténation aléatoire de radicaux. Il existe en français des patrons de composition. Pour la composition nominale, par exemple, on retrouve fréquemment les deux patrons suivants :

1 adjectif épithète + nom

BONHOMME ~ BONJOUR ~ BONNE-MAMAN ~ GRANDS-PARENTS

2 verbe + complément

CASSE-NOISETTE [ou CASSE-NOISETTES] ~
PORTE-MANTEAU [ou PORTEMANTEAU] ~
OUVRE-BOÎTE

Comme on peut le constater, il existe un certain flottement dans la façon dont s'orthographient les lexies composées, cette fluctuation étant le reflet du caractère flou de la frontière séparant parfois en français locutions et lexies composées.

Il faut aussi noter que l'on voit souvent dans les grammaires et dans les textes de linguistique la notion de composition étendue à la lexicalisation de groupes syntaxiques, c'est-à-dire à la formation de locutions :

FAIT DIVERS, POMME DE TERRE, COUP DE CŒUR, LAISSER TOMBER, etc. Cela s'explique par le fait que les mots-formes dont les locutions sont formellement constituées sont en quelque sorte « dégénérés » et que les locutions tendent à se comporter comme des blocs morphologiques, à la manière des mots-formes (voir Chapitre 3, Section *Les locutions*, page 38). Ce qui importe, pour une locution, c'est de posséder une structure syntaxique interne qui en fait une expression linguistique valide. Par exemple, JETER L'ÉPONGE est une locution verbale parce que sa structure interne est celle d'un groupe verbal type du français : verbe, qui admet toutes les formes fléchies verbales du français (*Je jette l'éponge, tu jettes l'éponge, etc.*), suivi d'un complément d'objet direct. On notera par contre qu'il serait impossible d'analyser PORTEMANTEAU de la même façon, même en ignorant l'orthographe qui fait que cette lexie s'écrit en un seul mot. En effet, PORTEMANTEAU, qui a une structure interne verbe + complément (sur le modèle de *Il porte un manteau*), est une lexie **nominale** et non verbale. Il n'y a donc aucune connexion entre la structure syntaxique interne que l'on pourrait chercher à imposer à cette lexie et son fonctionnement linguistique dans la phrase. Nous sommes bien ici en présence d'un nom composé et non d'une locution. La différence entre les deux notions de locution et de lexie composée existe donc bel et bien, même si elle n'est pas facile à cerner dans tous les cas.



La composition n'existe quasiment pas **en français** en tant que mécanisme morphologique **synchronique**. C'est uniquement la perspective diachronique (apparition de nouvelles lexies en français) qui permet de parler de composition dans le cas de cette langue.

Dans le cadre d'une véritable étude de la morphologie lexicale, et notamment dans le cadre d'une étude diachronique, on pourrait aller beaucoup plus loin dans l'examen des modes de formation des lexies. Je ne ferai que mentionner ici deux autres modes de formation lexicale n'existant, comme la composition en français, que d'un point de vue diachronique :

- 1 L'**abréviation** (APPART, AUTO, ...) relève en général du langage parlé ou familier et permet de produire une nouvelle lexie par troncation du radical d'une lexie initiale.

- 2 La **siglaison** (USA, PDG...) produit une lexie à partir d'une locution en concaténant les lettres initiales de chacune des lexies de la locution en question. En français, les sigles sont normalement des noms. Un sigle qui se prononce comme une suite de syllabes et non en épelant les lettres est appelé **acronyme** : OTAN (/ɔtã/ et pas /ɔteaen/), NASA (/naza/ et pas /enaesa/), ...

Lectures complémentaires

Revenir éventuellement sur la lecture Mel'čuk (1993) du chapitre précédent (page 45), pour les notions de mot-forme et lexème.

Nida, Eugene A. (1976) Chapter 1: Introduction to morphology. In : *Morphology: The Descriptive Analysis of Words*, Second Edition, Ann Arbor: The University of Michigan Press, pp. 1-5.

Ce court texte, d'une grande clarté, situe l'étude morphologique dans le cadre plus général de l'étude des langues. On notera certains écarts terminologiques par rapport à ce qui a été dit dans ce chapitre, notamment pour ce qui est de l'emploi du terme *word* (= mot). Cette remarque vaut aussi pour la lecture suggérée ci-dessous.

Lehmann, Alise et Françoise Martin-Berthet (1998) Deuxième partie : Morphologie lexicale. In : *Introduction à la lexicologie : Sémantique et morphologie*, Collection « Lettres Sup », Paris : Dunod, pp. 99-185.

La seconde partie de cet ouvrage est en quelque sorte un mini-manuel d'introduction à l'étude de la morphologie, centré sur le français. On peut donc le consulter si l'on désire en savoir plus sur cette branche de la linguistique et voir d'autres exemples d'analyses morphologiques de mots-formes français.

Exercices

- 1 Il a été dit des signes énumérés dans l'exemple (1) qu'ils n'étaient pas analysables en terme d'autres signes de la langue. Pourquoi le fait que *maison* contienne deux syllabes — /mɛ/ et /zɔ̃/ — ne contredit-il pas cette affirmation ?

- 2 Comparer, le plus exhaustivement possible, tous les types de flexions que l'on trouve en anglais et en français.
- 3 Examiner les formes fléchies du verbe ÊTRE. Cela permet-il de mettre en évidence un problème par rapport à la définition de la notion de flexion présentée dans ce chapitre ? (Penser à la notion de radical.)
- 4 Décrire le plus complètement possible en tant que signe linguistique (signifié, signifiant et propriétés de combinatoire) l'affixe *-age*, que l'on trouve dans *nettoyage, débauchage, etc.*
- 5 Comparer la façon dont s'écrivent les deux lexies suivantes :

(10) PORTEFEUILLE ~ PORTE-MONNAIE [ou, parfois, PORTEMONNAIE]

J'ai mentionné plus haut (page 58) qu'il y avait une certaine part d'arbitraire dans la façon dont s'orthographient les lexies composées. Est-il cependant possible de trouver une explication à la différence constatée ici ?

- 6 Est-ce que l'on est en présence d'un cas de dérivation dans

(11) *Il n'oublie jamais d'apporter son **manger** car il aime trop **manger** ?*

Qu'est-ce que cela implique au niveau de la notion de dérivation telle qu'elle a été définie page 55 ?

- 7 Le sigle ONU est-il un acronyme ?



Chapitre 5

Description de la structure du lexique

Lexique, vocabulaire d'un texte vs d'un individu, idiolecte, marque d'usage, dialecte, terminologie, partie du discours (= classe grammaticale, catégorie syntaxique), classe ouverte vs fermée, mot lexical vs grammatical, lien paradigmatique vs syntagmatique, introspection, enquête linguistique, corpus linguistique, fréquence d'emploi, occurrence, index (de signifiants lexicaux), lemmatisation, concordance (= KWIC), statistique lexicale, linguistique quantitative, hapax.

Toutes les utopies sont déprimantes, parce qu'elles ne laissent pas de place au hasard, à la différence, aux « divers ». Tout a été mis en ordre et l'ordre règne.

Derrière toute utopie, il y a toujours un grand dessein taxinomique : une place pour chaque chose et chaque chose à sa place.

Georges Perec, *Penser/Classer*

Les chapitres précédents nous ont permis de définir et caractériser avec une certaine précision la lexie, l'unité de base dont sont constitués les lexiques. Nous allons maintenant examiner comment ceux-ci sont **structurés**. Le fait de parler de la structure des lexiques nous amènera naturellement à aborder le sujet de l'accès aux données lexicales. Cela explique pourquoi le présent chapitre est relativement dense, comme l'indique la longue liste de notions énumérées ci-dessus.

Je vais tout d'abord préciser la notion de lexique, en la mettant en contraste avec celle de vocabulaire ; cela me conduira à examiner la question de la variation linguistique. Puis, je présenterai les parties du discours, qui sont des classes de lexies regroupées en fonction de leurs

caractéristiques grammaticales. Nous examinerons ensuite les différents types de liens que peuvent entretenir les lexies au sein du lexique d'une langue. (Cette section sera très brève dans la mesure où un chapitre entier du cours — Chapitre 7 — est consacré à l'étude et à la modélisation des liens lexicaux.) Le lexique étant un ensemble d'informations très riche et très complexe, il nous faudra ensuite examiner le problème de l'identification des connaissances lexicales ; nous verrons ainsi quelles méthodes d'accès aux données linguistiques peuvent être utilisées en lexicologie. Finalement, je conclurai ce chapitre par quelques observations sur les phénomènes statistiques liés à l'usage du lexique.



Malgré son titre, ce chapitre ne prétend pas démontrer que les lexiques possèdent **une** structure donnée. Et c'est pour faire contrepoids à l'ardeur classificatrice de certains de mes collègues que j'ai mis en exergue cette citation de Georges Perec. On peut toujours effectuer des classifications et des regroupements des lexies de la langue. Cependant, ces structures descriptives que l'on plaque sur le lexique ne servent qu'à mettre en évidence des aspects bien spécifiques de son organisation. Elles répondent à un besoin pratique de classement mais, prises individuellement, elles ne peuvent suffire à modéliser la structure du lexique, qui est multidimensionnelle. Cette particularité de la structure du lexique nous oblige à la considérer selon plusieurs angles simultanément.

Lexique *vs* vocabulaire

La notion de lexique

Le moment est venu de mettre noir sur blanc une définition de la notion de lexique :

Le **lexique** d'une langue est l'entité théorique correspondant à l'ensemble des lexies de cette langue.

Par *entité théorique*, je veux dire que le lexique n'est pas véritablement un ensemble dont les éléments, les lexies, peuvent être énumérés de

façon systématique. Le lexique ressemble plutôt à un « ensemble flou », un ensemble dont il n'est pas toujours possible de dire s'il contient tel ou tel élément particulier.

Voici trois cas de figure permettant d'illustrer cette caractéristique du lexique :

- 1 Un terme entendu dans une conversation ou lu dans un article de journal va souvent poser problème si c'est un anglicisme, c'est-à-dire si c'est un emprunt plutôt récent de l'anglais¹. Ainsi, le nom FAX [*Je vous envoie un fax immédiatement.*] fait-il ou non partie du lexique français ? Certains voudront l'inclure, parce qu'il est d'un emploi très fréquent. D'autres le rejeteront, parce qu'ils sont opposés par principe à l'usage d'anglicismes, et exigeront que l'on utilise TÉLÉCOPIE. Bien entendu, il y a aussi le comportement très « français » qui consiste à aller vérifier dans le dictionnaire pour trouver une réponse. Même les plus puristes toléreront donc peut-être FAX, puisque ce lexème apparaît dans le *Petit Robert* (avec le verbe correspondant FAXER, d'ailleurs). Il est à noter cependant qu'ils rejeteront le nom E-MAIL [*J'ai reçu plus de trente e-mails hier !*], qui est d'un emploi tout aussi courant, s'ils consultent le *Petit Robert* dans sa version 1996. Par contre, il l'accepteront peut-être s'ils ont une version plus récente de ce dictionnaire, qui introduit le vocable en question, avec mention de la terminologie officielle recommandée pour chacun de ses deux sens (ADRESSE ÉLECTRONIQUE [*Est-ce que tu connais son e-mail ?*] et COURRIER ÉLECTRONIQUE [*J'ai reçu son e-mail.*]). Le recours au dictionnaire n'est donc pas une solution miracle, surtout si vous êtes vous-même lexicologue ou lexicographe (rédacteur de dictionnaires) et êtes à la recherche de critères rationnels et cohérents pour mener à bien votre travail descriptif².

1. Seuls les emprunts récents nous intéressent ici. Discuter du fait de savoir si WAGON est ou non un anglicisme, alors que ce lexème est utilisé en français depuis plusieurs siècles, relèverait véritablement de l'ergotage.

2. Il faut souligner qu'il ne s'agit pas, dans le cas des anglicismes, d'un simple problème d'« attitude » : faire preuve de laisser-faire ou de purisme. Les emprunts lexicaux à l'anglais sont tellement nombreux en français contemporain qu'ils posent de véritables problèmes descriptifs, du fait notamment du mélange des codes que leur présence entraîne dans les domaines morphologique et phonologique. Voir à ce propos le texte de Josette Rey-Debove donné en lecture complémentaire pour ce chapitre.

- 2 Une expression comme *Défense de stationner* est-elle ou non une lexie ? C'est une entité linguistique qui forme un tout sémantique. Même si elle est autonome — comme l'est une phrase —, elle possède toutes les propriétés du signe linguistique. Une telle entité linguistique, comme d'ailleurs les proverbes (*Un tiens vaut mieux que deux tu l'auras*, *Qui vole un œuf vole un bœuf*, etc.) peut être ou non incluse dans le lexique, selon la perspective théorique que l'on adopte vis-à-vis de la notion de lexie.
- 3 Finalement, et c'est peut-être point le plus important, tout le monde ne partage pas la même connaissance de la langue. Le lexique de la langue doit-il donc être l'**union** ou l'**intersection** de la connaissance lexicale des locuteurs ? Dans le premier cas — si l'on prend la somme de **toutes** les connaissances lexicales de **tous** les locuteurs de la langue, on risque de se retrouver avec un lexique hybride qui ne pourra en aucun façon être considéré comme formant un système. Dans le second cas — si l'on ne considère **que** les connaissances qui sont véritablement communes à **tous** les locuteurs, on risque de laisser de côté une partie essentielle du lexique, certaines personnes ayant clairement une connaissance lexicale très inférieure à la moyenne.

Ces trois cas de figure suffisent à montrer que, lorsque l'on parle du lexique d'une langue, on postule une entité théorique qui, dans les faits, ne peut pas être décrite avec une précision et une certitude totales. Des choix théoriques et descriptifs doivent être faits lorsque l'on cherche à décrire le lexique. Ces choix sont loin d'être évidents.

Considérons maintenant la seconde partie de la définition donnée ci-dessus : le lexique est un ensemble de lexies. On pourrait bien entendu postuler que le lexique est un ensemble de mots-formes, c'est-à-dire de signes lexicaux, et représenter ainsi le lexique du français :

$$\text{Lex}_{\text{français}} = \{ \grave{a}, \dots, \text{avoir}, \text{ai}, \text{a}, \text{avons}, \dots, \text{maison}, \text{maisons}, \dots, \text{petit}, \text{petite}, \text{petits}, \text{petites}, \dots \}.$$

Cette modélisation revient à dire que, par exemple, *maison* et *maisons* sont deux éléments distincts de notre connaissance lexicale du français. Cependant, nous percevons bien que les deux mots-formes en question sont regroupés dans quelque chose de plus général, qui « factorise » tout ce que ces deux signes linguistiques ont en commun. Il s'agit de la **lexie** MAISON. Lorsque l'on apprend un nouvel élément du lexique d'une

langue, on apprend en fait quelque chose qui, potentiellement, est associé à plusieurs mots-formes distincts et qui correspond directement à la notion de lexie. La lexie est vraisemblablement une entité psychique qui structure notre connaissance des langues, une sorte de « métasigne ».

D'un strict point de vue ensembliste, il revient à peu près au même de considérer le lexique comme un ensemble de lexies ou un ensemble de mots-formes. La différence se fait sentir dès que l'on veut modéliser le lexique dans un dictionnaire. Il nous faut alors choisir une unité de description. Les mots-formes associés à une lexie étant en général calculables à partir du radical de la lexie et de règles de grammaire générales, il serait extrêmement redondant de construire un dictionnaire décrivant explicitement tous les mots-formes de la langue. On va donc considérer que le lexique d'une langue donnée est l'ensemble de ses lexies et non l'ensemble de ses mots-formes.

La notion de vocabulaire

La notion de lexique doit être mise en contraste avec celle de vocabulaire :

Le ***vocabulaire d'un texte*** est l'ensemble des lexies utilisées dans ce texte.

Le terme *texte* doit être compris ici dans un sens très large. Ainsi, un texte peut être :

- un texte ou un ensemble de textes ;
- un texte oral ou écrit ;
- un texte impliquant un locuteur unique ou un ensemble de locuteurs.

Texte et vocabulaire d'un texte sont donc du domaine de la parole et non de celui de la langue (cf. Chapitre 1).

J'ai employé systématiquement le terme *vocabulaire d'un texte*, car *vocabulaire* seul est ambigu. En effet, il faut distinguer la notion de vocabulaire d'un texte de celle de vocabulaire d'un individu³ :

Le ***vocabulaire d'un individu*** est le sous-ensemble du lexique d'une langue donnée contenant les lexies de cette langue que maîtrise l'individu en question.

Contrairement au vocabulaire d'un texte, le vocabulaire d'un individu est, en tant que sous-ensemble d'un lexique, une entité théorique. Le vocabulaire d'un individu est une composante de l'***idiolecte*** de cet individu, c'est-à-dire de la langue qu'il maîtrise et parle. En effet, personne ne parle véritablement de la même façon et personne n'a exactement la même connaissance de la langue. De ce point de vue, la langue (comme le lexique) n'existe pas : c'est une abstraction théorique qui synthétise les connaissances communes à l'ensemble des locuteurs. Cela a des conséquences très importantes sur la méthodologie de l'étude linguistique, dans la mesure où l'on ne peut jamais se fonder sur la façon de parler d'un individu pour en déduire une description de la langue en général. Il faut toujours se donner les moyens de prendre en compte les variations idiolectales.

En réalité, le problème de la variation linguistique ne se pose pas seulement au niveau de l'idiolecte. La langue varie en fonction des individus qui l'utilisent mais aussi en fonction des contextes d'utilisation. Nous allons nous attarder un instant sur cet important sujet.

La variation linguistique

On peut isoler au moins cinq axes de variation linguistique pouvant poser problème dans le cadre de l'étude de la langue et, notamment, du lexique : variation géographique, sociale, selon les générations, selon le domaine d'utilisation de la langue et, finalement, selon le mode de communication. Nous allons les examiner à tour de rôle.

Variation géographique C'est un ensemble de différences linguistiques associées à des régions ou à des pays particuliers. Pour ce qui est du

3. Il peut bien entendu aussi s'agir du vocabulaire d'un **ensemble** d'individus.

pas d'originaux dans ce pays. Il ne s'agit pas ici d'un véritable cas de variation **linguistique** géographique, dans la mesure où un français de France n'aura pas d'autre choix que d'utiliser ORIGINAL s'il veut désigner l'animal en question par son nom.

Les quelques remarques qui viennent d'être faites à propos des types de variations géographiques peuvent être transposées pour les autres variations linguistiques qu'il nous reste à examiner.

Variation sociale Il n'est pas rare que l'on puisse déterminer le « milieu social » dans lequel a grandi ou dans lequel évolue un individu en se basant sur la façon dont il s'exprime. Les indices peuvent être lexicaux, grammaticaux ou phonologiques. Au niveau de la description lexicale, on a souvent recours à des marques d'usage telles que **fam** (familier), **vulg** (vulgaire), **soutenu** (style soutenu), **offic** (officiel), etc. pour indiquer que l'emploi d'une lexie donnée est associé à un contexte particulier. Ce type d'étiquetage repose bien entendu souvent sur des présupposés idéologiques. Par exemple, si l'on introduit la marque d'usage **pop** (populaire), quelle différence fait-on véritablement entre du parler populaire et familial ? Pourquoi avoir **pop**, par exemple, et pas **bourg** (bourgeois) ? Mais je n'entrerai pas dans ce débat, me contentant de citer ici quelques exemples peu litigieux :

- CABINET ~ **fam** PETIT COIN ~ **vulg** chiottes ;
- CRIER ~ **fam** [ou **vulg** ?] GUEULER ;
- MOURIR ~ **fam** CASSER SA PIPE ~ **vulg** CREVER ~ **soutenu** PASSER DE VIE À TRÉPAS ~ **soutenu** TRÉPASSER ;
- KLAXON ~ **offic** AVERTISSEUR SONORE.

Ce type de phénomène a été fréquemment exploité par les écrivains, notamment pour provoquer des contrastes humoristiques, comme dans l'extrait suivant de *Zazie dans le métro* :

Une bourgeoise qui maraudait dans le coin s'approcha de l'enfant pour lui dire ces mots :

— Mais, voyons, ma petite chérie, tu lui fais du mal à ce pauvre meussieu. Il ne faut pas brutaliser comme ça les grandes personnes.

— Grandes personnes mon cul, répliqua Zazie. Il ne veut pas répondre à mes questions.

— Ce n'est pas une raison valable. La violence, ma petite chérie, doit toujours être évitée dans les rapports humains. Elle est éminemment condamnable.

— Condamnable mon cul, répliqua Zazie, je ne vous demande pas l'heure qu'il est.

Queneau, Raymond (1959) *Zazie dans le métro*, Collection « Folio », n° 103, Paris : Gallimard, p. 100.

Nous ne nous attarderons pas d'avantage sur ce type de variation linguistique. L'Exercice 3, en fin de chapitre, propose de mener une petite analyse linguistique du texte ci-dessus.

Notons que les deux types de variations qui viennent d'être présentés correspondent à ce que l'on appelle des ***dialectes*** : dialectes régionaux, dans le premier cas, et dialectes « sociaux », dans le second.

Variation selon les générations On ne parle sûrement pas la même langue que nos grands-parents ou que nos petits-enfants (si nous en avons), et cela, même si l'on ne tient pas compte de différences entrant dans la catégorie des variations géographiques ou sociales. Nous avons déjà vu que la langue pouvait s'étudier dans la diachronie, pour modéliser son évolution dans le temps. Cette évolution se manifeste aussi dans le contexte d'une étude synchronique, lorsque justement on doit prendre en compte la façon dont s'expriment les individus appartenant à des générations différentes. Là encore, on peut avoir recours à certaines marques d'usage pour rendre compte de ces variations. Par exemple, au niveau lexical :

- RADIO ~ **vieilli** T. S. F. (pour *transmission sans fil*) ;
- PNEU ~ **vieilli** PNEUMATIQUE.

Variation selon le domaine d'utilisation de la langue Au niveau du lexique, on sera surtout concerné ici par les ***terminologies*** (scientifiques ou techniques). La terminologie est une discipline en soi, ce qui se justifie par le fait que les lexiques terminologiques forment des tous lexicaux entretenant des liens complexes avec le lexique « général ». De plus, l'étude d'une terminologie donnée est normalement indissociable de l'étude poussée du domaine d'activité (informatique, chimie, chirurgie, pharmacie, génie civil, etc.) dont elle relève.

Variation due au mode de communication Il s'agit ici notamment de la distinction entre langue parlée et langue écrite, mentionnée dès le premier chapitre de ce livre (page 7). On peut cependant affiner beaucoup plus en se penchant sur les cas d'utilisation de modes de communication spécifiques : conversations téléphoniques, échanges de

courriers électroniques, etc. Je rappelle qu'un exercice portant sur l'identification des traces de langue parlée dans les courriers électroniques a été proposé à la fin du Chapitre 1 (Exercice 3, page 13).

Nous arrêterons ici cet examen des différents cas de variations linguistiques. On pourrait bien entendu approfondir beaucoup plus la question ; mais elle relève du domaine de l'étude sociolinguistique, domaine que nous ne pouvons qu'effleurer ici.

Remarque sur la définition des notions scientifiques

Nous avons terminé notre examen des notions de lexique et de vocabulaire. Nous allons maintenant passer aux parties du discours : le système le plus courant de classification des unités lexicales. Avant cela, je voudrais faire une remarque importante concernant la façon dont sont définies les notions introduites dans ce cours.



Les notions scientifiques ne peuvent être correctement comprises et définies qu'en considérant les liens qu'elles entretiennent avec d'autres notions. C'est pourquoi une bonne définition de la notion de lexique doit en fait être une définition de la notion de lexique **d'une langue**. C'est aussi pourquoi nous voyons clairement dans les définitions proposées plus haut qu'il existe en fait deux notions distinctes de vocabulaire, selon que l'on parle du vocabulaire **d'un texte** ou du vocabulaire **d'un individu**.

Toutes les définitions données dans cet ouvrage sont élaborées en fonction de cette contrainte. La clé pour la compréhension et l'assimilation d'une définition de ce type est de chercher à comprendre le tout formé par la notion et les autres notions qu'elle met en jeu : il faut se souvenir qu'un lexique est en fait le lexique d'une langue, qu'un vocabulaire est en fait le vocabulaire d'un texte ou d'un individu, etc.

Nous verrons plus loin (Chapitre 8) que cette méthode de définition des notions scientifiques est similaire à celle qui doit être employée pour définir les termes « ordinaires » de la langue.

Les parties du discours

Les *parties du discours* — nom, verbe, adjectif, etc. — sont des classes générales dans lesquelles sont regroupées les lexies de la langue en fonction de leurs propriétés grammaticales.

Remarque terminologique

Deux autres termes fréquemment utilisés dans la littérature linguistique et dans les manuels d'enseignement pour désigner les parties du discours sont *classe grammaticale* et *catégorie syntaxique*. Voyons pourquoi je n'utilise que le terme *partie du discours* dans le présent ouvrage.



Tout d'abord, pour éviter toute confusion, il est toujours préférable de n'utiliser qu'un seul terme pour désigner une notion donnée. Il faut donc faire un choix.

Le terme le plus approprié ici est, à mon avis, *classe grammaticale*, car les regroupements dont nous allons parler sont des classes (pas nécessairement très étanches d'ailleurs) qui regroupent les lexies en fonction de leurs propriétés grammaticales. Malheureusement, ce terme est très peu utilisé dans la littérature et il est toujours souhaitable, dans un cours de base comme celui-ci, de chercher à se conformer au maximum à la pratique courante. C'est la meilleure façon de s'assurer que les notions enseignées pourront être aisément réutilisées lors de l'approfondissement des connaissances (par la pratique d'une discipline ou son étude). Je dois donc ici faire mon deuil de *classe grammaticale*.

Catégorie syntaxique est sans doute le terme le plus fréquemment utilisé en linguistique moderne issue, notamment, de la tradition nord-américaine. Ce terme est cependant problématique, pour au moins deux raisons. Tout d'abord, il ne s'agit pas véritablement ici de catégories au sens d'un ensemble de valeurs mutuellement exclusives. On se souviendra, par exemple, de l'usage de ce terme dans *catégorie flexionnelle* (Chapitre 4, page 53). Les regroupements de lexies que nous considérons ici n'ont pas la systématisme, la rigidité de catégories véritables et le terme plus vague de *classe* me semble donc plus approprié. Mais surtout, ces regroupements ne se font pas uniquement en fonction de

critères syntaxiques. Les critères morphologiques (présence ou non de variation flexionnelle notamment) sont également très importants, même si, effectivement, les lexies vont avant tout être regroupées en fonction des rôles syntaxiques qu'elles peuvent avoir dans la phrase. Le terme *catégorie syntaxique* me semble trop spécifique, surtout dans le cadre de ce cours où, justement, il est important de percevoir la multiplicité des facteurs (sémantiques, syntaxiques et morphologiques) de rapprochement ou distinction des lexies.

Finalement, le terme *partie du discours*, dont l'usage s'est répandu à partir du Moyen-Âge (Latin *partes orationis*), ne veut pas dire grand-chose. (Il s'agissait initialement de désigner les « parties » dont la phrase est constituée.) C'est bien entendu un désavantage par rapport à *classe grammaticale*, mais c'est un avantage par rapport au terme trop spécifique *catégorie syntaxique* (plus justifié, peut-être, dans le cadre d'un cours de syntaxe). Son principal mérite, outre le fait d'être tellement vague qu'il s'applique sans problème à la notion qui nous intéresse ici, est d'être un terme très courant dans la littérature, notamment en grammaire française. Je peux donc l'utiliser sans risque, même si je voudrais insister sur le fait que les parties du discours sont, par définition, des classes grammaticales de lexies. Comme nous allons le voir immédiatement dans ce qui suit, j'y référerai donc toujours en tant que « classes ».

Passons maintenant à la présentation des principales parties du discours, que l'on regroupe traditionnellement selon qu'elles forment des classes ouvertes ou fermées de lexies.

Classes ouvertes de lexies

Une partie du discours est une ***classe ouverte*** de lexies si l'ensemble des éléments qu'elle contient peut varier sans que cela corresponde à une modification importante du fonctionnement de la langue. Les néologismes ou les emprunts appartiennent avant tout à ces classes et ce sont les lexies de ces classes qui tombent le plus facilement en désuétude.

L'ensemble des parties du discours identifiées pour le français varie d'une grammaire à l'autre, que ce soit pour les classes ouvertes ou fermées. On admet cependant généralement que le français, comme la plupart des langues, possède quatre classes ouvertes majeures, que je vais maintenant énumérer en les caractérisant (i) par une série d'exemples et (ii) par leur rôle syntaxique typique dans la phrase⁴ :

1 verbe

ÊTRE^{1.1} [*Je pense donc je suis.*], MANGER, SE CASSER LE NEZ, ... (J'utilise ici la numérotation du *Nouveau Petit Robert* (2001) pour identifier un sens donné de *être*.)

Le verbe se caractérise par le fait que c'est l'élément essentiel de la phrase, celui auquel se rattache (directement ou indirectement) tous les autres éléments (sujet **du verbe**, complément **du verbe**, etc.).

2 nom, aussi appelé **substantif**

NOURRITURE, POMME DE TERRE, IGOR, ...

Le nom est le sujet/complément typique du verbe (même si, bien entendu, un verbe peut avoir un complément non nominal : *Je veux que tu viennes.*).

3 adjectif

GÉNÉREUX, TYRANNIQUE, MAL EMBOUCHÉ, ...

L'adjectif est le modificateur typique du nom (*une arrivée soudaine*).

4 adverbe

TRÈS, LENTEMENT, À TOUTE ALLURE, ...

L'adverbe est le modificateur typique du verbe (*Il arriva soudainement*).

Noter que j'ai pris soin de donner dans toutes les séries d'exemples ci-dessus des lexies qui sont soit des lexèmes soit des locutions. Locution verbale, locution nominale, etc. sont ainsi des sous-classes des parties du discours correspondantes.

4. Les autres classes ouvertes que l'on trouve mentionnées dans la littérature (comme, par exemple, celle des interjections — MINCE !, OH !, ...) peuvent toutes être associées à une de ces quatre classes majeures.

Classes fermées de lexies

Une partie du discours est une **classe fermée** de lexies si l'ensemble des éléments qu'elle contient est stable. (Il n'accepte ni ajouts ni retraits.) De plus, les classes fermées de lexies sont très petites, comparées aux classes ouvertes.

Comme dans le cas des classes ouvertes, il n'y a pas de consensus absolu sur le nombre et sur la dénomination des classes lexicales fermées du français. On mentionne fréquemment les parties du discours suivantes :

- 1 **verbe auxiliaire**
ÊTRE¹V.1 [*Il est attaqué.*], ...
- 2 **pronom**
JE, TU, ...
- 3 **déterminant**
articles : LE¹, ... ; adjectifs démonstratifs : CE, ... ; adjectifs possessifs : MON, ... ; etc.
- 4 **conjonction**
ET, BIEN QUE, ...
- 5 **préposition**
DE, PAR, ...

Chaque classe fermée peut être associée, sur la base des propriétés grammaticales qui caractérisent ses lexies, à une des quatre classes ouvertes majeures mentionnées plus haut : les auxiliaires sont en fait des cas particuliers de verbes, les pronoms des cas particuliers de noms, les déterminants des cas particuliers d'adjectifs (au sens large), les conjonctions des cas particuliers d'adverbes et les prépositions des cas particuliers d'adverbes ou d'adjectifs.

Mots grammaticaux vs mots lexicaux

La distinction entre classes ouvertes et classes fermées de lexies correspond **approximativement** à l'opposition entre **mots lexicaux** et **mots grammaticaux**. Les lexies des classes ouvertes sont en quelque sorte des unités lexicales types, du point de vue de leur comportement en langue et, comme nous le verrons plus tard dans le cours, du point de vue de leur sens. D'où le terme de *mot lexical* fréquemment employé pour les désigner. La plupart des lexies appartenant aux classes fermées

sont, au niveau de leur comportement en langue et au niveau du sens qu'elles véhiculent, intimement liées à la grammaire de la langue (cf. l'utilisation des articles, des pronoms, etc.). Il faut cependant faire très attention à ne pas établir un parallèle trop strict entre ces deux paires de notions. Ainsi, les prépositions appartiennent à une classe fermée mais certaines prépositions semblent bien être des mots lexicaux. Comparons les deux lexies DE^{I.B.1} et DE^{II.1} du *Petit Robert* utilisées dans les exemples ci-dessous⁵ :

- (1) a. *Les salades **de** Lucien sont tendres.* → DE^{I.B.1}
b. *Il parle **de** Jean.* → DE^{II.1}

La préposition DE^{I.B.1} est utilisée en (1a) pour exprimer un sens clairement identifiable et l'exemple en question peut être aisément paraphrasé en remplaçant *de* par une expression sémantiquement équivalente. Je propose ci-dessous deux paraphrases, compte tenu du caractère vague du sens de la préposition en question :

- (2) a. *Les salades **qui appartiennent à** Lucien sont tendres.*
b. *Les salades **que fait pousser** Lucien sont tendres.*

Il n'en va pas de même pour DE^{II.1}, que l'on utilise simplement en (1b) pour lier *parle* à son complément et qui est un mot grammatical par excellence.

Nature grammaticale des parties du discours

Il est très important de toujours garder à l'esprit le fait que l'on regroupe les lexies dans des parties du discours en fonction d'un ensemble très hétérogène de propriétés grammaticales. De plus, les propriétés qui caractérisent les verbes, les noms, etc. varient considérablement d'une langue à l'autre. Par exemple, les noms français se caractérisent notamment par le fait qu'ils possèdent une flexion en nombre (singulier vs pluriel), qu'ils possèdent un genre grammatical (masculin ou féminin) et qu'ils impliquent une détermination (*le chat*, *une idée*, etc.). Les noms anglais, quant à eux, ont aussi une flexion en nombre et la

5. La numérotation de ces lexies est empruntée au *Nouveau Petit Robert* (2001).

détermination mais ne possèdent pas de genre grammatical ; les noms mandarins n'ont ni flexion ni genre grammatical ni détermination ; etc.

Bien que les parties du discours se caractérisent avant tout par des propriétés grammaticales — du type de celles qui viennent d'être données en exemple —, il existe certaines propriétés sémantiques partagées par les noms, les verbes, etc. Mais, comme nous le verrons dans le chapitre suivant lors de l'étude du sens lexical (Section *Prédictat sémantique vs objet sémantique*, page 107 et suivantes), ces propriétés sémantiques communes sont très vagues et ne permettent en aucun cas de caractériser de façon rigoureuse les parties du discours. Il faudra donc se méfier des définitions du type *Les noms désignent des objets, les verbes des actions*. De telles définitions ne sont que des approximations et donnent d'étranges résultats si on les utilise de façon littérale. L'exemple ci-dessous illustre ce problème :

- (3) a. *Il lui déclare son **amour**.*
b. *Ce rocher **pèse** deux tonnes.*

Le nom en (3a) ne désigne pas plus un « objet » que le verbe de (3b) ne désigne une « action ». En effet, le nom AMOUR désigne un sentiment — un état psychique d'un individu — que l'on peut tout aussi bien désigner en utilisant le verbe AIMER :

- (4) *Il lui déclare qu'il l'aime.*

Quant au lexème verbal PESER, utilisé en (3b), il ne désigne pas une action, mais bien une caractéristique. Cette dernière pourrait tout aussi bien être désignée au moyen du nom POIDS, comme le démontre la paraphrase suivante de (3b) :

- (5) *Le poids de ce rocher est de deux tonnes.*

Il faut donc se méfier des caractérisations sémantiques des parties du discours, même si ce sont sans doute elles que l'on entend le plus souvent répétées lorsqu'il s'agit d'ébaucher une définition de ces notions. Bien entendu, la particularisation sémantique des parties du discours a malgré tout un certain fondement. Personne ne niera le fait que le nom prototypique est un nom de « chose » et que le verbe prototypique désigne un « fait ». Ce qu'il faut rejeter, c'est la nature **défini-toire** d'une telle particularisation.

Cette façon de présenter les parties du discours remonte à très loin et a été « officialisée » par la tradition grammaticale française issue de la *Grammaire générale et raisonnée de Port-Royal*, ouvrage écrit au XVII^e siècle. Voici deux courtes citations tirées de la *Grammaire de Port-Royal* qui illustrent ce fait. Je tiens à souligner que mon but est ici de montrer d'où la vision « sémantique » des parties du discours peut tirer ses origines et non de dénigrer un texte dont je pense personnellement qu'il reste d'une lecture tout à fait fascinante, plus de trois siècles après sa rédaction :

Les objets de nos pensées étant, comme nous avons déjà dit, ou des choses, ou des manières des choses, les mots destinés à signifier tant les choses que les manières s'appellent *noms*. [Seconde partie, Chapitre I, page 167]

(...) le verbe, selon ce qui lui est essentiel, est un mot qui signifie l'affirmation. Mais si l'on veut mettre dans la définition du verbe ses principaux accidents⁶, on le pourra définir ainsi : *vox significans affirmationem cum designatione personæ, numeri et temporis* : un mot qui signifie l'affirmation (sic) avec désignation de la personne, du nombre et du temps. [Seconde partie, Chapitre II, page 180]

Antoine Arnaud et Claude Lancelot (1993) *Grammaire générale et raisonnée de Port-Royal*. Réimpression de l'édition de Paris 1846, Genève : Slatkine Reprints.

On pourra mettre en contraste ces définitions avec l'extrait suivant du *Bon usage* de Maurice Grevisse, qui sert très souvent de grammaire de référence pour le français contemporain :

Les listes de parties du discours ont beaucoup varié. La tradition utilisait, selon les catégories⁷, des critères sémantiques (pour le nom, l'adjectif et le verbe) ou des critères syntaxiques (pour la préposition et la conjonction notamment). Le procédé le plus sûr et le plus cohérent est de se fonder sur les critères morphologiques et les critères syntaxiques.

Grevisse, Maurice (1993) *Le bon usage*. Grammaire française refondue par André Goosse, 13^e éd. revue, § 139, Paris/Louvain-la-Neuve : Duculot, p. 178.

6. C'est-à-dire, ce qui peut s'appliquer au verbe mais n'est pas définitoire.

7. Noter l'emploi de *catégorie*. Le *Bon usage* n'établit pas de distinction entre les termes *classe* et *catégorie*, comme le démontre la phrase suivante, qui apparaît dans la grammaire trois paragraphes au-dessus de la présente citation : « On divise les mots en catégories ou classes, qu'on appelle traditionnellement parties du discours. »

Nous en avons terminé avec les regroupements de lexies effectués au moyen des parties du discours. Nous allons maintenant examiner brièvement les liens qui peuvent unir les lexies, ces liens servant à établir d'autres types de classements lexicaux que ceux fondés sur les parties du discours.

Liens entre lexies : le réseau lexical de la langue

Le lexique n'est pas un ensemble « plat » de lexies. Chaque lexie prend sa **valeur** sémantique en langue du fait des liens d'opposition, de similarité, de compatibilité, d'incompatibilité, etc. qui l'unissent aux autres lexies. (Je reviendrai sur cette notion de valeur en langue dans le chapitre suivant.) Le lexique est ainsi un réseau extrêmement riche et complexe d'unités lexicales connectées les unes aux autres.

Il existe deux types majeurs de liens entre lexies, qui ont été identifiés par F. de Saussure dans le *Cours de linguistique générale* (cf. lectures des Chapitres 1 et 2) :

- 1 Les **liens paradigmatiques** connectent les lexies **à l'intérieur du lexique** par des relations sémantiques, éventuellement accompagnées de relations morphologiques. Par exemple, la lexie BARBE est liée paradigmatiquement aux lexies BARBICHE, BOUC (comme types de barbe), BARBU, IMBERBE, GLABRE, BARBIER, POIL, etc.
- 2 Les **liens syntagmatiques** connectent les lexies **à l'intérieur de la phrase** par des relations de combinatoire. Par exemple, la lexie BARBE, toujours elle, s'emploie dans les expressions suivantes⁸ : *grande / longue / forte / épaisse / grosse / ... barbe, se couper / se tailler / se raser / se faire / ... la barbe, avoir / porter une barbe*, etc.



Les quelques exemples donnés ci-dessus illustrent le fait que le lexique est un gigantesque réseau où tout se tient. Il suffit de « tirer » sur une lexie de ce réseau pour que vienne avec elle toute une série d'autres lexies auxquelles elle semble attachée par des liens parfois très subtils.

8. Ces expressions sont des collocations, notion que nous étudierons en détail dans le Chapitre 7.

L'étude des relations lexicales est au cœur du travail en lexicologie et il nous faudra bien entendu revenir en détail sur cette question, notamment dans les Chapitres 7 et 8.

L'accès aux données linguistiques

Si l'on veut comprendre ce qu'est la langue, il faut pouvoir la décrire. Si l'on veut la décrire, il faut pouvoir l'observer. Les remarques qui ont été faites plus tôt dans ce chapitre à propos de la variation linguistique montrent qu'il n'est pas aisé d'identifier ce qu'est la langue et, donc, d'isoler quel est l'objet du travail descriptif en linguistique. En d'autres termes, il faut se poser la question suivante : quelle doit être la source de nos données linguistiques, sachant que ces dernières sont infiniment riches et variées ?

Trois méthodes principales d'accès aux données

En lexicologie (ou, plus généralement, en linguistique), on peut procéder de trois façons pour obtenir de l'information sur la langue, c'est-à-dire pour collecter des données linguistiques :

- 1 La méthode la plus simple, au niveau logistique, consiste à procéder par **introspection**, en tentant de mettre à jour notre propre connaissance linguistique (*Comment est-ce que je dirais ça ?*, *Est-ce que cette phrase est correcte ?*, *Que veut dire cette phrase ?*, ...). Cette façon de faire est bien entendu très limitée et risque de mener à la description de la compétence linguistique d'un individu particulier : nous-même.
- 2 Une méthode, qui semble en apparence plus « scientifique », consiste à mener des **enquêtes linguistiques**, en posant des questions à des locuteurs de la langue. Il ne faut pas se leurrer : une enquête linguistique doit aussi être interprétée, évaluée en fonction de multiples paramètres. Ce n'est donc pas non plus une technique d'accès aux données qui peut prétendre à une objectivité absolue. En fait, une telle technique n'existe tout simplement pas car il faut **toujours** évaluer ce qu'on observe.
- 3 Finalement, on peut procéder par examens de **corpus linguistiques**, qui peuvent être des ensembles de textes littéraires, de textes journalistiques, de transcriptions de dialogues, etc. L'avantage de cette technique est de donner accès à une quantité potentiellement gigantesque de

données, maintenant que l'ordinateur permet de manipuler automatiquement des ensembles de textes contenant des millions et même des milliards de « mots ». La grande fiabilité des logiciels et des équipements modernes de numérisation a permis de construire relativement aisément des corpus à partir de textes qui n'existaient que sous forme imprimée, sans que l'on soit obligé de les retaper entièrement. De plus, l'information textuelle est maintenant presque entièrement créée et stockée sous forme informatique, ce qui provoque un accroissement exponentiel de la masse de corpus potentiellement exploitables pour l'étude linguistique.

La bonne façon de procéder en lexicologie est d'utiliser un mélange des trois méthodes qui viennent d'être décrites. Il faut cependant noter que l'apparition de l'ordinateur comme outil de recherche a donné une place de plus en plus importante au travail sur corpus. L'informatique permet non seulement de stocker d'énormes quantités de textes sur disque, mais elle permet aussi d'analyser ces textes pour en extraire de l'information de façon rapide et systématique. La linguistique et la lexicologie « de corpus » se sont donc considérablement développées au cours des dernières années.

Même si le développement de la lexicologie informatique est un phénomène plutôt récent, le travail sur corpus a permis très tôt de mettre à jour certains phénomènes liés à l'utilisation du lexique dans les textes. On a pu ainsi proposer, dès la fin des années cinquante, un noyau lexical du français — le *Français fondamental* —, à partir de l'identification des vocables à haute **fréquence d'emploi** dans un corpus de référence. La technique mise en œuvre était relativement simple puisqu'il s'agissait de compter le nombre d'**occurrences** de vocables dans ce corpus.

On appelle **occurrence** d'un élément linguistique dans un corpus donné une instance d'utilisation de cet élément dans le corpus en question.

Par exemple, le paragraphe ci-dessus contient deux occurrences du nom ÉLÉMENT.

La recherche qui a été menée sur le *Français fondamental* (initialement appelé *Français élémentaire*) avait une finalité très pratique. Il s'agissait

référence plutôt réduit, j'en conviens, mais qui suffira pour illustrer les notions introduites ici) :

À	2	DU	1	LES	2	QU	1
APPROCHA	1	ELLE	1	LUI	2	QUESTIONS	1
BOURGEOISE	1	ÉMINEMENT	1	MA	2	QUI	1
BRUTALISER	1	ENFANT	1	MAIS	1	RAISON	1
ÇA	1	EST	3	MAL	1	RAPPORTS	1
CE	2	ÊTRE	1	MARAUDAIT	1	RÉPLIQUA	2
CES	1	ÉVITÉE	1	MES	1	RÉPONDRE	1
CHÉRIE	2	FAIS	1	MEUSSIEU	1	S	1
COIN	1	FAUT	1	MON	2	TOUJOURS	1
COMME	1	GRANDES	2	MOTS	1	TU	1
CONDAMNABLE	2	HEURE	1	N	1	UNE	2
CUL	2	HUMAINS	1	NE	3	VALABLE	1
DANS	2	IL	3	PAS	4	VEUT	1
DE	1	JE	1	PAUVRE	1	VIOLENCE	1
DEMANDE	1	L	2	PERSONNES	2	VOUS	1
DIRE	1	LA	1	PETITE	2	VOYONS	1
DOIT	1	LE	1	POUR	1	ZAZIE	2

Index de signifiants lexicaux : citation de « Zazie dans le métro »

L'ordre d'énumération des signifiants lexicaux peut être alphabétique, comme ci-dessus, ou alphabétique inverse. Les signifiants peuvent aussi être classés selon que les caractères sont lus de gauche à droite, ce qui est la façon standard, ou de droite à gauche, pour un classement par terminaisons. Ce dernier type de classement est utile en français quand, par exemple, on s'intéresse à isoler des familles de dérivations suffixales. Ainsi, avec un classement balayant les caractères de droite à gauche, tous les signifiants lexicaux se terminant par *-able*, *-age*, *-eur*, etc. apparaîtront de façon groupée. Par exemple, si l'on fait générer un index trié par terminaisons à partir de notre petit corpus de référence, on

obtient un tableau bien différent, dont je ne ne donne ici qu'un court extrait :

(...)	
JE	1
LE	1
VALABLE	1
CONDAMNABLE	2
ELLE	1
COMME	1
(...)	

On voit que cette méthode de classement permet d'isoler facilement des mots-formes apparentés par la suffixation (ici, *valable* et *condamnable*).

Les signifiants lexicaux ont tous été mis en majuscules dans les index ci-dessus. J'aurais pu tout aussi bien faire générer par le programme que j'ai utilisé des index avec les formes en minuscules, ou même faire respecter la capitalisation originelle du texte. Dans ce dernier cas, cependant, le programme aurait indexé les deux occurrences *condamnable* et *Condamnable* comme correspondant à deux formes distinctes. Ce que le programme considère alors comme une « forme » s'éloigne de notre notion de signifiant lexical.

Finalement, pour conclure sur les index, il convient de remarquer que l'on pourrait vouloir travailler sur des index de **vocables**, où les formes fléchies auraient été identifiées et fusionnées dans une seule entrée de tableau. Par exemple, les deux données suivantes de notre index :

EST 3 ÊTRE 1

seraient remplacées par l'entrée unique :

ÊTRE 4

Pour obtenir ce genre d'index, il faut disposer d'un programme dit de **lemmatisation** qui, grâce à une analyse morphologique automatique, remplace toutes les occurrences de mots-formes dans le texte originel par le nom canonique du vocable correspondant, accompagné de codes indiquant la flexion appliquée à chaque occurrence. Le programme d'indexation peut ensuite tourner en ne s'attachant qu'au nom de vocable pour construire l'index.

Les concordances Examinons maintenant ce qu'est une **concordance** — fréquemment appelée **KWIC**, pour *KeyWords In Context*. C'est une structure de données où chaque occurrence d'un mot-forme dans le corpus est énumérée (généralement suivant l'ordre alphabétique), accompagnée de son contexte d'emploi. Ce dernier est constitué de la suite de caractères apparaissant immédiatement à gauche et à droite de l'occurrence en question. Le contexte est bien entendu ajustable par l'utilisateur du concordancier (le programme de génération de concordances). Voici un extrait d'une concordance produite à partir de notre mini-corpus de référence, avec une (très petite) « fenêtre » de contexte de 70 caractères :

dans le coin s'approcha de l'enfant pour lui dire ces mots :
 Zazie, je ne vous demande pas l'heure qu'il est.
 Ce n'est pas une raison valable. La violence, ma petite chérie,
 bourgeoise qui maraudait dans le coin s'approcha de l'enfant
 ne faut pas brutaliser comme ça les grandes personnes. – Grandes
 doit toujours être évitée dans les rapports humains. Elle est

Si l'on pense au fait que les formes *l'*, *le*, *la* et *les* peuvent correspondre à des signifiant d'article (*le chien*) ou de pronom (*Il le regarde.*), on voit immédiatement l'intérêt d'utiliser une concordance. L'extrait de concordance donné ci-dessus nous permet de voir d'un coup d'œil que seul $LE_{Article}$ et non LE_{Pronom} est utilisé dans notre corpus de référence. Cette information serait bien plus longue à obtenir s'il nous fallait examiner mot par mot le texte brut. On peut imaginer l'énormité de la tâche s'il s'agissait de travailler sur un corpus qui ne serait plus constitué de quelques dizaines d'occurrences mais de plusieurs millions !

Pour conclure sur ce sujet, rappelons que le travail sur corpus a beaucoup évolué depuis ses premiers balbutiements. Non seulement parce que la taille des corpus informatisés a considérablement augmentée, mais aussi parce que les programmes permettant leur gestion et, surtout, leur consultation, se sont sophistiqués.

Voici trois aspects de cette évolution, choisis parmi les plus significatifs :

- 1 Les ordinateurs sont de plus en plus à même de traiter de façon robuste et conviviale d'autres systèmes d'encodage de l'écrit que le seul alphabet anglais (comme c'était le cas il n'y encore pas si longtemps). Graduellement, ce sont tous les caractères utilisés de par le monde qui

peuvent être manipulés de façon simple et uniforme grâce au développement des normes internationales de type ISO (*International Organization for Standardization*) ou Unicode (du consortium Unicode).

- 2 Les index, concordances et autre données extraites des corpus ne sont plus nécessairement générés comme auparavant de façon centralisée, sur un serveur auquel doivent se connecter les utilisateurs. Il est possible de faire tourner les concordanciers sur les ordinateurs individuels et de bénéficier ainsi de plus de souplesse d'utilisation.
- 3 Les index et les concordances sont de plus en plus souvent produits dans des environnements où d'autres ressources d'exploration de corpus sont disponibles. Il s'agit notamment d'analyseurs morphologiques permettant la lemmatisation, comme je l'ai mentionné, mais aussi de programmes statistiques, qui permettent le repérage automatique de certains patrons de cooccurrence de termes ou, plus généralement, de phénomènes statistiques liés à l'utilisation du lexique.

Je vais d'ailleurs maintenant conclure ce chapitre sur une brève présentation de quelques phénomènes de statistique lexicale.

Fréquence d'emploi et autres phénomènes statistiques

La recherche en linguistique quantitative

Il existe des lois de *statistique lexicale* s'appliquant au vocabulaire des textes, lois qui peuvent être exploitées de multiples façons. La statistique lexicale est une discipline de recherche que l'on peut inclure dans une discipline plus générale, appelée *linguistique quantitative*. Comme son nom l'indique, la linguistique quantitative se penche sur l'étude des phénomènes linguistiques quantifiables (nombre d'occurrences de lexies, de patrons syntaxiques, etc. dans les corpus), en se fondant sur des méthodes statistiques.

Les premières applications de la statistique lexicale se sont faites dans le domaine de l'analyse des textes littéraires. Il s'agissait notamment de caractériser le style d'auteurs classiques en fonction des particularités lexicales de leurs textes, d'identifier une évolution de leur style au cours des années, de mettre en évidence des particularités lexicales de certains de leurs textes, etc. Éventuellement, les lois statistiques et les méthodes

développées dans ce contexte peuvent aussi servir à confirmer ou infirmer le fait qu'un auteur présumé a bien la paternité d'un texte donné : il s'agit de techniques d'identification d'auteurs. Bien entendu, pour que les méthodes statistiques puissent prétendre à une certaine efficacité, il faut nécessairement travailler sur des corpus de grande taille. Cela n'aurait aucun sens de vouloir utiliser ces méthodes pour établir, par exemple, qu'un petit quatrain anonyme a bien été écrit par tel auteur du XVII^e siècle !

Certains ont abusé des méthodes statistiques en voulant leur faire dire plus qu'elles ne peuvent. Malgré cela, c'est un domaine d'étude qui est potentiellement très utile pour la recherche linguistique et ses applications pratiques. La statistique lexicale est exploitée maintenant bien au-delà du seul domaine littéraire. Elle trouve des applications notamment dans les logiciels d'aide à la traduction, d'extraction automatique d'informations contenues dans de très larges bases de données textuelles, etc.

Il n'est pas possible d'introduire véritablement les notions de base de la statistique lexicale dans le cadre du présent ouvrage. Je vais me contenter de présenter un type particulier de régularité statistique constatée au niveau de l'utilisation du lexique dans les textes. Ce n'est qu'un exemple parmi d'autres des phénomènes qui peuvent être exploités pour mettre au point des techniques performantes d'analyse statistique lexicale.

Mesure de la richesse lexicale d'un corpus

Le phénomène dont il va être question ici concerne la mesure de la richesse lexicale d'un corpus. Si le corpus en question se réduit à un simple texte de longueur moyenne, on peut bien entendu en répertorier directement tout le vocabulaire. Cependant, la situation est rarement aussi simple. On peut notamment se trouver dans un des trois cas de figure suivants :

- 1 Il arrive fréquemment que l'on veuille examiner le vocabulaire de corpus très vastes, ou même, de corpus dont la taille n'est pas fixe et qui continuent de croître. On peut donc être forcé de n'étudier en détail qu'une partie d'un corpus.

À ces propriétés « visibles » de la courbe, il faut ajouter deux autres propriétés, qui apparaissent à l'examen des données elles-mêmes :

- 3 Lorsque la courbe devient plate, l'accroissement se fait essentiellement avec des *hapax*, c'est-à-dire des signifiants lexicaux n'apparaissant qu'une seule fois dans le corpus.
- 4 On arrive très rapidement à une saturation complète des signifiants de mots grammaticaux, qui sont les lexies à plus haute fréquence : elles apparaissent très vite dans le corpus et l'apparition de nouveaux mots grammaticaux devient un phénomène de plus en plus rare au fur et à mesure que T croît.

Ces observations ont une valeur universelle : elles s'appliquent à tous les corpus, pour toutes les langues. Bien entendu, les statistiques lexicales peuvent être basées sur différents types de comptages ; on peut ainsi s'intéresser à compter les signifiants lexicaux, les mots-formes, les lexies, les vocables ou même, prendre en compte plusieurs types d'entités linguistiques à la fois.

Ceci termine cette brève incursion dans la statistique lexicale. On trouvera, dans la liste de lectures complémentaires, deux textes de référence qui pourront aider le lecteur à mieux se familiariser avec ce domaine d'étude. Ici s'achève aussi ce long chapitre, qui m'a permis de boucler l'introduction de toute les notions de linguistique générale dont nous avons besoin dans ce cours. Nous allons maintenant entrer dans ce qui relève spécifiquement de l'analyse lexicale, notamment dans le domaine de la sémantique lexicale, qui fera l'objet des trois prochains chapitres.

Lectures complémentaires

Perrot, Jean (1968) Le lexique : Grammaire et lexique. In André Martinet (dir.) : *Le langage*, Encyclopédie de La Pléiade, Paris : Gallimard, pp. 283-299.

9. Une droite est asymptote d'une courbe si la distance qui la sépare de la courbe tend vers zéro, lorsque l'on progresse le long de cette droite à l'infini.

À lire pour une présentation générale du lexique, contrasté avec le système de la grammaire. C'est aussi une bonne lecture pour se préparer aux chapitres qui vont traiter spécifiquement de la sémantique, notamment de la sémantique lexicale. On trouvera dans ce texte de nombreuses références aux notions de morphologie qui ont été examinées dans le chapitre précédent.

Palmer, F. R. (1981) Chapter 6, Section 6.3, Grammar and lexicon. In : *Semantics*, Cambridge et al. : Cambridge University Press, pp. 130-135.

Très court texte qui complète utilement Perrot (1968).

Rey-Debove, Josette (1998) 4. Effets des anglicismes lexicaux sur le système du français. In : *La linguistique du signe : Une approche sémiotique du langage*, Chapitre III, Collection U, série « Linguistique », Paris : S.E.S.J.M./Armand Colin, pp. 185-192.

On trouvera dans ce texte (mentionné plus haut, page 65) une perspective intéressante sur l'influence que la présence « massive » d'anglicismes exerce sur le système non seulement lexical mais aussi morphologique et phonologique du français. J. Rey-Debove y voit une menace pour notre langue. Je n'ai pas d'opinion à formuler sur le sujet, mais je pense qu'il est utile de prendre connaissance des arguments avancés ici.

Ducrot, Oswald et Jean-Marie Schaeffer (1995) Sociolinguistique. In : *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris : Éditions du Seuil, pp. 143-148.

À lire pour la notion de variation linguistique. Ce court texte ne donne pas plus d'informations sur la variation lexicale que ce que l'on trouve dans le présent chapitre. Il présente cependant une bonne synthèse du domaine de l'étude sociolinguistique, avec de nombreux pointeurs bibliographiques.

Muller, Charles (1979) La statistique lexicale. In : *Langue française et linguistique quantitative (Recueil d'articles)*, Genève : Slatkine, pp. 229-242.

Voici un excellent texte d'introduction à la recherche en statistique lexicale, écrit par le père de la linguistique quantitative en France. Cet article est assez ancien et a considérablement vieilli pour ce qui est des aspects informatiques de la discipline. Il introduit cependant, de façon limpide et précise, les notions

de base de statistique lexicale, qui sont toujours actuelles. Surtout, il propose une vision pleine de sagesse de l'intérêt **et** des limites de l'utilisation d'outils statistiques pour l'analyse lexicale.

McEnery, Tony and Michael Oakes (2000) Authorship Identification and Computational Stylometry. In R. Dale, H. Moisl and H. Somers (eds.): *Handbook of Natural Language Processing*, New York/Basel : Marcel Dekker, pp. 545-562.

Ce texte est à lire pour une présentation de l'état de l'art en linguistique quantitative appliquée à l'identification de l'auteur d'un texte. Il est intéressant pour nous dans la mesure où les principales techniques utilisées dans ce domaine d'application se fondent avant tout sur des indices stylistiques relevant de la statistique lexicale (plus que sur l'identification de patrons syntaxiques récurrents).

Exercices

- 1 Les langages formels ont eux aussi un lexique. Identifier de la façon la plus exacte possible le lexique du calcul arithmétique simple (le calcul que l'on utilise pour faire ses comptes, remplir ses déclarations d'impôt, etc.).
- 2 Chacune des phrases ci-dessous contient une incohérence, au niveau de l'emploi de la terminologie linguistique. Expliquer.
 - (6) *Cette étude a recensé tout le lexique de « Notre-Dame de Paris », le chef-d'œuvre de Victor Hugo.*
 - (7) *La plupart des lexies du français ont plus d'un sens.*
- 3 Relire l'extrait de *Zazie dans le métro* donné à la page 70. Identifier les indices linguistiques introduits par l'auteur pour marquer la différence d'origine sociale entre la « bourgeoise » et l'« enfant » (c'est-à-dire, Zazie).
- 4 Quelle différence de sens peut-on trouver entre les deux phrases suivantes ?
 - (8) a. *Mon opinion est différente de la sienne.*
 - b. *Mon opinion diffère de la sienne.*

Quelle conclusion peut-on en tirer quant à la caractérisation sémantique des parties du discours ? Trouver d'autres exemples de ce type permettant de tirer le même genre de conclusion.

- 5 Trouver tous les hapax présents dans la citation de *Zazie dans le métro*. Ceci est un exercice que l'on peut faire en deux minutes, pourvu que l'on utilise les bonnes données...



Chapitre 6

Le sens linguistique

Sens linguistique, paraphrase, structure communicative, valeur (du signe linguistique), dénotation / dénoter, référent, déictique, sens logique, valeur de vérité, connotation, évidence linguistique, sens lexical vs grammatical, prédicat sémantique, argument (d'un prédicat), objet sémantique, réseau (ou graphe) sémantique, nœud d'un réseau sémantique.

DUPOND — C'est un projet ridicule !... Et puis, à votre âge, ce serait de la folie !...

DUPONT — Je dirais même plus : ce serait de la folie à votre âge !...

Hergé, Objectif lune

Il n'est pas nécessaire de faire des études, et surtout pas des études universitaires, pour parler une langue. On apprend la langue par imprégnation, en étant en contact avec elle. Ainsi, les personnes qui savent « bien écrire » — quel que soit le sens que l'on veut donner à cette expression — sont généralement des personnes qui lisent ou ont lu beaucoup ; les personnes qui savent raconter des histoires sont généralement des personnes qui en ont beaucoup entendu ; etc. Donc, étudier une langue n'est pas véritablement ce qui permet de **parler** une langue avec aisance. Par contre, pour pouvoir parler **d'**une langue, ou des langues, il faut que l'on nous ait enseigné comment le faire.

Parmi tous les aspects de la connaissance linguistique, celui dont il est sans doute le plus difficile de parler (pour enseigner une langue, pour analyser des textes, etc.) est le sens linguistique. Non pas parce qu'il faut faire appel pour cela à un appareillage théorique très complexe, mais

tout simplement parce que le sens semble aller de soi. On a tendance à oublier qu'il relève d'une abstraction totale, liée au monde « réel » bien entendu, mais qui projette sur ce monde une grille d'analyse en grande partie arbitraire.

Nous allons donc faire ici un premier pas en direction de l'apprentissage de méthodes de description et d'analyse du sens. Je commencerai par définir quelques notions sémantiques élémentaires : sens linguistique, référent, sens logique et connotation. Puis je proposerai un système très général de classification des sens linguistiques. Finalement, j'aborderai la question de la représentation formelle du sens des énoncés, en introduisant le formalisme graphique des réseaux sémantiques.

Notions sémantiques élémentaires

Le sens linguistique

La façon la plus naturelle d'appréhender le **sens** d'une expression linguistique consiste avant tout en une mise en relation de cette expression avec d'autres. Le petit dialogue ci-dessous entre une personne apprenant le français et son professeur illustre cette particularité du sens linguistique :

- (1) — *Qu'est-ce que ça veut dire « passer un savon à quelqu'un » ?*
— *Ça signifie « le réprimander », « le gronder ».*

Pour parler du sens d'une expression, pour le décrire, on met normalement cette expression en relation d'équivalence ou de quasi-équivalence avec une autre expression :

passer un savon à quelqu'un \cong *réprimander quelqu'un*.



$\langle \text{Expression}_1 \rangle \cong \langle \text{Expressions}_2 \rangle$ signifie que les deux expressions sont quasi-équivalentes au niveau sémantique. Le symbole \cong , quant à lui, sera utilisé pour désigner l'équivalence exacte.

Deux expressions linguistiques ayant (approximativement) le même sens sont appelées des **paraphrases**. Il n'y a pratiquement pas d'autre façon naturelle de procéder pour décrire le sens que de faire appel à des paraphrases. C'est d'ailleurs ainsi que fonctionnent les définitions de la

plupart des dictionnaires, comme nous le verrons au Chapitre 9. Cela nous amène à définir le sens linguistique de la façon suivante :

Le **sens** d'une expression linguistique est la seule propriété qu'elle partage avec toutes ses paraphrases.

Cette définition pourrait paraître circulaire dans la mesure où elle revient *grosso modo* à dire que le **sens** d'une expression linguistique est la seule propriété que partage cette expression avec toutes les autres expressions ayant le **même sens**. Cependant, cette circularité n'est qu'apparente : *avoir le même sens* (ou *être une paraphrase*) est, comme *être grammaticalement correct*, une propriété immédiatement perçue par le locuteur, sans qu'il soit nécessaire de la définir précisément. Parce que vous êtes des locuteurs du français, vous pouvez immédiatement dire si les trois phrases françaises (2a-c) ci-dessous sont des paraphrases — si elles ont le même sens —, et cela sans avoir besoin de suivre un cours de linguistique :

- (2) a. *Je pense donc je suis.*
b. *Le fait que je pense démontre que j'existe.*
c. *Ma pensée est la preuve de mon existence.*

Une des caractéristiques de la langue, qui l'oppose à bien des systèmes sémiotiques « artificiels » comme la logique formelle, les langages de programmation, etc., est d'offrir à la personne qui l'emploie un très grand nombre d'options plus ou moins équivalentes pour exprimer un contenu donné. La relation de paraphrase est en quelque sorte une donnée première du sens, quelque chose que nous n'allons pas définir mais que nous prendrons comme un concept primitif permettant de définir la notion de sens elle-même.

Le lien de paraphrase est cependant de nature très complexe, même si sa perception par le locuteur se fait de façon instantanée. Il faut notamment distinguer la paraphrase exacte, qui est finalement assez difficile à obtenir, et différentes variétés de paraphrases approximatives. Le lien de paraphrase renvoie en fait à un continuum ; c'est ce qu'illustrent les exemples suivants, où se manifeste un écart sémantique croissant par rapport à la phrase de référence (3a) :

- (3) a. *Cette pierre est très lourde.*
 b. *Cette pierre pèse très lourd.*
 c. *Le poids de cette pierre est élevé.*
 d. *Cette pierre est difficile à transporter.*

La phrase (3b) peut être considérée comme une paraphrase exacte de (3a), le remplacement de ÊTRE par PESER n'introduisant aucune nuance de sens. Par comparaison, (3c) s'écarte légèrement du sens initial, et cela de deux façons :

- 1 Dire que le poids de la pierre est élevé est moins précis que dire que cette pierre est lourde ; un peu comme si, en disant (3c), le locuteur cherchait justement à éviter d'affirmer que la pierre est clairement lourde.
- 2 La structure de (3c), où POIDS est le sujet grammatical et ÉLEVÉ attribut, diminue l'importance de PIERRE, qui était sujet dans la première phrase : (3a) dit quelque chose à propos d'une pierre alors que (3c) dit quelque chose à propos du poids de cette pierre. Bien entendu, il est clair que ce sont les mêmes informations qui sont communiquées, et donc que l'on a bien affaire à deux paraphrases. Mais une nuance existe tout de même au niveau de la façon dont l'information communiquée est « emballée » dans la phrase. On dira que (3a) et (3c) n'ont pas la même **structure communicative**. La notion de structure communicative des énoncés est fort importante en sémantique, mais elle déborde largement le contexte de l'étude de la sémantique lexicale. Je n'approfondirai donc pas la question ici, me contentant de suggérer une lecture sur ce sujet à la fin de ce chapitre (voir Halliday 1985) ainsi qu'un petit exercice (Exercice 1, page 114).

Enfin, la phrase (3d) n'est pas vraiment une paraphrase de (3a). On peut l'utiliser dans un contexte donné pour transmettre la même « idée » générale, mais le contenu littéral des deux phrases est très différent. Ce qui les unit, c'est un lien logique et non un véritable lien linguistique : si cette pierre est difficile à transporter c'est vraisemblablement parce qu'elle est très lourde ; ou, à l'inverse, si cette pierre est très lourde, elle doit être difficile à transporter. On appelle parfois *paraphrase conceptuelle* le type de lien unissant (3a) à (3d), par opposition à la « vraie » paraphrase, appelée *paraphrase linguistique*.

La notion de paraphrase sur laquelle je me suis appuyé plus haut pour définir le sens est bien la paraphrase linguistique véritable. Cette approche de la définition du sens est tout à fait compatible avec la façon dont le sens est décrit dans les dictionnaires de langue (voir le Chapitre 9, sur la lexicographie) et avec la notion de *valeur* du signe linguistique proposée par F. de Saussure (voir le texte de J. Picoche donné dans les lectures pour le présent chapitre). Ainsi, le sens d'une lexie se conçoit en fonction du rapport qu'elle entretient dans le réseau lexical de la langue avec d'autres lexies ayant un sens plus ou moins équivalent ou entretenant une certaine relation de sens avec elle.

Prenons un exemple simple pour illustrer ce fait. Si quelqu'un arrive chez un ami en disant :

(4) *J'ai garé mon véhicule devant la porte.*

son interlocuteur va peut-être penser qu'il est venu à bord d'un véhicule un peu spécial. Parce qu'il existe en français des noms courants pour désigner des véhicules spécifiques (VOITURE, CAMION, MOTO, etc.), l'emploi de VÉHICULE dans (4) n'est pas perçu comme neutre. En disant *véhicule*, c'est comme si le locuteur avait volontairement évité de dire *voiture*, *camion*, ... On voit donc que le sens de VÉHICULE, comme le sens de toute lexie de la langue, est perçu par les locuteurs non de façon autonome, mais en relation avec le sens d'autres lexies que la langue met à notre disposition. C'est à ce phénomène que renvoie la notion de valeur du signe linguistique.

Nous en avons terminé avec la présentation générale de la notion de sens. Je vais conclure par deux courtes remarques de nature terminologique :

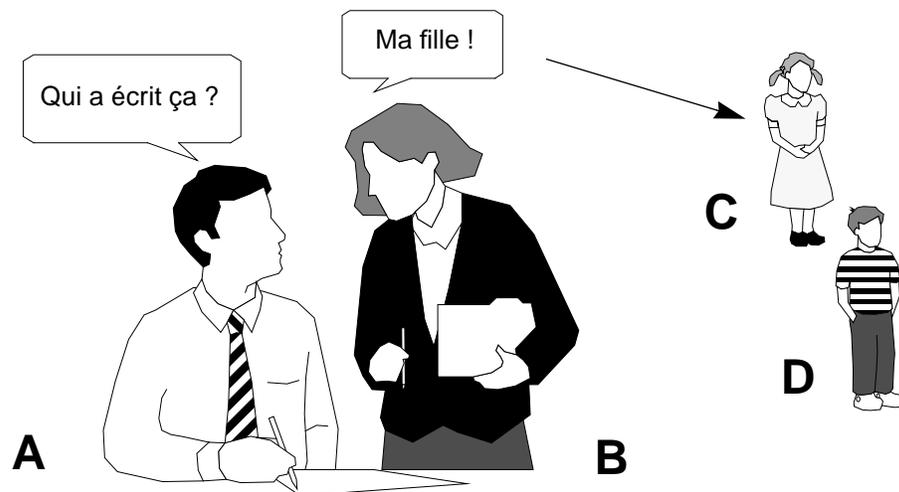
- 1 Il faut noter que le sens est à l'expression linguistique ce que le signifié est au signe linguistique. On va parler du *sens d'une expression linguistique* et du *signifié d'un signe*, mais ces notions renvoient toutes deux à un même type d'entité : un contenu informationnel.
- 2 Le terme de *dénotation* est souvent utilisé en linguistique soit comme un équivalent pour *sens linguistique*, soit au contraire pour désigner une notion séparée du sens. On trouvera à la fin de ce chapitre la référence d'un texte de J. Lyons qui établit une distinction terminologique entre sens linguistique et dénotation. Je n'aurai pas à faire usage de cette distinction dans le présent ouvrage. Par contre, il m'arrivera d'employer

le verbe **dénoter** dans le sens suivant : X dénote $Y \equiv$ La lexie ou l'expression linguistique X sert, de par son sens, à désigner Y [Ex. Les lexies *DÉSIR*, *ANGOISSE* et *SOULAGEMENT* dénotent toutes des sentiments.].

Le référent

Une expression linguistique est, sur le modèle du signe, une association entre un sens (le contenu véhiculé par cette expression) et une forme (orale ou écrite). Cependant, lorsqu'une expression linguistique est utilisée dans la parole par le locuteur ou perçue par le destinataire, elle fonctionne généralement en pointant vers un élément de la « réalité », que l'on appelle le **référent** de l'expression.

Pour bien comprendre la différence entre le sens d'une expression linguistique et son référent, prenons un cas concret, illustré par le dessin ci-dessous :



Lorsque la personne B répond à la question de A par — *Ma fille !*, elle utilise une expression linguistique (une phrase qui est syntaxiquement un groupe nominal) au moyen de laquelle elle désigne une autre personne impliquée dans cette situation : la petite fille C.

Doit-on considérer que l'idée associée à la phrase *Ma fille !*, son contenu, est la fillette C ? La réponse est **non** car n'importe qui peut utiliser cette phrase pour désigner un individu autre que C, pour peu que cette personne soit sa fille. Or c'est bien la même expression qui est

utilisée à chaque fois ; c'est la même association entre un sens et une forme donnés.

On peut décrire le sens de la phrase énoncée par B au moyen de la paraphrase suivante : 'L'enfant de sexe féminin dont je suis un des parents'. C'est ce sens que l'on va toujours exprimer par *Ma fille !* lorsque l'on parle de quelqu'un et c'est ce sens qui est linguistiquement associé à la phrase en question. La personne qui peut être désignée en utilisant *Ma fille !* n'est donc pas le sens de cette phrase. C'est un élément externe, impliqué dans une situation donnée d'utilisation de l'expression linguistique considérée : son référent.

Le **référent** d'une expression linguistique est un élément du « monde » que cette expression permet de désigner dans un contexte donné d'utilisation.



Le sens appartient à la langue alors que le référent relève de la parole : ce n'est que lorsque l'on considère une instance particulière d'utilisation, de manifestation, d'une expression linguistique que l'on peut identifier un référent donné. Notons que, lorsque l'on décrit les sens lexicaux, on décrit en même temps, d'une certaine façon, leurs référents potentiels. Ainsi, lorsque le *Nouveau Petit Robert* définit le sens de GUERRE par

Lutte armée entre groupes sociaux, et spécialement entre États, considérée comme un phénomène social.

il caractérise en même le type de situations concrètes que l'on va pouvoir appeler *des guerres*.

Pour conclure sur la notion de référent, examinons le cas particulier de certaines expressions appelées **déictiques**¹. Ce sont des expressions dont le sens ne peut se décrire qu'en mentionnant une « entité » impliquée dans la situation de communication. Voici trois cas, choisis parmi les plus typiques :

- 1 Le pronom de première personne MOI a pour sens la seule désignation du locuteur (la personne qui dit *moi* ou *je*).
- 2 Le pronom de deuxième personne TOI a pour sens la seule désignation du destinataire (la personne à qui le locuteur dit *toi* ou *tu*).

- 3 L'adverbe de temps DEMAIN a pour sens 'le jour qui suit le moment où je parle' [*Je viens demain* ≡ *Je viens le jour qui suit le moment où je parle*], c'est-à-dire que son sens implique la désignation du moment où la phrase contenant *demain* est énoncée.

Les déictiques sont très intéressants dans la mesure où ils sont l'illustration de signes linguistiques intermédiaires : des symboles qui sont aussi en partie des indices. Il faut se souvenir ici de la notion de signe indiciel, introduite au Chapitre 2 (page 19) : un indice est un signe qui implique une relation de proximité entre sa manifestation et le contenu qu'il exprime. Or, un déictique ne prend son sens véritable que dans un contexte de communication donné, là où existe véritablement un locuteur, un destinataire et un moment de la parole. C'est donc un signe hybride. J'avais mentionné au Chapitre 2 (page 22) le cas des onomatopées, signes linguistiques, donc symboliques, qui sont en même temps fortement iconiques. Les déictiques nous donnent maintenant une illustration de l'autre type d'hybridation : des signes linguistiques qui fonctionnent en partie comme des indices.

Les déictiques pronominaux MOI et TOI sont les exceptions qui confirment la règle, pour ce qui est de la distinction fondamentale entre sens et référent. Ces lexèmes bien particuliers semblent avoir un sens qui coïncide en fait avec leur référent en parole. Plus précisément, un lexème comme MOI ne peut pas être défini par une paraphrase : la seule façon de procéder pour rendre compte de son sens est d'employer une expression comme *La personne qui parle en ce moment* ou, encore mieux, de montrer du doigt le locuteur.

Je me suis beaucoup attardé sur la distinction entre sens et référent parce qu'elle pose souvent problème aux personnes non entraînées à mener des analyses sémantiques. La tendance générale, lorsque l'on parle du sens des énoncés (et donc des lexies), semble être de faire l'amalgame entre sens et référent. La raison en est vraisemblablement que le sens est une entité totalement abstraite, qui se laisse difficilement appréhender. La confondre avec le référent donne une fausse impression de compré-

1. Le cas des déictiques a déjà été brièvement mentionné dans le corrigé de l'Exercice 2 du Chapitre 1 (page 200, à la fin de l'ouvrage). Vous trouverez la référence d'un court texte de R. Jakobson sur ce sujet dans la liste des lectures complémentaires pour le présent chapitre.

hension, parce que c'est une façon commode d'ancrer le sens dans la « réalité ». Cependant, la seule façon de mettre en lumière le sens d'une expression linguistique, c'est toujours de la paraphraser. Nous reviendrons en détail sur ce point essentiel dans le Chapitre 8, sur la définition lexicale.

Je vais maintenant dire quelques mots sur un approche « concurrente » de la modélisation du sens : l'approche logique.

Le sens logique (ou valeur de vérité)

Il ne faut pas confondre le sens linguistique avec l'interprétation logique que peuvent recevoir les expressions linguistiques — ce qu'on pourrait appeler leur *sens logique*. En effet, d'un point de vue strictement logique, l'interprétation du sens se ramène à une interprétation en terme de *valeurs de vérité* : vrai ou faux. Ainsi, deux propositions ont le même sens (logique) si elles ont la même valeur de vérité. Il n'existe pas d'autre façon de comparer le sens logique de deux propositions que de regarder si elles sont toutes deux vraies, toutes deux fausses ou si l'une est vraie et l'autre est fausse.

Lorsque l'on adopte le point de vue de la logique pour décrire le sens des énoncés, on ne peut plus rendre compte du sens linguistique puisque dans un contexte de parole donné, deux phrases peuvent tout à fait être vraies sans pour autant être des paraphrases — c'est-à-dire, sans avoir le même sens linguistique :

- (5) a. *Vous êtes en train de lire un cours de lexicologie.*
b. *Vous comprenez le français.*

Ces deux phrases sont vraies dans le présent contexte ; elles ont donc le même sens logique. Mais il est clair qu'elles n'ont pas du tout le même sens linguistique. Bien entendu, pour utiliser le sens logique en sémantique, on dira que ce qui distingue les phrases (5a) et (5b) c'est le fait qu'elles ne sont pas substituables dans tous les contextes en gardant la même valeur de vérité. Si je vois dans la rue quelqu'un parlant français avec un ami et que je lui dis — *Vous comprenez le français*, la phrase que j'énonce se trouve être vraie ; ce n'est pas le cas si j'énonce — *Vous êtes en train de lire un cours de lexicologie*. On arrive donc à montrer qu'il existe une différence de sens entre deux expressions en

ayant recours à la notion de sens logique. Cependant, cette façon de procéder n'offre pas une véritable description, une explicitation, du sens linguistique. De plus, on ne parviendra jamais par cette méthode à démontrer une identité de sens linguistique, puisqu'il faudrait pour cela être capable de tester l'équivalence de sens logique dans tous les contextes d'énonciation possibles et imaginables. Le fait de ne pas pouvoir isoler une quelconque différence de sens logique, dans le cas d'une paire de phrases donnée, ne permet donc pas de déduire que les deux phrases en question ont le même sens linguistique. Pour cela, il faut avoir recours à la perception intuitive du lien de paraphrase, comme je l'ai dit au début de ce chapitre.

L'approche logique permet de faire bien des choses dans le contexte de la modélisation des phénomènes linguistiques, et il est très important de pouvoir s'en inspirer. Cependant, le système de la logique formelle est avant tout un outil de modélisation des différents types de **raisonnements** — à travers la notion d'inférence logique — et non un outil fait pour la modélisation du sens linguistique. Nous ne nous attarderons donc pas plus ici sur la logique formelle.



Le fait que ce cours ne comprenne pas de véritable introduction aux notions de la logique formelle ne signifie pas que l'on puisse faire l'économie de celles-ci. Je pense au contraire qu'elles sont trop importantes et trop utiles dans le cadre d'études poussées en linguistique (et dans bien d'autres disciplines) pour pouvoir être présentées de façon succincte. Je recommande donc vivement au lecteur de cet ouvrage de se familiariser, si ce n'est déjà fait, avec les notions de base de la logique.

Pour information, je signale le texte suivant, que je trouve excellent du fait de sa relative simplicité et, en même temps, de sa grande profondeur :

Grize, Jean-Baptiste (1967) Logique : Historique. Logique des classes et des propositions. Logique des prédicats. Logiques modales. In Jean Piaget (dir.) : *Logique et connaissance scientifique*, Encyclopédie de La Pléiade, Paris : Gallimard, pp. 135-289.

Le sens et son rapport au monde

Les sens linguistiques « dressent une carte » du monde tel que nous le percevons :

What conditions need to be met for the signs of language, limited in number, to designate reality, which is infinite? The first condition is that reality must be segmented. Whenever we manipulate an object we separate it from its environment. Part of the act of separating it is the act of naming it: a cumulus cloud, a wall, a stick, a laugh. Language gives us a map of reality in which everything is covered but much detail is left out.

Bolinger, Dwight (1968) *Aspects of Language*. New York *et al.* : Harcourt, Brace & World, p. 221.

D. Bolinger nous dit deux choses dans ce paragraphe :

- 1 Le signes linguistiques entretiennent un lien étroit avec le « monde » par le biais de leur sens même si, comme nous l'avons vu dans la section sur le référent, le sens d'une expression linguistique ne doit surtout pas être confondu avec le segment du monde que cette expression désigne en parole.
- 2 L'ensemble des sens linguistiques représente une grille d'analyse qui façonne notre perception du monde.

Il convient d'ajouter à cela que les lexies de la langue influencent la façon dont percevons le monde non seulement à travers leur sens, mais aussi à travers leurs possibles connotations :

Une **connotation** est un contenu informationnel associé à une lexie qui, contrairement au sens, n'est pas nécessairement exprimé quand cette lexie est utilisée.

Ainsi, TIGRE connote en français la férocité, ce qui se manifeste dans des expressions comme *féroce comme un tigre* ou *se battre comme un tigre* (en parlant d'un soldat, d'un maquisard, etc.). On peut cependant tout à fait parler d'un tigre poltron, qui se cache dans les taillis au moindre bruit. La férocité ne fait donc pas partie du sens de TIGRE, à la différence de 'animal', qui est une composante de ce sens². En disant :

(6) *J'ai vu un tigre.*

on dit **nécessairement** que l'on a vu un animal, mais pas nécessairement que c'était un animal féroce.



Les expressions du type *féroce comme un tigre* et *se battre comme un tigre*, mentionnées ci-dessus, sont des **évidences linguistiques** démontrant que la lexie française TIGRE connote la férocité. Il est essentiel de pouvoir présenter de telles évidences pour supporter l'identification d'une connotation ou, plus généralement, d'une caractéristique sémantique d'une lexie que l'on cherche à décrire.

Cela conclut ma présentation de la notion de sens et des autres notions qui y sont directement attachées. Je vais maintenant examiner le problème de la classification des sens, en restant à un niveau très général. Nous verrons d'abord la distinction entre deux grandes familles de sens linguistiques : sens lexicaux *vs* sens grammaticaux. Nous examinerons ensuite la subdivision des sens lexicaux en prédicats sémantiques *vs* objets sémantiques.

Classification des sens linguistique

Sens lexical *vs* grammatical

On peut distinguer deux types de sens contenus dans les ressources sémantiques de chaque langue :

- 1 les **sens lexicaux**, qui sont généralement exprimés par des lexies de la langue et se décrivent assez bien au moyen des définitions standard des dictionnaires ;
- 2 les **sens grammaticaux**, qui ne sont pas associés aux lexies de la langue — sauf dans le cas des mots grammaticaux (cf. Chapitre 5, page 76) — et qui peuvent difficilement être décrits autrement qu'en faisant référence à la grammaire de la langue.

Pour bien comprendre la différence entre sens lexical et sens grammatical, je suggère de comparer la relative facilité avec laquelle on peut

2. Nous reviendrons en détail sur la notion de composante du sens d'une lexie ou de sa définition aux Chapitres 7 et 8.

paraphraser un lexème comme VOISIN et la quasi-impossibilité de faire la même chose avec l'article LE³ :

- (7) a. *son voisin* \equiv *la personne qui habite près de chez lui*
b. *le chat* \equiv ? *chat auquel on pense dans la présente situation*⁴

Dans toute langue, les sens lexicaux constituent l'écrasante majorité des sens disponibles. Ils sont, par excellence, les sens que l'on cherche à communiquer. Par contraste, les sens grammaticaux sont en nombre très réduit (variable selon les langues) et leur expression nous est imposée par la langue. En lexicologie, nous nous concentrons bien entendu sur l'étude des sens lexicaux.

Prédicat sémantique vs objet sémantique

On considère habituellement deux grandes classes de sens lexicaux :

- 1 Les **prédicats sémantiques** sont des sens de lexies qui dénotent des faits ou des entités impliquant au moins un « participant » appelé **argument** (du prédicat). Les arguments d'un prédicat sont habituellement désignés par des variables du type *X, Y, Z*, etc. — 'X mange Y', 'X donne Y à Z', '[X est] petit', 'amour de X pour Y', 'nez de X', ...



Une remarque doit être faite à propos de ce tout dernier exemple. On associe généralement les lexies prédicatives à des faits (actions, événements, états, etc.). Il est cependant très important de garder à l'esprit qu'elles peuvent aussi dénoter des entités. Ainsi, la lexie NEZ, comme toutes les lexies dénotant des parties du corps, est un exemple de lexie prédicative dénotant une entité (et non un fait). Un nez est nécessairement le nez de quelqu'un et la mention de la personne à qui appartient le nez — *le nez de Cléopâtre* — permet d'exprimer l'argument de ce prédicat (un participant nécessaire de la situation « avoir un nez »).

3. Le lecteur pourra aussi regarder la façon dont son dictionnaire favori définit le sens de ces deux lexèmes.

4. Cette tentative désespérée de paraphrasage se fonde bien entendu sur l'opposition *le chat* vs *un chat*.

- 2 Les **objets sémantiques** sont des sens de lexies qui dénotent des entités n'impliquant intrinsèquement aucun participant — 'tomate', 'sable', 'Émile Zola', ...

Les prédicats sémantiques types sont des verbes. En fait, un verbe est nécessairement un prédicat sémantique. Cependant, les adjectifs et les adverbes sont eux aussi des prédicats. Des lexèmes comme GROGNON ou MÉCHAMMENT signifient nécessairement '[quelqu'un est] grognon' et '[quelque chose se produit/est fait] méchamment'.

On voit donc que le fait d'être un prédicat est une propriété sémantique de ces trois parties du discours. Cependant, les noms peuvent aussi être des prédicats ; ils le sont même très fréquemment. Ainsi, lorsqu'on emploie le lexème AMOUR, on sous-entend, cf. (8a) ci-dessous, ou on exprime explicitement, cf. (8b), deux arguments de ce prédicat : celui qui éprouve de l'amour et la personne pour laquelle ce sentiment est éprouvé :

- (8) a. *C'est un amour platonique.*
b. *L'amour de Léonce pour Justine est platonique.*

Même si ni Léonce ni Justine ne sont mentionnés dans (8a), on sait que deux « participants » sont nécessairement impliqués dans la situation dont il est question. Cette information nous est communiquée par le sens même du lexème AMOUR, dont une caractéristique est d'être un prédicat à deux arguments. On peut ainsi contraster 'amour' avec un autre prédicat nominal, à quatre arguments celui-là :

- (9) a. *La vente s'est faite dans la matinée.*
b. *La vente par l'agence[= X] d'une maison[= Y] au client impatient[= Z] pour une somme astronomique[= W] s'est faite dans la matinée.*

Comme on le verra plus loin dans le cours, lors de l'examen de la structure des définitions lexicographiques (Chapitre 8), il est impossible de décrire correctement un sens lexical sans considérer sa nature de prédicat sémantique. Plus précisément, si un sens est un prédicat, il importe de déterminer combien d'arguments il contrôle pour pouvoir dégager ce que nous appellerons les « composantes » de sa définition. De plus, certains phénomènes de combinatoire lexicale — comparer [X]

effectue une vente, [Y] est l'objet de la vente et [W] est le montant de la vente — ne peuvent être bien modélisés qu'une fois mise en évidence la nature prédicative de la lexie.

Importante mise en garde sur la notion d'argument

Il est assez fréquent de voir mises **en opposition** les deux notions de prédicat et d'argument. J'ai ainsi très souvent entendu des questions du type : *Est-ce que ce sens est un prédicat ou un argument ? ; Dans cet exemple, combien y a-t-il de prédicats et combien d'arguments ? ; etc.*



Il me semble difficile de répondre à de telles questions puisque les deux notions en cause ne renvoient pas à des phénomènes comparables :

- Être un prédicat est une propriété intrinsèque d'un sens en langue.
- Être un argument (d'un prédicat) est une « fonction sémantique » d'un sens (prédicat ou objet sémantique) dans un message linguistique donné. Ce n'est aucunement une propriété intrinsèque de ce sens.

Demander si un sens est un prédicat ou un argument est aussi bizarre que de demander si une lexie est un nom commun (propriété intrinsèque de la lexie dans le lexique) ou un complément d'objet direct (fonction grammaticale que peut avoir la lexie dans une phrase donnée).

Pour bien enfoncer le clou, prenons un exemple concret :

(10) *Léo veut rencontrer Lida.*

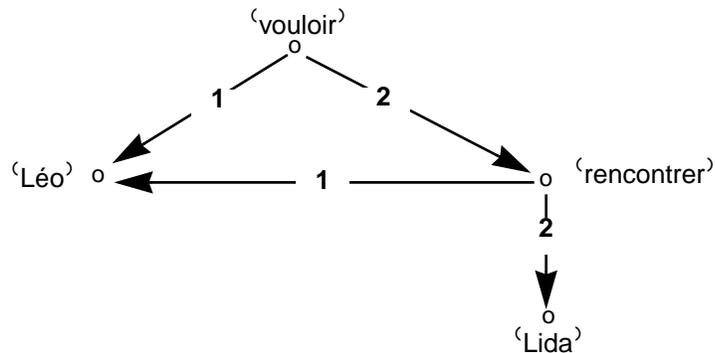
Le sens 'rencontrer' est un prédicat à deux arguments ('X rencontre Y'). Dans (10), ses deux arguments sont 'Léo' et 'Lida'. Mais, dans cet exemple, il se trouve aussi être lui-même le second argument du prédicat 'vouloir', dont le premier argument est 'Léo'. L'argument d'un prédicat peut donc tout à fait être lui-même un prédicat !

Représentation formelle du sens des énoncés

Il est utile de disposer de moyens formels de visualisation du contenu des messages linguistiques lorsque, comme avec l'exemple (10) ci-dessus, on se met à analyser des configurations complexes de sens. Il

existe pour ce faire un formalisme graphique très puissant appelé **réseau sémantique** (ou **graphe sémantique**).

Ainsi, la configuration de sens lexicaux exprimée dans l'exemple (10) ci-dessus peut être visualisée au moyen du réseau sémantique suivant :



Dans cette figure, une flèche ('X—*n*→Y') indique que le sens ('Y') est le *n*^{ième} argument du prédicat ('X'). Un réseau sémantique est donc une figure constituée de flèches, représentant des liens prédicat-argument, qui connectent les sens lexicaux exprimés dans la phrase correspondante. Les points connectés par les flèches et étiquetés par les sens lexicaux sont appelés **nœuds** du réseau.



Il est aussi possible de représenter dans un réseau sémantique les sens **grammaticaux** exprimés dans une phrase (temps grammatical, singulier vs pluriel des noms, etc.). Pour simplifier ma présentation, je ne tiendrai cependant pas compte ici de ce type de sens.

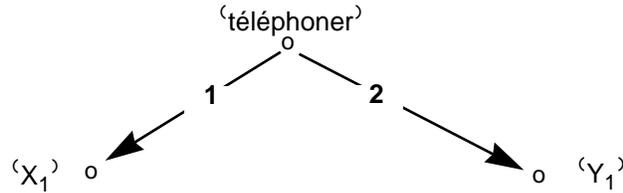
Le formalisme des réseaux sémantiques permet de mettre clairement en évidence la structure sémantique des phrases. Ainsi, la figure ci-dessus rend explicite le fait qu'en (10) le prédicat 'rencontrer' est le second argument du prédicat 'vouloir'.

Voyons à partir d'un autre exemple simple comment il faut procéder pour représenter le contenu d'une phrase au moyen d'un réseau sémantique⁵. Soit la phrase suivante :

(11) *Léo téléphone souvent à son ami José.*

La meilleure façon de commencer notre analyse sémantique est d'identifier quel est le prédicat central de la phrase, celui autour duquel

grave tout le message exprimé en (11). C'est bien évidemment 'téléphoner', qui est un prédicat à deux arguments :



Je mets ici des indices sur les variables parce qu'il va me falloir nommer plus bas les arguments d'autres prédicats. J'utiliserai donc $X_1, X_2, \dots, Y_1, Y_2, \dots$

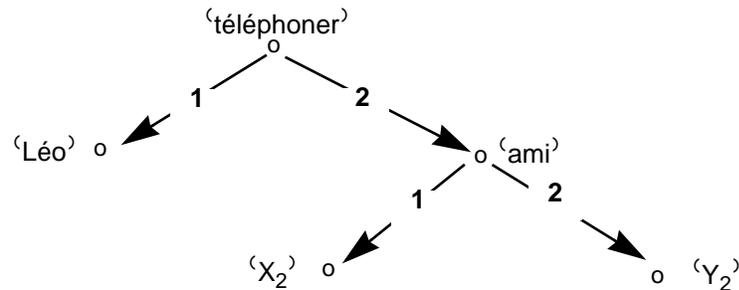
Qui est X_1 ? La personne qui téléphone est désignée en (11) par le nom propre LÉO, qui est un objet sémantique. On peut donc remplacer ' X_1 ' dans la figure ci-dessus par 'Léo'.

Qui est Y_1 ? On pourrait être tenté de dire que c'est José et donc remplacer ' Y_1 ' par 'José' dans notre réseau en construction. Cependant, la phrase n'est pas *Léo téléphone souvent à José, qui est son ami* mais *Léo téléphone souvent à son **ami** José*. Ce qui est communiqué par la phrase, c'est le fait de téléphoner à un ami, caractérisé comme étant José, et non le fait de téléphoner à José, caractérisé comme étant l'ami de Léo. Le sens 'ami' est donc le second argument de 'téléphoner'.

Maintenant, il faut faire attention au fait que 'ami' n'est pas un objet sémantique. C'est un prédicat à deux argument : ' X_2 est l'ami de Y_2 '.

5. Je ne tiens compte, dans ces représentations sémantiques, que des connections prédicat-argument. Le formalisme des réseaux peut être étendu pour modéliser, notamment, la structure communicative des phrases (voir page 98). Je n'entre pas ici dans de telles considérations puisque l'étude de la structure communicative dépasse le cadre du présent ouvrage.

Dessignons donc, pour y voir plus clair, la configuration de sens que nous avons identifiée jusqu'à présent :



On remarquera qu'aucun nœud sémantique n'est associé à la préposition *à* (*téléphoner à quelqu'un*). En effet, cette dernière est un mot grammatical, qui n'a donc pas de sens lexical. Elle est utilisée ici simplement pour établir la connexion syntaxique entre TÉLÉPHONER et son complément AMI. Son emploi est imposé par les propriétés de combinatoire du verbe TÉLÉPHONER. Il ne résulte pas du besoin d'exprimer un sens particulier, qui serait associé à *À*. Si l'on choisit d'utiliser le lexème APPELER au lieu de TÉLÉPHONER, on fait l'économie de l'emploi de cette préposition, puisque la combinatoire de APPELER nous indique que ce verbe réalise son second argument par un complément d'objet **direct** : *Léo appelle José*. Revenons maintenant à notre analyse.

La valeur de $\langle X_2 \rangle$ nous est fournie par le pronom SON, qui réfère ici au nom LÉO. On peut donc faire coïncider le nœud étiqueté $\langle X_2 \rangle$ avec le nœud 'Léo'. Cela se fait tout simplement en effaçant $\langle X_2 \rangle$ et en reconnectant au nœud 'Léo' la flèche qui le liait à 'ami'. Quant à la valeur de $\langle Y_2 \rangle$, c'est évidemment 'José'. (José est l'ami auquel Léo téléphone souvent.)

Il ne nous reste plus qu'à régler le cas du sens du lexème SOUVENT. Ce dernier est un adverbe et, comme tous les adverbes, c'est un prédicat sémantique : ' X_3 a lieu souvent'. Bien entendu, ce sont les appels téléphoniques qui sont fréquents ici et, donc, 'téléphoner' est l'argument

cupations dans le présent cours, et l'organisation sémantique de la phrase, que nous n'étudierons pas véritablement. L'approche linguistique sur laquelle s'appuie M. Halliday est la théorie de la grammaire systémique fonctionnelle, dont il est le créateur. Les notions qu'il présente ici se retrouvent cependant dans la plupart des écoles linguistiques, notamment dans les approches relevant de ce qu'on appelle de façon très générale le structuralisme européen.

Lyons, John (1978) Chapitre VII. Référence, sens et dénotation. In : *Éléments de sémantique*, Paris : Larousse, pp. 143-186.

Il est nécessaire de maîtriser de très nombreuses notions de base pour pouvoir mener à bien l'étude du lexique. Ces notions seront utilisables dans le cadre d'applications de la lexicologie uniquement si elles forment un **système**. Elles doivent être connectées et se compléter mutuellement. Voilà pourquoi j'ai choisi de limiter au maximum dans le présent ouvrage la discussion des différentes approches linguistiques et des terminologies dont elles font usage. Je pense qu'il est préférable de commencer par bien maîtriser un tout notionnel cohérent, pour ensuite le relativiser en le confrontant à des approches complémentaires ou contradictoires. De ce point de vue, la référence ci-dessus complète très bien le présent chapitre. Elle discute de façon assez détaillée les différentes terminologies linguistiques liées à la notion de sens.

Jakobson, Roman (1963) 1. Embrayeurs et autres structures doubles. In : *Essais de linguistique générale*, Chapitre IV, Collection « Arguments », Paris : Éditions de Minuit, pp. 176-181.

C'est le texte de référence qui est peut-être le plus souvent cité à propos de la notion de déictique. Les déictiques sont ici appelés *embrayeurs*, une traduction du terme anglais *shifter* utilisé par l'auteur dans la version originale. On trouvera dans ce texte une présentation claire des déictiques selon une perspective sémiotique, en tant que signes linguistiques à caractère indiciel. Jakobson emploie le terme d'*index* — les embrayeurs sont des symboles-index —, suivant en cela la terminologie du sémioticien américain Charles Peirce.

Exercices

- 1 Se reporter à la citation donnée en exergue au tout début de ce chapitre. Est-ce que Dupont dit véritablement « plus » que Dupond dans sa réplique ? Penser aux notions de paraphrase et de structure communicative.

- 2 En français québécois, un sacre est un blasphème, c'est-à-dire un juron qui offense le sentiment religieux. Essayer d'expliquer, en s'appuyant sur les notions de sens *vs* référent, ce qui est comique dans la façon dont Michel Tremblay, jeune enfant, réplique à sa mère à la fin de la citation suivante :

— J'veux pus lire ! Jamais !

— Voyons donc ! Y veut pus lire ! T'es rien que rendu à page neuf de ton premier livre ! Essaie encore un peu, bonyeu, tu vas finir par t'habituer ! C'est quand même pas toi qui vas montrer à la comtesse de Ségur comment écrire des livres, verrat !

Elle porta une main à sa bouche, l'autre à son cœur.

— Ça y est, y m'a faite sacrer un matin de Noël !

— Grand-moman Tremblay a dit, l'autre jour, que « verrat » c'tait pas un sacre ! Un verrat, c'est un cochon, pis un cochon, ça peut pas être un sacre !

Tremblay, Michel (1994) *Un ange cornu avec des ailes de tôle*. Montréal : Leméac, p. 40.

- 3 Les deux phrases ci-dessous ont-elles le même sens linguistique ? Ont-elles le même sens logique ? Quelle conclusion peut-on en tirer ?

(12) a. *Cet homme est vivant ou mort.*

b. *Soit il pleut soit il ne pleut pas.*

- 4 Est-ce que 'rusé' fait partie du sens de RENARD ? De sa connotation ?
- 5 Les lexies suivantes sont-elles des prédicats ? Si oui, combien d'arguments ont-elles ?

- DORMIR [*Jean dort depuis trois heures.*]
- PRÊTER [*Jean a prêté son livre.*]
- SOMMEIL [*Il dort d'un profond sommeil.*]
- DÉPART [*Le départ a lieu à trois heures.*]
- LUNE [*La lune est pleine ce soir.*]
- DIFFÉRENT [*Jules est très différent d'Émile.*]

- 6 Traduire en français la phrase anglaise ci-dessous. Constate-t-on quelque chose de spécial au niveau du fonctionnement des prédicats sémantiques anglais et français impliqués ici ?

(13) *I miss you.*

- 7 Représenter sous forme de réseau sémantique le sens de la phrase ci-dessous :

(14) *Rencontrer Lida a bouleversé la vie de Léo.*



Chapitre 7

Relations sémantiques lexicales

Relation sémantique lexicale, composante sémantique, identité / intersection / inclusion / disjonction de sens, sens plus simple (qu'un autre), hyperonymie vs hyponymie, sens plus riche (qu'un autre), hiérarchie sémantique des lexies, synonymie exacte vs approximative, paraphrase, antonymie, lexies contrastives, conversivité, homonymie / homographie / homophonie, polysémie, causativité, fonction lexicale paradigmaticque, (principe de) compositionnalité sémantique, locution (= expression idiomatique), collocation (= expression semi-idiomatique), base d'une collocation, collocatif, fonction lexicale syntagmatique.

*La guerre éclate
C'est tout ce qu'elle sait faire
Les bombes hachent
C'est tout ce qu'elles savent faire
Brigitte Fontaine, Il pleut*

Le chapitre précédent s'est attaché à décrire la nature du sens lexical. Nous allons maintenant examiner les différents types de liens sémantiques pouvant exister entre lexies : les **relations sémantiques lexicales**. Je procéderai en trois étapes :

- 1 modélisation des liens sémantiques par comparaison « ensembliste » des sens lexicaux (*grosso modo*, inclusion et intersection de sens) ;
- 2 présentation des relations sémantiques fondamentales, qui sont à la base de la structuration sémantique du lexique ;
- 3 modélisation des relations lexicales au moyen d'un outil descriptif formel appelé *fonction lexicale*.

Les sens lexicaux conçus comme des ensembles

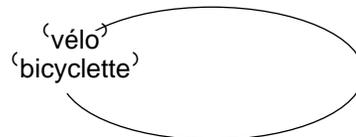
On peut se représenter un sens lexical comme étant un ensemble structuré d'autres sens lexicaux. Par exemple, le sens 'lit' contient les sens 'meuble' (un lit est un meuble), 's'allonger' (il est conçu pour qu'on s'y allonge), 'dormir ou se reposer' (on l'utilise avant tout pour dormir ou se reposer), etc.

Il est relativement aisé de démontrer l'inclusion de sens. Ainsi, 'meuble' est inclus dans 'lit' parce que l'expression *ce lit* dénote nécessairement un meuble, l'inverse n'étant pas vrai : l'expression *ce meuble* peut dénoter un fauteuil, une armoire, etc. De plus, 's'allonger' est aussi inclus dans 'lit' : *ce lit* dénote un meuble servant à s'allonger alors que *Je m'allonge* ne dénote pas une situation où il est nécessairement fait usage d'un lit.

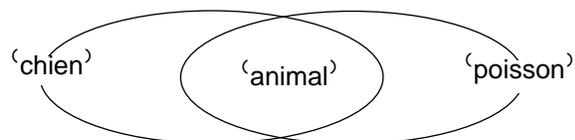
Lorsqu'un sens est inclus dans un autre, nous dirons qu'il en est une **composante**. Il est aussi de ce fait une composante de la définition de la lexie correspondante ; par exemple, 'meuble' est une composante de la définition de LIT.

Si l'on considère les sens lexicaux comme des ensembles, quatre types de relations sémantiques lexicales peuvent logiquement exister :

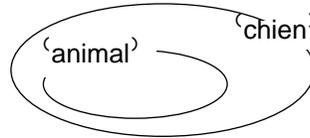
- 1 **identité de sens** — par ex., 'vélo' = 'bicyclette'



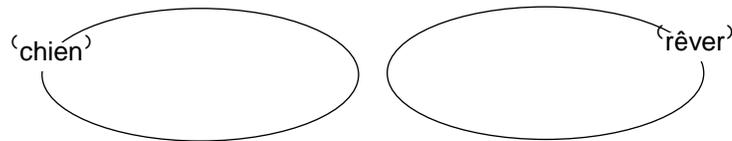
- 2 **intersection de sens** — par ex., 'chien' \cap 'poisson' = 'animal'



3 **inclusion de sens**¹ — par ex., ‘animal’ \subset ‘chien’



4 **disjonction de sens** — ‘chien’ \cap ‘rêver’ = \emptyset



Ces petits schémas² sont sans doute très jolis mais ils ne font que donner un caractère superficiellement « scientifique » à la description des liens sémantiques lexicaux. Ils ne disent presque rien sur les liens sémantiques en question. Ainsi, ‘chien’ contient non seulement ‘animal’ mais aussi ‘poil’ et ‘queue’ — un chien est un animal à poils possédant une queue. Or, le rapport qu’entretient ‘chien’ avec ‘animal’ est très différent de celui qu’il entretient avec ‘poil’ et ‘queue’.

La description des liens sémantiques uniquement en terme d’inclusion de sens ne permet pas de rendre compte de ces phénomènes. Nous reviendrons sur ce problème à la fin du chapitre. Le but de cette première section était avant tout de démontrer qu’il est pertinent de considérer qu’un sens « contient » d’autres sens. En fait, l’inclusion et l’identité de sens sont les relations sémantiques premières ; ce sont celles sur lesquelles se fondent, de façon directe ou indirecte, toutes les autres relations sémantiques fondamentales, comme nous le verrons dans la section suivante.



Lorsqu’un sens ‘s₁’ est inclus dans un sens ‘s₂’, on dira que ‘s₁’ est **plus simple** que ‘s₂’. L’emploi de ce terme est justifié par le fait que ‘s₁’ entre, en quelque sorte, dans la composition du sens ‘s₂’ ; il en est un des éléments constitutifs. Bien entendu, dans le cas où aucune relation d’inclusion ne peut être établie entre deux sens, il devient absurde de

1. Comme en logique ensembliste, une inclusion de sens est un cas particulier d’intersection de sens, où l’intersection correspond en fait à un des deux sens en cause.

2. Ils sont appelés *diagrammes de Venn* en mathématiques. On les désigne aussi parfois sous le nom plus familier de *patates*.

parler de sens plus simple. Ainsi, il ne serait pas pertinent de se demander si 'nager' est plus simple que 'fauteuil'.

Relations sémantiques fondamentales

Les relations sémantiques qui vont être présentées ici sont considérées comme étant les relations sémantiques « fondamentales » car elles forment la charpente de la structuration sémantique du lexique de toute langue. Chaque lexie se positionne dans le réseau lexical de la langue tout d'abord en fonction de ces relations.

Hyperonymie et hyponymie

Il s'agit ici de deux relations sémantiques converses entre lexies, correspondant à un cas particulier d'inclusion de sens :

La lexie L_{hyper} est un **hyperonyme** de la lexie L_{hypo} si

- le sens ' L_{hyper} ' est inclus dans le sens ' L_{hypo} '
- et si ' L_{hypo} ' peut être considéré comme un cas particulier de ' L_{hyper} '.

La lexie L_{hypo} , quant à elle, est appelée **hyponyme** de L_{hyper}

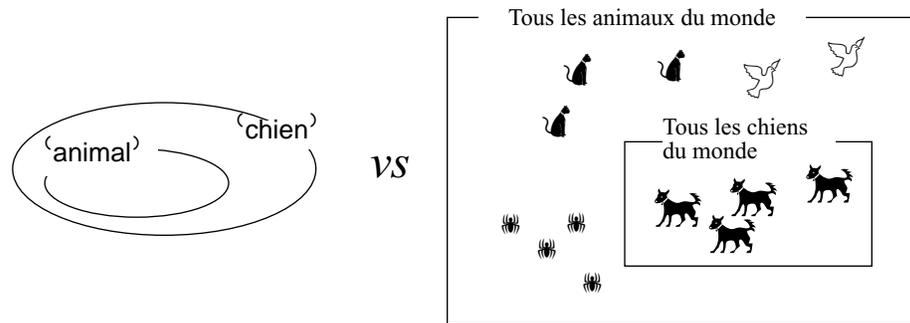
Pour reprendre un exemple déjà examiné plus haut, on dira que ANIMAL est un hyperonyme de CHIEN et que CHIEN est un de ses hyponymes, avec d'autres lexies telles que CHAT, CHEVAL, DROMADAIRE, POISSON, etc.

La définition ci-dessus implique qu'un hyperonyme est nécessairement un sens plus simple que son ou ses hyponymes. Cependant, comme l'hyperonymie et l'hyponymie renvoient à une situation beaucoup plus spécifique que la seule inclusion, on évitera désormais de parler dans un tel cas de sens plus (ou moins) simple. On dira que le sens d'un hyponyme est **plus riche** que celui de son hyperonyme et, *vice versa*, qu'un hyperonyme est moins riche que ses hyponymes.



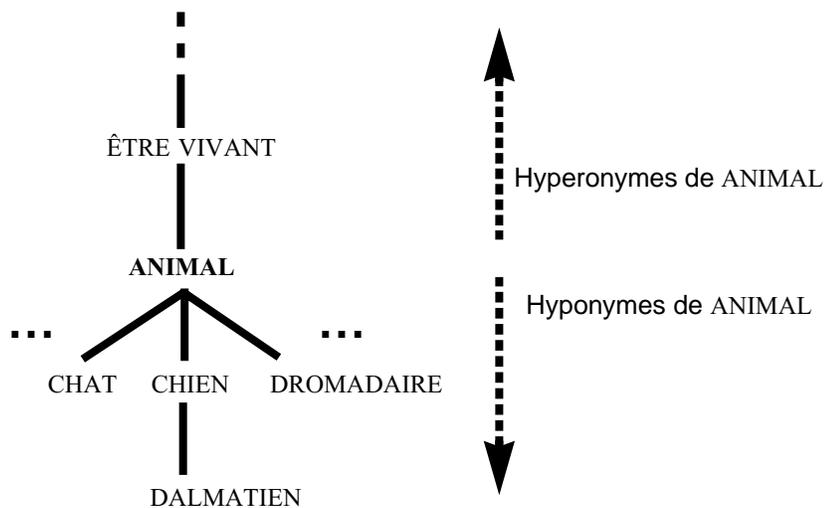
Il est intéressant de remarquer que si L_{hypo} est un hyponyme de L_{hyper} , l'ensemble des référents possibles de L_{hypo} est inclus dans celui des référents possibles de L_{hyper} ; en contrepartie, le sens de L_{hyper} est, lui,

inclus dans celui de L_{hypo} . Ainsi, CHIEN se définit par ‘**animal** domestique ...’ — donc, le sens de ANIMAL est inclus dans celui de CHIEN — mais l’ensemble de tous les chiens du monde est inclus dans l’ensemble des animaux :



Ce phénomène est une bonne illustration de la distinction existant entre la notion de sens et celle de référent (voir chapitre précédent, page 100 et suivantes).

La relation d’hyponymie-hyponymie est transitive³ et permet donc de construire une *hiérarchie sémantique des lexies*, hiérarchie qui peut se représenter sous la forme d’un « arbre » :



Extrait de la hiérarchie sémantique des lexies françaises (centré autour de ANIMAL)

En général, on ne lie par les relations d'hyperonymie et d'hyponymie que des lexies appartenant à la même partie du discours. On pourrait cependant admettre dans une certaine mesure que, par exemple, le nom SENTIMENT est un hyperonyme non seulement du nom AMOUR mais aussi du verbe AIMER. Il existe donc une certaine flexibilité dans l'utilisation des relations d'hyperonymie et d'hyponymie. Il en va tout autrement des autres relations qu'il nous reste à examiner : elles ne valent que pour des lexies appartenant à la même partie du discours.

Synonymie

La synonymie, c'est-à-dire l'identité de sens, est la relation lexicale sémantique par excellence :

Deux lexies L_1 et L_2 appartenant à la même partie du discours sont des ***synonymes exacts*** (ou synonymes absolus) si ' L_1 ' \equiv ' L_2 '.

Ce sont des ***synonymes approximatifs*** si ' L_1 ' \equiv ' L_2 '. Dans ce dernier cas, il y a soit intersection soit inclusion de sens telle que L_1 et L_2 peuvent être considérées comme ayant une valeur sémantique suffisamment proche pour que l'une puisse être utilisée à la place de l'autre pour exprimer sensiblement la même chose.

Il est essentiel de noter que la synonymie lexicale exacte est rarissime ; on cite toujours les mêmes exemples pour le français : VÉLO et BICYCLETTE, AUTOMOBILE et VOITURE, etc. La synonymie lexicale est avant tout une synonymie approximative. Elle peut être testée en effectuant des substitutions en contexte :

L_1 et L_2 peuvent être considérées comme étant synonymes si, en remplaçant L_1 par L_2 dans une phrase, on obtient une nouvelle phrase à peu près équivalente sémantiquement — c'est-à-dire une ***paraphrase*** approximative⁴.

3. Si ANIMAL est un hyperonyme de CHIEN et CHIEN un hyperonyme de DALMATIEN, alors ANIMAL est aussi un hyperonyme de DALMATIEN.

4. Nous reviendrons sur le test de substitution en contexte dans le Chapitre 8, qui traite de l'analyse sémantique lexicale. Je rappelle que la notion de paraphrase, quant à elle, a été examinée au chapitre précédent.

Par exemple :

(1) *Nestor éprouve de la **haine** pour Bianca.*

≡

*Nestor éprouve de l'**aversion** pour Bianca.*

Les deux lexies HAINE et AVERSION, en plus d'être des synonymes approximatifs, sont en même temps des co-hyponymes de SENTIMENT. Leur intersection de sens est *grosso modo* : 'sentiment négatif éprouvé envers quelqu'un'.

Il peut aussi tout à fait arriver, comme l'implique la définition ci-dessus, qu'un hyperonyme d'une lexie soit en même temps un de ses synonymes approximatifs (cas d'inclusion de sens entre L₁ et L₂) :

(2) *Le **déluge** d'hier soir a endommagé les récoltes.*

≡

*La **pluie** d'hier soir a endommagé les récoltes.*

Mais cela n'est pas toujours le cas, surtout si l'on considère un hyperonyme qui ne se trouve pas immédiatement au-dessus de la lexie en question dans la hiérarchie sémantique des lexies. Ainsi, il serait très bizarre d'affirmer que (3b) ci-dessous est une paraphrase de (3a) :

(3) a. *Regarde ce **dalmatien** !*

b. *Regarde cet **être vivant** !*

L'écart de sens entre DALMATIEN et ÊTRE VIVANT est bien trop grand pour que l'on puisse considérer ces deux lexies comme étant des synonymes approximatifs.

Les synonymes approximatifs se distinguent en général non seulement par leur sens — qui n'est pas exactement identique — mais aussi par leur combinatoire. Il ne faut donc pas s'attendre à ce que le test consistant à substituer une lexie à son synonyme dans une phrase pour voir si l'on obtient des paraphrases s'applique dans n'importe quel contexte. Par exemple, TRAVAIL [*Il a un travail intéressant.*] et EMPLOI [*Il a un emploi intéressant.*] sont clairement des synonymes approximatifs. Cependant, leur combinatoire n'est pas totalement identique, comme le montrent les exemples suivants :

- (4) a. *une offre d'emploi ~ *une offre de travail*
 b. *perdre son emploi ~ ?perdre son travail*
 c. *créer un emploi ~ *créer un travail*
 d. *un emploi/travail de comptable*
 e. *les emplois de comptable ~ *les travaux de comptable*

Les synonymes ne sont donc pas nécessairement mutuellement substituables dans tous les contextes. Il suffit cependant que l'on puisse facilement trouver des contextes où la substitution paraphrastique est possible pour que le lien de synonymie soit établi.

Antonymie

Bien que l'antonymie s'oppose logiquement à la synonymie, ces deux relations sont somme toute très proches puisqu'elles lient des lexies présentant une forte parenté sémantique :

Deux lexies L_1 et L_2 appartenant à la même partie du discours sont des **antonymes** si ' L_1 ' et ' L_2 ' se distinguent par la négation ou, plus généralement, la mise en opposition d'une composante de leur sens.

Comme dans le cas des synonymes, on peut distinguer les antonymes exacts — cf. (5a) — et les antonymes approximatifs — cf. (5b) :

- (5) a. *Cet arbre est **près/loin** de moi.*
 b. *Il **aime/déteste** le fromage.*

Je propose de travailler sur l'analyse de cet exemple dans l'Exercice 4, en fin de chapitre.

Il est possible d'identifier des types particuliers d'antonymes, selon la nature de l'opposition sémantique mise en jeu. Ainsi, on trouve fréquemment citées dans la littérature, comme exemples d'antonymes, des paires de lexies dites « réversives » :

- BOUTONNER vs DÉBOUTONNER,
- COLLER vs DÉCOLLER,

- CONSTRUIRE vs DÉTRUIRE,
- etc.

On peut aussi identifier une antonymie dite « scalaire », qui lie des lexies dénotant des valeurs situées aux deux extrêmes d'une échelle de valeurs possibles :

- CHAUD vs FROID,
- GRAND vs PETIT,
- etc.

L'antonymie recouvre donc des connexions lexicales très variées et je ne fais que donner ici une caractérisation minimale de cette notion.



Attention cependant ! Il ne faut pas confondre les antonymes véritables et les lexies dites **contrastives**. Ainsi BLANC et NOIR sont en opposition contrastive dans la langue, comme en témoignent les expressions *écrire noir sur blanc*, *en noir et blanc*, *dire « blanc » quand quelqu'un d'autre dit « noir »*, etc. Pourtant, ces lexies ne sont pas des antonymes : dire *cette chose est noire* ne revient pas à **exprimer** l'opposé de *cette chose est blanche*. *Sa chemise est noire* signifie que la chemise en question est d'une couleur particulière, associée à la nuit (à l'absence de lumière), etc. L'opposition contrastive est bien une relation sémantique lexicale. Cependant, contrairement à l'antonymie, elle ne repose pas entièrement sur une opposition se situant au niveau du sens de chacune des lexies en cause. En décomposant le sens de deux antonymes comme CHAUD et FROID, on va pouvoir démontrer la relation d'antonymie ; il suffit pour cela que la seule différence existant entre les deux décompositions soit la négation ou la mise en opposition d'une composante de sens :

CHAUD \cong 'dont la température est **plus** élevée que la normale'
 FROID \cong 'dont la température est **moins** élevée que la normale'

L'opposition contrastive, en revanche, ne peut pas être démontrée par l'analyse des sens lexicaux car elle est en grande partie conventionnelle. Elle trouve sa justification dans la présence d'expressions toutes faites, de proverbes, etc. qui sont des évidences linguistiques de la présence de cette opposition entre deux lexies données de la langue — comme dans le cas de BLANC et NOIR examiné ci-dessus. Je rappelle que la notion

d'évidence linguistique a été introduite dans le chapitre précédent (page 106), lorsqu'il s'agissait de trouver comment démontrer la présence de connotations associées aux lexies de la langue.

Conversivité

Pour bien comprendre la notion de conversivité, il faut utiliser la modélisation des sens lexicaux en tant que prédicats sémantiques, qui a été présentée au chapitre précédent (page 107) :

Deux lexies L_1 et L_2 appartenant à la même partie du discours sont **conversives** si

- ce sont des prédicats sémantiques dénotant une même situation
- et qui intervertissent l'ordre d'au moins deux de leurs arguments.

Les paires de conversifs peuvent appartenir à n'importe quelle partie du discours, pourvu que ces conversifs soient des prédicats sémantiques possédant au moins deux arguments. Par exemple, il peut s'agir de verbes

(6) X **emploie** Y $\equiv Y$ **travaille** pour X .

aussi bien que de noms

(7) X est l'**employeur** de Y $\equiv Y$ est l'**employé** de X .

Maintenant que la relation de conversivité a été introduite, nous pouvons revenir sur l'Exercice 6 du chapitre précédent (page 115). Si l'on accepte d'élargir la conversivité à la mise en relation de lexies appartenant à des langues différentes, on peut dire que la lexie anglaise MISS_V [*I miss you.*] fonctionne comme un conversif de MANQUER [*Tu me manques.*] (et *vice versa*).

Il est important de noter que de nombreuses langues, dont le français, offrent un moyen grammatical pour effectuer une paraphrase basée sur l'inversion de l'ordre des arguments dans la phrase : la voix passive. C'est ce qu'illustrent les phrases suivantes, où la conversion est effectuée tout d'abord lexicalement, (8b), puis grammaticalement, (8c) :

Notons que l'on peut en fait distinguer deux cas d'homonymie, selon le type de signifiant que l'on prend en compte :

- 1 Il y a **homographie** lorsque les deux lexies sont associées aux mêmes signifiants écrits — par exemple, *Il a acheté trois livres* vs *Ça pèse trois livres*. La plupart des dictionnaires courants distinguent les noms de lexies homographiques par des numéros en exposant : LIVRE¹ vs LIVRE². Tel qu'indiqué dans le Chapitre 3 (page 43), j'utilise la même convention d'écriture dans cet ouvrage.
- 2 Il y a **homophonie** lorsque les deux lexies sont associées aux mêmes signifiants sonores — par exemple, *Il est sot* vs *Va remplir ton seau*.

Une homographie peut coïncider avec une homophonie ; c'est le cas de LIVRE¹ vs LIVRE² ci-dessus. Mais il n'en va pas toujours ainsi : *du pain bis (/bi/)* vs *Le public a réclamé un bis (/bis/)*.

On voit donc que l'homonymie n'est pas une relation véritablement **sémantique** entre lexies. C'est bien plutôt une relation de forme très forte — une identité de signifiants —, qui est particulière puisqu'elle s'accompagne justement d'une absence de lien sémantique !

Polysémie

Le terme de *polysémie* ne désigne pas au sens strict une relation de sens entre lexies mais une caractéristique d'un vocable :

Un vocable est **polysémique** s'il contient plus d'une lexie.

Bien entendu, on peut dériver de la polysémie une notion importante qui est, elle, une relation sémantique véritable. On peut ainsi dire que la lexie VERRE² [*Il boit dans un verre.*] est un « partenaire » de la lexie VERRE¹ [*Le verre est un matériau transparent.*] au sein du vocable polysémique VERRE. Mais il n'existe pas à proprement parler de terme courant pour désigner cette relation. Par contre, comme nous allons le voir immédiatement ci-dessous, il existe des termes pour désigner des rapports sémantiques spécifiques entre lexies, rapports qu'entretiennent fréquemment les lexies d'un même vocable.

La plupart des vocables courants de la langue sont polysémiques. Dans le cadre de l'étude sémantique, on se doit d'examiner les différents schémas de polysémie, qui sont souvent fondés sur des types de

relations sémantiques différents de ceux qui ont été examinés jusqu'à présent. Pour ne pas surcharger trop ce chapitre (déjà assez consistant), je ne citerai ici que la relation sémantique de **causativité** :

La lexie L_1 est un causatif de la lexie L_2 si ' L_1 ' \cong 'causer L_2 '.

La causativité est souvent impliquée dans les rapports sémantiques entre lexies d'un même vocable :

- (9) a. *L'eau bout.*
b. *Je bous de l'eau.*

Même si certains diront peut-être plus fréquemment *faire bouillir de l'eau* que *bouillir de l'eau*, la phrase (9b) reste tout à fait normale. On peut la comparer avec (10b) ci-dessous, qui est, elle, incorrecte :

- (10) a. *L'eau frémit dans la bouilloire.*
b. **Je frémis de l'eau dans la bouilloire.*

La causativité peut bien évidemment aussi lier des lexies n'appartenant pas au même vocable, avec présence ou non d'un lien morphologique entre les deux lexies en cause :

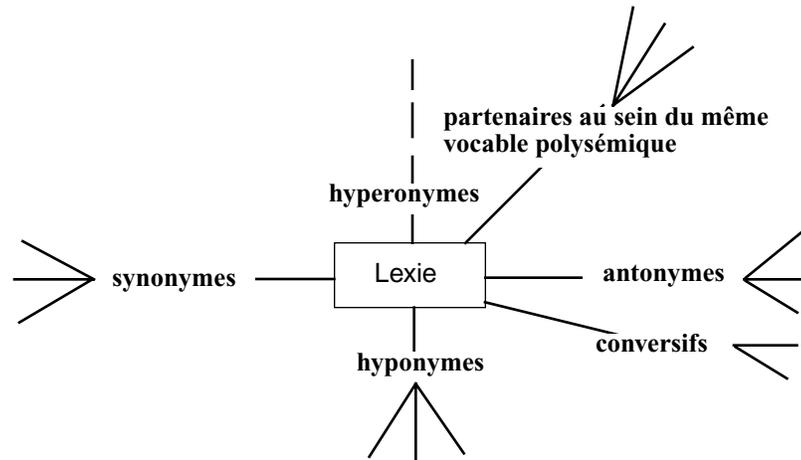
- (11) a. *Le bébé dort.*
b. *J'endors le bébé.*

- (12) a. *César meurt.*
b. *Brutus tue César.*

Je reviendrai beaucoup plus en détail sur la causativité et sur d'autres types courants de liens sémantiques entre lexies d'un même vocable polysémique dans le Chapitre 8 (Section *Structure sémantique des vocables*, page 165).

Nous en avons terminé avec la présentation des relations sémantiques fondamentales. Je rappelle que toute lexie se positionne avant toute

chose dans le réseau lexical de la langue en fonction de ces relations, comme l'illustre la figure ci-dessous :



Il existe un très grand nombre d'autres relations sémantiques lexicales récurrentes en langue, et l'on pourrait consacrer tout un ouvrage à leur étude. L'examen systématique des relations lexicales déborde donc largement le cadre de ce cours d'introduction. Il est cependant important, arrivé à ce stade de l'exposé, de montrer que la structure sémantique du lexique, bien que complexe et irrégulière, peut être étudiée et modélisée de façon relativement systématique. C'est ce que je vais faire dans la section qui suit.

La modélisation formelle des relations lexicales

Nous allons étudier les fonctions lexicales, un outil proposé dans le cadre de la théorie linguistique Sens-Texte pour modéliser les relations entre lexies. On peut distinguer deux aspects de l'étude des fonctions lexicales : la notion elle-même vs sa formalisation. Même si des éléments de formalisation seront introduits ici, je vais surtout chercher à faire comprendre la notion de fonction lexicale en tant que telle, à montrer comment elle rend compte des différents types de liens qui peuvent unir les éléments du réseau lexical de la langue.

La notion de fonction lexicale

Comme nous l'avons vu dans la section précédente, les relations lexicales telles que la synonymie et l'antonymie lient entre elles des lexies qui sont substituables l'une à l'autre dans une phrase puisqu'elles appartiennent à la même partie du discours :

- synonymie → *Jean a une* $\begin{bmatrix} \text{voiture} \\ \text{automobile} \end{bmatrix}$.
- antonymie → *Jean est* $\begin{bmatrix} \text{petit} \\ \text{grand} \end{bmatrix}$.

On peut cependant concevoir d'autres types de relations lexicales, liant des lexies appartenant à des parties du discours différentes. En voici deux exemples.

- 1 Deux lexies peuvent appartenir à des parties du discours différentes tout en ayant un sens identique, comme le démontrent les deux paires verbe-nom ci-dessous :
 - COURIR [*Jean court vite.*] → COURSE [*La course de Jean est rapide.*] ;
 - DORMIR [*Jean dort profondément.*] → SOMMEIL [*Le sommeil de Jean est profond.*].
- 2 Certaines lexies prédicatives (certains verbes, notamment) sont liées de façon privilégiée à des lexies nominales qui correspondent au nom standard d'un de leurs arguments :
 - COURIR ['X court'] → COUREUR [= 'X qui court'] ;
 - DORMIR ['X dort'] → DORMEUR [= 'X qui dort'] ;
 - VOLER ['X vole Y à Z'] → VOLEUR/COUPABLE [= 'X qui vole'], BUTIN [= 'Y qui a été volé'] et VICTIME [= 'Z à qui Y a été volé'].

Bien entendu, un coupable n'est pas nécessairement le coupable d'un vol, remarque qui vaut aussi pour une victime. De plus, la lexie BUTIN veut dire plus que simplement 'quelque chose qui a été volé'. Il n'en demeure pas moins que la façon standard de référer aux arguments de VOLER est d'utiliser les lexies mentionnées ci-dessus, comme le montre le texte suivant :

(13) *Un collier d'une valeur inestimable a été volé hier. Le coupable/voleur s'est enfui avec son butin sans que la victime, Mme Bianca C., puisse intervenir.*

La famille de relations examinée ici peut aussi unir deux lexies appartenant à la même partie du discours : la relation qui existe entre VOLER et VOLEUR est la même que celle existant entre VOL et VOLEUR ; celle existant entre DORMIR et DORMEUR est identique à celle existant entre SOMMEIL et DORMEUR ; etc.

Il a été proposé, dans le cadre de la théorie linguistique Sens-Texte (voir lectures complémentaires pour ce chapitre), de décrire toutes les relations sémantiques lexicales au moyen d'un outil formel conçu sur le modèle des fonctions mathématiques : les fonctions lexicales.

Une **fonction lexicale** f décrit une relation existant entre une lexie L — l'argument de f — et un **ensemble** de lexies ou d'expressions figées appelé la *valeur* de l'application de f à la lexie L .

- L'expression $f(L)$ représente l'application de la fonction f à la lexie L .
- Chaque élément de la valeur de $f(L)$ est lié à L de la même façon.
- Il existe autant de fonctions lexicales qu'il existe de types de liens lexicaux et chaque fonction lexicale est identifiée par un nom particulier — **Syn**, **Anti**, etc.

Cette définition est sans doute assez « abstraite » et doit être illustrée par des cas précis de modélisation de liens lexicaux. C'est ce que nous allons faire maintenant. Lors de l'étude d'exemples concrets de fonctions lexicales, il sera important de se reporter régulièrement à la définition ci-dessus pour bien comprendre la nature théorique de cette notion.

Fonctions lexicales paradigmatiques

Je vais tout d'abord démontrer que chacune des relations lexicales qui ont été examinées jusqu'à présent dans ce chapitre peut être modélisée au moyen d'une fonction lexicale particulière. Remarquons qu'il s'agissait dans tous les cas de liens lexicaux paradigmatiques (cf.

3 S_0 est la fonction lexicale qui associe à une lexie verbale, adjectivale ou adverbiale sa contrepartie nominale :

- $S_0(\textit{courir}) = \textit{course}$
- $S_0(\textit{dormir}) = \textit{sommeil}$

On voit qu'il s'agit ici de la modélisation au moyen d'une fonction lexicale du cas de figure n° 1 examiné ci-dessus, à la page 131.

Le symbole « S » renvoie au terme *substantif*. Il forme la base du nom de plusieurs fonctions lexicales dont l'application retourne des éléments nominaux (voir notamment S_i ci-dessous).

4 V_0 est le pendant verbal de S_0 , qui associe donc à une lexie nominale, adjectivale ou adverbiale sa contrepartie verbale :

- $V_0(\textit{sommeil}) = \textit{dormir}$

5 S_i lie une lexie prédicative au nom standard de son i^{e} argument (premier, deuxième, ... argument) :

- $S_1(\textit{courir}) = \textit{coureur}$
- $S_1(\textit{dormir}) = \textit{dormeur}$
- $S_1(\textit{voler [qqch.]}) = \textit{voleur, coupable}$
- $S_2(\textit{voler [qqch.]}) = \textit{butin}$
- $S_3(\textit{voler [qqch.]}) = \textit{victime}$

Il s'agit ici de la modélisation au moyen de fonctions lexicales du cas de figure n° 2 de la page 131.

Les différents exemples qui viennent d'être présentés montrent bien qu'une application de fonction lexicale — $f(L)$ — retourne une valeur qui est un ensemble. Celui-ci peut contenir un seul élément, comme $S_2(\textit{voler [qqch.]})$, deux éléments, comme $S_1(\textit{voler [qqch.]})$, ou plus. Mais la valeur peut aussi être l'ensemble vide ; comparons, par exemple, les deux applications suivantes de S_0 :

$$S_0(\textit{marcher}) = \textit{marche} \sim S_0(\textit{ramper}) = \emptyset.$$

Je n'irai pas plus loin dans l'illustration de la modélisation de liens paradigmatiques au moyen de fonctions lexicales. Cependant, les fonctions lexicales permettent aussi de rendre compte de la combinatoire des lexies — de liens syntagmatiques —, en encodant les affinités particulières que peut avoir une lexie de la langue avec d'autres lexies

au sein de ce que l'on appelle des collocations. Je vais tout d'abord introduire la notion fort importante de collocation et montrer ensuite comment elle se modélise au moyen de fonctions lexicales dites *syntagmatiques*.

Les collocations

En théorie, les énoncés obéissent au principe suivant, déjà mentionné dans le Chapitre 3 (page 40) :

Selon le principe de **compositionalité sémantique**, le sens d'un énoncé est la résultante de la composition du sens des éléments qui le constituent.

Ce principe rend compte du fait que le sens d'une phrase comme

(14) *Les trois escargots se ruèrent sur Gustave.*

est directement calculable à partir du sens de chacune des lexies qui le composent. Pour pouvoir calculer le sens de la phrase (14), il « suffit » de connaître :

- 1 le sens des lexies LE, TROIS, ESCARGOT, SE RUER, GUSTAVE,
- 2 les sens grammaticaux de pluriel et de passé,
- 3 les règles syntaxiques et morphologiques du français.

Les énoncés sont cependant truffés de cas qui contredisent ce principe général. Ainsi, nous avons déjà vu dans le Chapitre 3 (page 38) le cas des **locutions**, qui sont constituées formellement d'éléments dont le sens ne se retrouve pas dans le sens de la locution : PASSER À TABAC, CASSER LES PIEDS, TIRER LES VERS DU NEZ, etc. Les locutions sont ce qu'on appelle des **expressions idiomatiques**.

Il existe un autre cas de transgression du principe de compositionalité sémantique ; il s'agit des **collocations**, qui sont des **expressions semi-idiomatiques** :

L'expression *AB* (ou *BA*), formée des lexies A et B, est une collocation si la lexie A est sélectionnée librement — strictement d'après son sens ^(A) —, alors que la lexie B est sélectionnée pour exprimer un sens ^(C) en fonction de A.

Voici quelques exemples de collocations en français :

- (15) a. *grosse*_[= B] *tempête*_[= A]
b. *dormir*_[= A] *profondément*_[= B]
c. *pleuvoir*_[= A] *des cordes*_[= B]

Une collocation est contrôlée par l'élément qui retient son sens dans la collocation et qui est sélectionné librement par le locuteur : il s'agit de la **base de la collocation**. Dans les exemples ci-dessus, les bases des collocations sont les éléments étiquetés *A*. Les éléments étiquetés *B* sont appelés **collocatifs**.

On dit que la base contrôle la collocation car, du point de vue du locuteur, c'est le collocatif qui est choisi en fonction de la base, et non l'inverse. Si j'entre dans une pièce mal éclairée et que je veux signaler ce fait, je peux dire, par exemple, (16a) ou (16b) :

- (16) a. *Il fait noir ici.*
b. *Il fait sombre ici.*

Cependant, si je veux revenir sur ce que je viens de dire, pour le modifier et insister sur le fait qu'on n'y voit vraiment rien, mon choix initial va contraindre la façon dont je peux maintenant m'exprimer :

- (17) a. *Il fait même noir **comme dans un four**.*
b. *Il fait même **vraiment très** sombre.*
c. **Il fait même sombre **comme dans un four**.*

Il n'y a aucune raison **logique** pour que (17a) soit correcte et non (17c), si ce n'est le fait que la locution FAIRE NOIR admet *comme dans un four* comme collocatif, ce que n'admet pas la locution FAIRE SOMBRE. Il est certain qu'en disant cela, je n'explique rien. Je ne fais que constater les faits. Nous allons voir que s'il n'est pas possible de prédire que *comme dans un four* se combine avec FAIRE NOIR et non avec FAIRE SOMBRE, on peut cependant prédire qu'un problème de combinatoire peut être anticipé ici.

Les collocations, en tant que phénomène linguistique, possèdent les trois caractéristiques suivantes :

- 1 Elles sont universellement présentes dans toutes les langues.
- 2 Elles sont omniprésentes dans les textes, qu'ils soient oraux ou écrits.
- 3 Elles semblent plus ou moins arbitraires, ne peuvent pas se traduire mot à mot d'une langue à l'autre et sont donc très difficiles à acquérir.



Toutes ces caractéristiques nous indiquent qu'il serait très utile de disposer d'un mécanisme qui nous permette non seulement de décrire les collocations de façon rigoureuse, mais aussi de faire des prédictions à leur propos : *Cette expression doit être une collocation et je dois faire attention en la traduisant / en l'enseignant, Ce sens doit sûrement s'exprimer par une collocation*, etc. C'est ici que les fonctions lexicales vont intervenir.

Fonctions lexicales syntagmatiques

Commençons par une remarque très importante concernant les collocations, vues selon une perspective sémantique :

Certains sens très généraux, universellement exprimés dans toutes les langues, tendent à s'exprimer de façon collocationnelle.

Les collocations que l'expression de ces sens engendre seront décrites au moyen de **fonctions lexicales syntagmatiques**. Pour ne pas saturer ce chapitre avec trop de nouvelles notions, je ne ferai que présenter quelques exemples de fonctions lexicales de ce type.

Une collocation se caractérise selon deux axes : le sens exprimé par le collocatif, bien entendu, mais aussi le rôle syntaxique que joue le collocatif auprès de la base. Pour bien illustrer la richesse autant sémantique que syntaxique des phénomènes collocationnels, nous examinerons des fonctions lexicales syntagmatiques rendant compte des deux types de collocations suivants :

- 1 collocations où le collocatif est un modificateur de la base ;
- 2 collocations où le collocatif est un verbe support, gouvernant syntaxiquement la base.

Collocatif qui est un modificateur de la base J'entends ici par *modificateur* un élément de la phrase qui fonctionne auprès de la base comme un adjectif épithète, si la base est nominale, ou comme un adverbe, si la base est verbale ou adjectivale⁷. Il s'agit donc de structures syntaxiques du type suivant :

- (18) a. *course* —modif→ *effrénée*
 b. *courir* —modif→ *à fond de train*
 c. *gros* ←modif— *bobo*
 d. *sale* ←modif— *regard*
 e. *regarder* —modif→ *de travers*
 f. *méchant* —modif→ *comme une teigne*

Nous allons voir deux fonctions lexicales décrivant des collocations de ce type, chacune de ces fonctions lexicales étant associée à une signification bien particulière.

- 1 La fonction lexicale **Magn** (nom tiré du latin *magnus* 'grand') associe à une lexie l'ensemble des lexies ou expressions linguistiques qui expriment auprès d'elle (en tant que modificateurs) l'intensification, c'est-à-dire le sens général 'intense', 'très', 'beaucoup', etc. L'intensification est un sens très vague, qui peut se combiner avec un peu n'importe quel autre sens et qui tend universellement à s'exprimer de façon collocationnelle :

- **Magn**(*chagrin*) = *grand, gros* < *énorme, immense*
- **Magn**(*amour*) = *grand* < *immense*
- **Magn**(*courir*) = *vite* < *à fond de train, à perdre haleine*

Dans les exemples ci-dessus, le symbole « inférieur à » indique une gradation entre les différents éléments de la valeur retournée. Ainsi, un énorme chagrin est plus « intense » qu'un gros chagrin.

7. Sur le fonctionnement syntaxique de l'adjectif et de l'adverbe, voir Chapitre 5, page 75.

2 La fonction lexicale **Bon** (nom tiré du latin *bonus* 'bon') associe à une lexie l'ensemble des lexies ou expressions linguistiques qui expriment auprès d'elle le sens général 'bon', 'bien', etc. — c'est-à-dire, l'évaluation positive ou l'approbation du locuteur :

- **Bon**(*colère*) = *saine, sainte*
- **Bon**(*compliment*) = *bien tourné*

Toutes les fonctions lexicales qui ont été présentées jusqu'à présent sont dites « simples », car elles ne sont pas analysables en terme d'autres fonctions lexicales. Les fonctions simples peuvent être combinées pour former des fonctions lexicales dites « complexes », comme **AntiMagn** (combinaison de **Anti** et **Magn**, c'est-à-dire le contraire d'un intensificateur) ou **AntiBon** :

- **AntiMagn**(*chagrin*) = *petit*
- **AntiBon**(*compliment*) = *fade, gauche, maladroit, mal tourné*

Afin de mieux mettre en évidence la richesse des données que ces fonctions lexicales syntagmatiques permettent de modéliser, je vous donne dans le tableau ci-dessous d'autres exemples de valeurs retournées en français pour **Magn**, **Bon**, **AntiMagn** et **AntiBon** :

Lexie (numérotation du <i>Nouveau Petit Robert</i>)	Fonc. lexic.	Valeur
ABOIEMENT 1 [Entends-tu ces aboiements ?]	Magn	<i>furieux, féroces</i>
	AntiMagn	<i>faibles</i>
	AntiBon	<i>intempestifs</i>
DÉFAITE 1 [Il a reconnu sa défaite.]	Magn	<i>écrasante, grave, sévère, sérieuse, terrible < complète, totale; cuisante; sanglante</i>
	AntiBon	<i>honteuse, humiliante</i>
FRISSON [Un frisson lui parcourut le dos.]	Magn	<i>grand</i>
	AntiMagn	<i>petit, léger</i>
	Bon	<i>agréable, délicieux</i>
FUNÉRAILLES [J'ai assisté à ses funérailles.]	Magn	<i>imposantes < grandioses</i>
	AntiMagn	<i>discrètes, intimes</i>
GALOP 1 [J'écoutais le galop des chevaux.]	Magn	<i>grand, double, triple, rapide</i>
	AntiMagn	<i>petit</i>
GOÛT 1.2 [La réunion d'hier avait un goût de fête.]	Magn	<i>clair, fort, marqué</i>
	AntiMagn	<i>petit, vague</i>
	Bon	<i>bon</i>
	AntiBon	<i>sale, mauvais</i>
REGARDER 1.1 [Il nous regardait sans bouger.]	Magn	<i>intensément; fixement; droit dans les yeux</i>
	AntiMagn	<i>discrètement, du coin de l'œil</i>
	AntiBon	<i>sournoisement</i>

Quelques illustrations de valeurs de **Magn**, **Bon**, **AntiMagn** et **AntiBon**

Bien entendu, toutes les options données pour une même valeur de fonction lexicale ne sont pas strictement équivalentes. *Regarder intensément* ne veut pas dire exactement la même chose que *regarder fixement* ou *droit dans les yeux*. C'est pourquoi j'utilise parfois des points-virgules pour indiquer des écarts sémantiques non négligeables. Ce qui importe ici c'est que toutes les options présentées restent interprétables comme des cas particuliers d'expression de la fonction lexicale correspondante.



L'étiquetage d'un lien base-collocatif au moyen d'un nom de fonction lexicale ne donne pas une description fine, parfaite, absolue (que de **Magn**(*description*)!) du lien en question. Il offre cependant une généralisation très intéressante, qui permet notamment de classer les collocations, de les anticiper lors de la description d'une lexie donnée (quels sont ses **Magn**, **AntiMagn**, ... ?), d'anticiper des problèmes d'apprentissage chez les personnes étudiant la langue, etc.

Collocatif qui est un verbe support Un verbe support est un collocatif verbal sémantiquement vide dont la fonction linguistique est de « verbaliser » une base nominale, c'est-à-dire de la faire fonctionner dans la phrase comme si elle était elle-même un verbe :

- (19) a. *éprouver* —compl. d'objet→ *du regret* ≡ *regretter*
b. *donner* —compl. d'objet→ *un coup* ≡ *frapper*

On considère que le verbe support est sémantiquement vide car il n'est pas utilisé par le locuteur pour exprimer un sens donné, qui ne serait pas déjà contenu dans la base de la collocation. Ainsi, tous les sens **lexicaux** contenus dans la phrase (20a) ci-dessous sont aussi contenus dans l'expression (20b) (où n'apparaît pas le verbe ÉPROUVER) :

- (20) a. *Jean éprouve du regret.*
b. *le regret de Jean*

La seule différence sémantique perceptible entre (20a) et (20b) se réduit à la structure communicative (voir Chapitre 6) et au temps grammatical, qui n'est pas exprimé dans le second cas. Les mêmes observations pourraient être faites à partir de la collocation *donner un coup*.

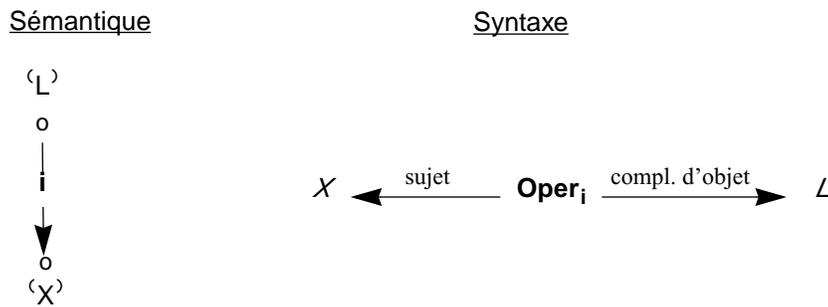
On mentionne le plus souvent les constructions à verbe support où la base est un complément du verbe collocatif (comme dans les exemples ci-dessus). Un verbe support peut cependant servir à faire fonctionner la base de la collocation comme sujet, par exemple :

- (21) a. *Un danger* ←sujet— *menace* [Jean].
b. *Les applaudissements* ←sujet— *se font entendre*.

Cette opposition — base qui est complément vs sujet du verbe support — me permet d'introduire ici deux nouvelles fonctions lexicales syntagmatiques : **Oper_i** et **Func_i**.

- 1 La fonction lexicale **Oper_i** (du latin *operari* 'faire') associée à une lexie prédicative nominale L l'ensemble des verbes supports qui prennent l'expression du i^e (premier, deuxième, etc.) argument de L comme sujet et prennent L comme complément d'objet direct ou indirect.

Cette caractérisation de **Oper_i** montre que, pour bien comprendre le fonctionnement de cette fonction lexicale, il est nécessaire d'avoir une vision claire des structures sémantiques et syntaxiques qu'elle conditionne. Le formalisme graphique des réseaux sémantiques, introduit au chapitre précédent, va nous permettre de visualiser cette correspondance entre configurations sémantique et syntaxique :



*Correspondance sémantique-syntaxe définissant la fonction lexicale **Oper_i***

Examinons immédiatement un exemple. Soit la phrase suivante où une collocation apparaît en gras :

(22) *Nicolas **donne un coup** à Alceste.*

Si nous menons une brève analyse sémantique des éléments de cette collocation, nous pouvons faire les constatations suivantes :

- La lexie COUP est, sémantiquement, un prédicat à deux arguments : 'coup de X sur Y'.
- En (22), *Nicolas* est l'expression du premier argument de COUP (et *Alceste* l'expression de son second argument).

- Dans cette phrase, *donne* est utilisé comme verbe support (sémantiquement vide) ; il prend *Nicolas* (premier argument de *coup*) comme sujet et *coup* comme complément d'objet direct.

Tout cela correspond exactement au rapport de configurations sémantique-syntaxe définissant la fonction lexicale **Oper_i** et représenté graphiquement dans la figure ci-dessus. Le sujet du verbe support étant le premier argument de la base de la collocation, nous sommes plus exactement en présence d'un **Oper₁**:

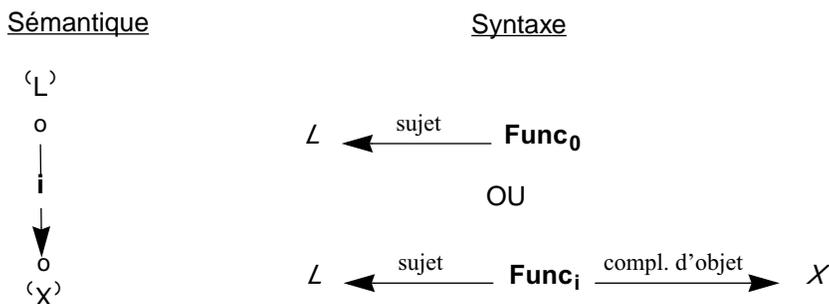
Oper₁(coup) = *administrer, asséner, donner, fam flanquer, porter*

La notion de **Oper_i**, telle qu'elle a été définie, nous permet de faire l'hypothèse que la lexie COUP pourrait aussi posséder des **Oper₂**, c'est-à-dire être la base de collocations du même type mais où le sujet du verbe support est l'expression du second argument de la lexie. Et tel est bien le cas :

Oper₂(coup) = *fam encaisser, fam manger, recevoir*

Passons maintenant à la seconde fonction lexicale syntagmatique de verbe support que j'ai choisi d'introduire ici.

- 2 La fonction lexicale **Func_i** (du verbe latin imaginaire *functionare* 'fonctionner') associée à une lexie prédicative nominale L l'ensemble des verbes supports qui prennent L comme sujet grammatical et prennent l'expression du i^e (premier, deuxième, etc.) argument de L comme d'objet direct ou indirect. Lorsque les verbes en question sont intransitifs, la fonction lexicale est appelée **Func₀**.



Correspondance sémantique-syntaxe définissant la fonction lexicale **Func_i**

Voici maintenant quelques exemples de **Func_i** ; je laisse au lecteur le soin de les analyser sémantiquement et syntaxiquement pour vérifier qu'ils correspondent bien à la situation décrite dans la figure ci-dessus :

Func₀(*applaudissements*) = *crépiter, résonner, retentir*

Func₁(*exclamation*) = *échapper* [à qqn.]

Func₂(*accusation*) = *peser* [sur qqn.]

Ici s'achève ma présentation des fonctions lexicales syntagmatiques. Comme on le voit, bien que les locuteurs utilisent les collocations de façon mécanique, leur modélisation et leur compréhension demande une très bonne maîtrise d'un ensemble important de notions linguistiques de base. Pour comprendre comment fonctionne une collocation, il faut être capable de mener une analyse en profondeur des structures sémantique et syntaxique qu'elle met en jeu.

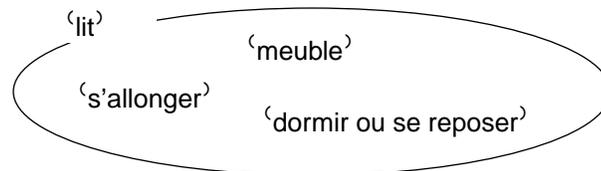
En guise de transition...

Le système des fonctions lexicales de la langue est très riche (plus d'une cinquantaine de fonctions lexicales, qui peuvent souvent se combiner entre elles). Il serait justifié de consacrer tout un cours de sémantique à cette notion : ses fondements, son utilisation potentielle en lexicologie, en enseignement des langues, etc. Je dois cependant m'arrêter ici. On trouvera en fin de chapitre plusieurs suggestions de lectures qui peuvent donner une meilleure idée de l'amplitude des phénomènes que modélisent les fonctions lexicales. On trouvera aussi dans les exercices plusieurs questions permettant d'aller au-delà des quelques cas de figure qui viennent d'être examinés.

De façon plus générale, nous en avons maintenant terminé avec l'examen du problème de la modélisation des relations lexicales. On voit bien que modéliser les relations entre lexies, même au moyen d'un outil aussi puissant que les fonctions lexicales, ne suffit pas à véritablement **décrire** le sens des lexies en question. C'est même plutôt l'inverse qui se produit : une véritable analyse des sens lexicaux est requise si l'on veut pouvoir identifier clairement les relations sémantiques entre lexies.

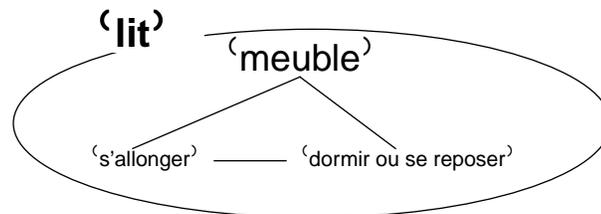
Reprenons notre exemple du sens 'lit' mentionné au début de ce chapitre (page 118). Sa représentation strictement ensembliste, sous forme de « patate », ressemblerait à quelque chose comme (23a) ci-dessous :

(23) a.



Cependant, cette visualisation du sens ne fait que mettre en évidence la relation d'inclusion. Il manque la modélisation de l'organisation interne du sens, quelque chose qui ressemblerait à :

b.



Ce que (23b) cherche à modéliser (à travers l'usage du gras et des caractères de tailles différentes), c'est en réalité la structure de la **définition** de la lexie LIT. La définition du sens d'une lexie n'est pas que la liste de tous les sens plus simples le constituant ; c'est une « formule » rendant explicite l'organisation interne de ce sens.

Seules des modélisations structurées sous forme de définitions véritables permettent d'explicitier les liens lexicaux du type de ceux étudiés dans le présent chapitre. Ainsi, pour pouvoir démontrer pourquoi une relation Syn_{\cap} existe entre LIT et HAMAC mais pas entre LIT et TABLE D'OPÉRATION (bien qu'une intersection de sens importante soit aussi présente dans ce second cas), il faut disposer d'une représentation structurée des sens.

En fait, tout le système des fonctions lexicales, pour être bien utilisé, présuppose que l'on est capable de mener à bien une analyse sémantique des lexies. La fonction lexicale **Magn**, par exemple, correspond nécessairement, dans un cas donné de $\text{Magn}(L)$, à l'intensification d'une

composante sémantique particulière du sens de la lexie L, composante qu'il faut être capable d'identifier. Voici, à fin d'illustration, deux séries d'exemples de **Magn**(ÉPIDÉMIE) mettant chacun en jeu une composante particulière du sens de cette lexie⁸ :

(24) a. *En 1894, une **effroyable épidémie** de peste fait rage à Hong Kong.*

→ **Magn** s'applique à la composante 'maladie'.

b. *Nous assistons à la plus **vaste épidémie** de peste qui ait ravagé le monde.
Depuis quelques années, l'ex U.R.S.S. est endeuillée par une très **grosse épidémie** de diphtérie.*

→ **Magn** s'applique à 'ensemble de personnes'.

Tous ces fait démontrent qu'il est maintenant grand temps d'aborder le problème central de la sémantique lexicale : l'analyse du sens. C'est ce que nous allons faire dans le chapitre suivant.

Lectures complémentaires

Palmer, F. R. (1981) Chapter 5 Lexical Semantics: Sense Relations. In : *Semantics*, Cambridge *et al.*: Cambridge University Press, pp. 83-108.

Je suggère de lire ce texte pour deux raisons. D'une part, il est très bien écrit et présente clairement les notions centrales examinées dans ce chapitre. D'autre part, il est écrit en anglais et propose des exemples dans cette langue. Il est en effet toujours bon de raisonner sur des exemples empruntés à une autre langue pour bien comprendre les problèmes de sémantique. On se distancie plus facilement d'une langue étrangère ; il est plus facile de la considérer uniquement comme un objet d'étude. On trouvera aussi dans ce texte une ébauche de formalisation de la représentation du sens au moyen du langage de la logique formelle. Sans qu'il faille nécessairement assimiler complètement cette méthode de description du sens, il est utile d'en comprendre le fonctionnement, pour le contraster notamment avec le formalisme des réseaux

8. Ces exemples ont été récupérés sur Internet.

sémantiques, introduit au chapitre précédent (Section *Représentation formelle du sens des énoncés*, page 109).

Mel'čuk, Igor A. (1997) *Vers une linguistique Sens-Texte*. Leçon inaugurale, Paris : Collège de France, pp. 41-57.

Mel'čuk, Igor A. (1994) Fonctions lexicales dans le traitement du langage naturel. In A. Clas & P. Bouillon (éd.) : *TA-TAO : Recherches de pointe et applications immédiates*. Beyrouth / Montréal : FMA/AUPELF-UREF, pp. 193-219.

Ces deux textes offrent une présentation relativement complète de la notion de fonction lexicale.

Exercices

- 1 Quelle relation ensembliste existe-t-il entre le sens de la lexie LIVRE [*J'ai lu un livre passionnant.*] et celui de la lexie POÈTE [*C'est un des plus grands poètes canadiens.*] ? En quoi la caractérisation des liens sémantiques sur la base de relations ensemblistes est-elle peu satisfaisante ?
- 2 Donner tous les hyperonymes de ARBRE [*On a planté un arbre.*]. Donner cinq de ses hyponymes.
- 3 Démontrer que EN ROUTE [*Le train est en route vers son terminus.*] est un antonyme de À L'ARRÊT [*Le train est à l'arrêt.*]. Pourquoi doit-on considérer que ces deux lexies sont des antonymes approximatifs ? (Penser au nombre d'arguments des prédicats correspondants.)
- 4 En se référant à l'Exemple (5a-b), expliquer pourquoi PRÈS et LOIN sont des antonymes exacts alors que AIMER et DÉTESTER sont des antonymes approximatifs.
- 5 La lexie FEU [*Ces hommes préhistoriques avaient déjà domestiqué le feu.*] est-elle un antonyme de la lexie EAU [*L'eau est nécessaire à la vie.*] ? Justifier.
- 6 Le vocable ROUTE est-il polysémique ? Justifier.
- 7 Lire attentivement la citation suivante. Comment pourrait-on caractériser la relation sémantique existant entre ŒIL et REGARD ? Comment cette relation est-elle exploitée ici par A. Nothomb ?

Les yeux des êtres vivants possèdent la plus étonnante des propriétés : le regard. Il n'y a pas plus singulier. On ne dit pas des oreilles des créatures qu'elles ont un « écoutard », ni de leurs narines qu'elles ont un « sentard » ou un « reniflard ».

Amélie Nothomb, *Métaphysique des tubes*

- 8 Décrire avec un maximum d'exactitude la relation sémantique existant entre S'ALIMENTER [*Jean s'alimente peu.*] et MANGER [*Jean mange peu.*].
- 9 Trouver en quoi les phrases (25) et (26) transgressent en partie le principe de compositionnalité sémantique :
- (25) *Quand il a appris la nouvelle, son sang n'a fait qu'un tour.*
- (26) *Ils l'avaient bien annoncé à la radio qu'il ferait un froid de canard.*
- 10 L'expression marquée en gras dans la phrase ci-dessous est-elle une expression idiomatique, semi-idiomatique ou sémantiquement compositionnelle ? Justifier la réponse donnée.
- (27) *Jean **court un danger.***
- 11 Même question avec la phrase ci-dessous :
- (28) *Devant la difficulté de la tâche, tout le monde **baisse les bras.***
- 12 Décrire les collocations ci-dessous au moyen de fonctions lexicales. Justifier :
- (29) *Il a un **cousin éloigné** qui vit en Ontario.*
- (30) a. *Un **grave différend** les oppose.*
b. *Un **léger différend** les oppose.*
- (31) a. *Il **dort sur ses deux oreilles.***
b. *Il **dort à demi.***
c. *Il **dort comme un bébé.***
- 13 Compléter les formules ci-dessous en donnant de façon aussi exhaustive que possible les valeurs retournées pour chacune des applications de fonctions lexicales. Je donne, lorsque cela est utile, des exemples permettant d'identifier l'acception qui doit être considérée.
- **Syn**(*manger* [*Il mange un steak saignant.*]) = ...
 - **Anti**(*permettre* [*Je lui permets d'utiliser mes skis.*]) = ...

- **S₀**(*tomber* [*Il est tombé sur le sol.*]) = ...
 - **S₂**(*acheter* [*Il a acheté une nouvelle chemise.*]) = ...
 - **Magn**(*pleurer* [*On a tous pleuré quand il est parti.*]) = ...
 - **AntiMagn**(*appétit* [*Cette ballade m'a ouvert l'appétit.*]) = ...
 - **Bon**(*temps* [*Quel temps fait-il ?*]) = ...
 - **AntiBon**(*temps*) = ...
- 14** Trouver les collocations présentes dans le texte ci-dessous ; les décrire au moyen de fonctions lexicales.
- Il entra dans cette pièce qu'il connaissait comme le fond de sa poche et s'aperçut aussitôt qu'une chaleur étouffante y régnait. Il fit un large sourire et dit de sa petite voix nasillarde : « Je sais que tu es là ! ».
- 15** L'expression *satisfaire une envie* est une collocation. Pourquoi n'est-ce pas une illustration d'un cas de **Oper₁** ?
- 16** Relire la citation donnée en exergue pour ce chapitre (page 117).
- Paraphraser en termes très simples et très généraux *éclate* dans *La guerre éclate*.
 - Pourrait-on décrire cette collocation au moyen des fonctions lexicales qui ont été introduites dans ce chapitre ?
 - Le jeu de mot lugubre *Les bombes hachent* n'est bien entendu pas une collocation du français. Essayer d'expliquer pourquoi ?
 - Quelles sont en revanche les collocations disponibles en français pour dire ce que « font » les bombes ?

Chapitre 8

L'analyse du sens



Définition lexicale, genre prochain (= sens générique) et différences spécifiques, définition analytique, (test de) substitution en contexte, circularité (dans une définition), cercle vicieux, champ sémantique, ambiguïté lexicale vs syntaxique, vague, analyse componentielle, sème, sémème, lexie de base (d'un vocable), sens propre vs figuré, causativité, métaphore, métonymie.

— Mais qu'est-ce que signifie « éphémère » ? répéta le petit prince qui, de sa vie, n'avait renoncé à une question, une fois qu'il l'avait posée.

— Ça signifie « qui est menacé de disparition prochaine ».

— Ma fleur est menacée de disparition prochaine ?

— Bien sûr.

« Ma fleur est éphémère, se dit le petit prince, et elle n'a que quatre épines pour se défendre contre le monde ! Et je l'ai laissée toute seule chez moi ! »

Antoine de Saint-Exupéry, *Le Petit Prince*

Cette citation de Saint-Exupéry illustre parfaitement les deux faits suivants, qui seront le point de départ du présent chapitre :

- 1 pour rendre apparent le sens d'une lexie, comme ÉPHÉMÈRE, il faut l'analyser ;
- 2 une analyse typique consiste en une décomposition de l'élément analysé en éléments « plus simples » qui le constituent.

Le but de ce chapitre est de donner une bonne vue d'ensemble des moyens dont on dispose pour analyser les sens lexicaux et des problèmes spécifiques que pose l'analyse sémantique lexicale.

Je procéderai en quatre étapes :

- 1 étude de l’outil privilégié d’analyse sémantique des lexies : la définition lexicale ;
- 2 examen rapide d’une méthode concurrente d’analyse du sens : l’analyse dite « componentielle » ;
- 3 étude des principales relations sémantiques entre lexies d’un même vocable, relations que les définitions lexicales doivent mettre en évidence ;
- 4 exercice pratique portant sur la relation métaphorique.

La définition lexicale : outil d’analyse du sens

Définition par genre prochain et différences spécifiques

Il existe plusieurs méthodes d’analyse des sens lexicaux. La méthode d’analyse sémantique la plus communément utilisée est celle basée sur l’écriture de **définitions lexicales** présentant les trois caractéristiques suivantes :

- 1 elles proposent une paraphrase du sens de la lexie définie ;
- 2 elles sont formulées à l’aide de lexies sémantiquement plus simples que la lexie définie ;
- 3 elles se subdivisent en deux parties :
 - une composante centrale appelée **genre prochain** ou **sens générique** — qui est en quelque sorte une paraphrase approximative minimale du sens de la lexie définie ;
 - un ensemble de composantes sémantiques périphériques appelées **différences spécifiques** — qui caractérisent le sens de la lexie définie par rapport au sens de toutes les autres lexies de la langue ayant le même genre prochain.

La dernière caractéristique mentionnée ci-dessus donne son nom au type de définition que nous étudierons : la **définition par genre prochain et différences spécifiques**. Ce terme est très ancien ; en fait, les notions de *genre* et de *différences spécifiques* ont été proposées par Aristote (384-322 avant J.-C.) pour fonder sa théorie de la définition, exposée dans *Les Topiques* (voir référence exacte dans la liste de lectures

en fin de chapitre). Il est à noter que l'on trouve souvent dans la littérature linguistique le terme *définition par genre prochain et différences spécifiques* écrit avec *différence spécifique* au singulier. Il me semble cependant que l'emploi du pluriel est préférable, dans la mesure où il rend compte du fait que l'on a normalement affaire ici à un ensemble de composantes sémantiques caractérisant la lexie définie. Ce détail terminologique n'aura de toute façon pas d'incidence dans ce qui suit, puisque, pour alléger, j'appellerai dorénavant les définitions en question **définitions analytiques**.

La définition (1a) ci-dessous est une définition analytique telle qu'on pourrait la trouver dans un dictionnaire comme le *Petit Robert*. Par contraste, (1b) est, plutôt qu'une définition, une simple énumération de synonymes approximatifs :

- (1) a. **LABEUR** *travail*_[genre prochain] *long et pénible*_[diff. spécifiques]
b. **LABEUR** *travail, corvée, besogne*

La plupart des dictionnaires de langue proposent des définitions analytiques, mais ils ne le font pas toujours de façon systématique, comme nous le verrons dans le Chapitre 9.

Avant de passer à la méthodologie d'élaboration des définitions analytiques, il convient de noter qu'il existe au moins deux familles de lexies qu'il n'est pas possible de définir de cette façon :

- 1 Tel que mentionné au Chapitre 6 (page 106), les lexies véhiculant un sens grammatical, comme les articles ou les verbes auxiliaires, ne peuvent être véritablement paraphrasées. On ne peut donc pas les définir au moyen de définitions analytiques. Leur sens sera généralement décrit en faisant référence directement à des notions de grammaire (tel article exprime le défini, tel autre l'indéfini, etc.).
- 2 Les interjections (OH !, ZUT !, etc.) sont des lexies particulières qui ne sont pas non plus paraphrasables au moyen de définitions analytiques. On pourra se reporter au petit texte donné dans l'Exercice 2 du Chapitre 6, où justement la difficulté de paraphraser le juron québécois VERRAT ! peut expliquer la relative maladresse avec laquelle un petit enfant essaie de le décrire.

Méthodologie d'élaboration d'une définition

Aristote, qui a établi de façon explicite les fondements théoriques de la notion de définition par genre prochain et différences spécifiques, a très clairement indiqué toute la difficulté inhérente à l'activité de définir :

Qu'il soit plus difficile d'établir une définition que de la ruiner, on le verra clairement d'après ce que nous allons dire.

Aristote, *Les Topiques*, Livre VII, Chapitre 5.

Il faut bien comprendre que c'est toute l'activité scientifique qui se fonde sur la construction et l'utilisation de définitions. Que ce soit dans l'élaboration de notions scientifiques ou dans leur application proprement dite, la réfutation (d'une définition, de la validité d'une description ou d'une expérimentation, etc.) est toujours plus facile que la « construction ». Et on sait que la preuve de la validité d'une construction scientifique est, quant à elle, impossible à établir : on ne peut pas démontrer qu'une définition ou une modélisation est valide ; on ne peut qu'échouer ou réussir en tentant de montrer qu'elle est invalide. Cela est bien entendu aussi vrai dans le cas de la définition lexicale. Un corollaire de la citation d'Aristote, donnée ci-dessus, est qu'il est plus aisé d'élaborer une méthode permettant de falsifier une définition que d'élaborer une méthode permettant de bien définir.



C'est seulement par l'acquisition d'un ensemble complexe de techniques bien précises, et surtout par la pratique, que l'on apprend à bien définir les lexies.

Le présent ouvrage n'est pas un manuel de lexicographie, contrairement à Mel'čuk *et al.* (1995) mentionné plus bas dans la liste des lectures complémentaires. Je vais donc me contenter de donner ci-dessous une ébauche de méthode que l'on peut suivre lorsque l'on cherche à élaborer une bonne définition analytique. Cette ébauche est assez grossière ; elle a cependant le mérite d'offrir une grille d'analyse sémantique solide, applicable toutes les lexies de la langue.

Nous examinerons le cas de la lexie DÉVISAGER [*Il me dévisageait d'un air mauvais.*], dont nous allons construire la définition en cinq étapes.

- 1 **Identification de la nature prédicative de la lexie** La première question à se poser est de savoir si la lexie que l'on définit est un prédicat sémantique et, si oui, quel peut être son nombre d'arguments. DÉVISAGER, en tant que verbe, est bien évidemment un prédicat et on peut faire l'hypothèse qu'il a deux arguments : X qui dévisage et Y qui est dévisagé. Ce que l'on va définir maintenant, ce n'est donc pas tout simplement 'dévisager' mais 'X dévisage Y'. Notre définition analytique devra être une paraphrase de cette proposition simple.
- 2 **Identification du genre prochain** Il faut maintenant trouver le genre prochain de la lexie, sa paraphrase minimale. Dans le cas de DÉVISAGER, le problème est facile à résoudre : dévisager c'est **regarder** d'une certaine façon. On peut donc proposer une première paraphrase approximative définissant grossièrement cette lexie :

(2) 'X dévisage Y' \equiv 'X regarde Y d'une certaine façon'.

On peut tester la validité de cette définition en montrant (i) que si l'on dévisage cela veut nécessairement dire que l'on regarde et (ii) que l'on peut tout à fait regarder sans dévisager. Le premier point est démontré par l'incohérence sémantique de la phrase (3a) ci-dessous et le second par le fait que (3b) est, quant à elle, sémantiquement tout à fait cohérente :

- (3) a. [#]Léo dévisageait Sylvain sans le regarder.
b. Léo regardait Sylvain, mais sans toutefois oser le dévisager.



Je fais précéder les exemples sémantiquement incohérents par le symbole dièse (#). Il est important d'utiliser ici un symbole différent de celui de l'agrammaticalité (l'astérisque *). En effet, la phrase (3a) est tout à fait grammaticale et l'anomalie qu'elle contient est strictement sémantique.

- 3 **Caractérisation sémantique des arguments** Avant de passer à l'identification des différences spécifiques, qui distinguent DÉVISAGER de REGARDER, nous pouvons essayer de caractériser sémantiquement les arguments de notre prédicat en répondant à ces deux questions :

Qu'est-ce qui peut dévisager ?
Qu'est-ce qui peut être dévisagé ?

Les exemples ci-dessous montrent que X peut être soit un individu soit un animal — mais pas un objet, comme une caméra — et que Y ne peut être qu'un individu :

- (4) a. *Le voisin me dévisageait d'un air mauvais.*
b. *Le chien du voisin me dévisageait d'un air mauvais.*
c. *#La caméra de sécurité dévisageait les visiteurs.¹*
- (5) a. *Il dévisageait le voisin.*
b. *#Il dévisageait le chien du voisin.*
c. *#Il dévisageait un tableau de Renoir.*

On peut remarquer que REGARDER ne se comporte pas de la même façon que DÉVISAGER, puisque le deuxième argument de ce verbe peut être absolument n'importe quelle entité visible :

- (6) *Il regardait le voisin/un chien/un tableau/le ciel.*

Il faudra aussi que notre définition mette en évidence le fait que le premier argument de DÉVISAGER ne peut pas être un objet, car cela distingue cette lexie d'un verbe sémantiquement très proche comme SCRUTER. Comparez l'exemple ci-dessous avec (4c) :

- (7) *La caméra de sécurité scrutait les visiteurs.*

Cela démontre que les caractéristiques sémantiques des arguments correspondent en fait à une première série de composantes de sens constituant les différences spécifiques !

Nous pouvons maintenant affiner notre première définition de DÉVISAGER de la façon suivante :

1. Bien entendu, on peut toujours produire une telle phrase. On pourrait même la considérer comme étant très bien tournée. Mais on va nécessairement percevoir une sorte d'effet de style dans le fait de dire qu'une caméra dévisage quelqu'un : attribution d'un « comportement humain » à un objet. Le même genre de remarque s'applique aux deux autres phrases précédées de # : (5b) et (5c).

(8) 'X dévisage Y' \equiv 'L'individu ou l'animal X regarde l'individu Y d'une certaine façon'.

- 4 **Identification des autres différences spécifiques** Nous devons maintenant remplacer ce 'd'une certaine façon', qui est bien trop vague, par une ou plusieurs composantes sémantiques qui caractérisent véritablement l'acte de dévisager par rapport à celui de regarder. En fait, nous devons faire une définition qui va aussi distinguer DÉVISAGER de SCRUTER, puisque ce sont deux synonymes approximatifs, co-hyponymes de REGARDER. Tout d'abord, on voit que DÉVISAGER et SCRUTER signifient toutes deux 'regarder avec une certaine intensité'. On pourrait dire que dévisager c'est regarder avec beaucoup d'attention alors que scruter est plus que cela : c'est regarder avec beaucoup d'attention comme si on cherchait à trouver ou à voir quelque chose. La phrase (9) montre que l'on peut dévisager simplement en signe de réprobation, sans qu'il soit question de chercher à voir quelque chose de précis :

(9) *Léo en était à son troisième verre de scotch et Sylvain le dévisageait d'un air à la fois surpris et réprobateur.*

Un autre fait important qu'il faut remarquer à propos de DÉVISAGER, c'est que, précisément, on **dévisage** quelqu'un en regardant son **visage** :

(10) a. *#Sylvain dévisageait Léo de la tête aux pieds.*

b. *#Sylvain dévisageait les mains de Léo.*

(11) a. *Sylvain scrutait Léo de la tête aux pieds.*

b. *Sylvain scrutait les mains de Léo.*

Dans le cadre d'une véritable analyse sémantique fouillée, il serait nécessaire de pousser plus loin notre investigation pour voir si nous avons bien identifié toutes les composantes du sens de DÉVISAGER². Je vais cependant présupposer que le travail est complété et que nous pouvons maintenant écrire la définition finale de notre lexie :

(12) 'X dévisage Y' \equiv 'L'individu ou l'animal X regarde très attentivement le visage de l'individu Y'.

2. Je propose en fait, à la fin de ce chapitre, deux exercices dans lesquels il est demandé d'améliorer cette définition (voir Exercices 1 et 2).

- 5 **Validation de la définition** Une dernière étape de la construction d'une définition est d'utiliser certains tests permettant de vérifier sa validité. Je propose ici le plus connu de ces tests ; il s'agit du **test de substitution en contexte**, qui permet de voir si la définition est une paraphrase valide de la lexie définie en la substituant à celle-ci dans différents contextes :

- (13) a. *Léo dévisageait Sylvain avec envie.*
b. *Léo regardait très attentivement le visage de Sylvain avec envie.*

On peut considérer que (13b) est une bonne paraphrase de (13a), même si elle est stylistiquement assez lourde. Bien entendu, en effectuant la substitution, on doit laisser de côté les composantes de sens qui ne sont pas pertinentes ; il s'agit ici des contraintes sur la valeur sémantique des arguments du prédicat.



Il serait possible d'aller plus loin dans le processus de validation, en cherchant notamment des emplois de DÉVISAGER dans des corpus de textes. Il s'agirait alors de vérifier que notre définition est compatible avec chacun des contextes sémantiques trouvés dans ces corpus.

Le problème des cercles vicieux

Un écueil à éviter à tout prix lorsque l'on fait une définition analytique est de se retrouver avec des cas de **circularité**. Je fais ici référence aux fameux **cercles vicieux**, sur lesquels nous reviendrons dans le Chapitre 9, à propos des définitions des dictionnaires commerciaux. Il y a cercle vicieux lorsque l'on utilise dans la définition d'une lexie L_1 la lexie L_2 qui elle-même va être définie directement ou indirectement (c'est-à-dire, via la définition d'une de ses composantes sémantiques) au moyen de L_1 .

Par exemple, il y a un cercle vicieux dans les définitions d'un dictionnaire s'il décrit ainsi le sens des lexies SCIER et SCIE :

- (14) a. 'X scie Y' ≡ 'X coupe Y avec une scie'
b. '[une] scie' ≡ 'instrument servant à scier'

(15) 'stylo' \cong 'instrument_[= genre prochain] servant à écrire grâce à l'encre qu'il contient'.

Les lexies regroupées dans un même champ sémantique n'appartiennent pas nécessairement à la même partie du discours. Par exemple, STYLO, CRAYON, PAPIER, etc. mais aussi ÉCRIRE, RÉDIGER, etc. appartiennent au champ sémantique de l'écriture.

De plus, un champ sémantique est un regroupement de **lexies**. Cela n'aurait aucun sens de regrouper des vocables au complet dans un champ sémantique. Par exemple, TABLE III.2 [*Il affiche tous ses résultats dans une table à quatre colonnes.*] du *Nouveau Petit Robert* n'appartient pas au champ sémantique des meubles.

La notion de champ sémantique est très utile en sémantique et en lexicologie. Ainsi, une étude de la sémantique lexicale conduite par champs sémantiques mènera nécessairement à de bien meilleurs résultats qu'une étude qui se ferait en choisissant les lexies une à une de façon arbitraire (selon l'ordre alphabétique, par exemple). Nous avons déjà eu un aperçu de ce phénomène lorsque nous avons mis en contraste SCRUTER et DÉVISAGER, pour construire la définition de cette dernière lexie. Il faudrait en fait considérer à la fois **toutes** les lexies sémantiquement liées à DÉVISAGER — des synonymes approximatifs comme SCRUTER aussi bien que des lexies comme REGARD, OBSERVATION, EXAMEN, etc. — pour en faire une étude qui révèle sa véritable valeur linguistique (sur la notion de valeur, voir Chapitre 6).

Pour conclure sur ce sujet, il faut souligner qu'une lexie peut tout à fait être considérée comme appartenant simultanément à plusieurs champs sémantiques. Ainsi, la lexie GARAGISTE appartient à la fois au champ sémantique des professions et à celui de l'automobile.

Remarque sur les notions d'ambiguïté et de vague

Lorsque l'on cherche à définir une lexie, la première chose à faire est de l'isoler en tant qu'élément distinct du lexique de la langue. C'est pourquoi je fais presque systématiquement suivre le nom des lexies sur lesquelles nous travaillons d'un exemple entre crochets, qui permet notamment de distinguer cette lexie des autres acceptions d'un même

vocable (ou des lexies homonymes). Il s'agit en fait ici de résoudre une possible ambiguïté :

Une expression est **ambiguë** lorsqu'elle peut être associée à plus d'un sens.

Ainsi, *dévisager* n'est pas ambigu alors que *scruter* l'est, comme le démontrent les deux exemples suivants :

- (16) a. *Il scrutait le paysage.*
b. *Nous devons scruter attentivement ses intentions avant de le prendre comme partenaire sur ce projet.*

On fait généralement la distinction, dans les ouvrages de sémantique, entre l'**ambiguïté lexicale** — qui tire son origine de la possibilité d'associer une expression donnée à plus d'une lexie — et l'**ambiguïté syntaxique** — qui se manifeste dans la possibilité de mener plusieurs analyses syntaxiques différentes d'une même expression (généralement, une phrase). Par exemple, la phrase *Il me parle de la chambre* est syntaxiquement ambiguë car elle peut être associée à deux structures syntaxiques⁴ :

- 1 soit *de la chambre* est le complément d'objet indirect de *parle* → la chose dont il me parle est la chambre ;
- 2 soit *de la chambre* est un complément circonstanciel de *parle* → il est dans une autre pièce, la chambre, lorsqu'il me parle.

Dans ce qui suit, il ne va être question que d'ambiguïté lexicale.



L'ambiguïté est un problème que le lexicologue doit résoudre dans son analyse, puisqu'il est impossible de produire une bonne définition lexicale si l'on n'a pas auparavant isolé de la façon la plus précise

4. Noter que pour certains auteurs la notion d'ambiguïté syntaxique n'est pas nécessairement liée à la présence d'interprétations sémantiques distinctes. Il suffit pour eux que plusieurs structures syntaxiques puissent être identifiées pour qu'une phrase soit syntaxiquement ambiguë, les analyses en question pouvant s'avérer équivalentes sur le plus sémantique. Je n'approfondirai pas cet aspect de la question ; l'exemple donné ici met en jeu une différence d'analyse syntaxique s'accompagnant d'une différence sémantique.

possible la lexie particulière sur lequel on veut travailler. En effet, l'ambiguïté est un phénomène « parasite », découlant de la possible polysémie des vocables (ou de l'homonymie) ; ce n'est en aucune façon une propriété d'une lexie. Il serait par exemple tout à fait illogique de dire d'une lexie qu'elle est ambiguë : une lexie est associée à un et un seul sens et seules les **formes** peuvent être ambiguës.

Il n'est pas toujours évident de déterminer si une forme est ambiguë car les sens lexicaux ne sont pas nécessairement très spécifiques. Examinons les exemples suivants :

(17) a. *J'ai vu ton cousin Marcel.*

b. *Il y avait un énorme cousin sur le mur du salon.*

On n'aura vraisemblablement aucun problème à identifier que le cousin dont il est question en (17a) est un individu alors qu'il s'agit d'un moustique dans la phrase suivante. Bien entendu, il se peut que l'on ne sache pas qu'un cousin est un type de moustique ; mais dans ce cas, on trouvera la seconde phrase bizarre et on sera peut-être amené à faire l'hypothèse qu'il existe un sens de *cousin* que l'on ne connaît pas. Ce qui importe pour nous ici, c'est que personne ne songera à considérer que la même lexie COUSIN a été employée dans les deux phrases ci-dessus, une fois établi qu'il est bien question d'un animal en (17b). Il en va ainsi parce que l'écart sémantique entre les deux sens pouvant être exprimés par *cousin* est très important. Cet écart nous empêcherait notamment d'énoncer la phrase suivante pour signaler qu'un individu et un animal se trouvent dans une même pièce (à moins, bien sûr, de faire un jeu de mots) :

(18) *Il y avait deux cousins dans la pièce.*

En d'autres termes, on ne peut employer une seule fois *cousin(s)* dans une phrase pour désigner simultanément un parent et un moustique. Nous pouvons donc affirmer sans hésiter qu'il existe en français deux homonymes COUSIN¹ ('individu') et COUSIN² ('moustique'), apparaissant respectivement dans les exemples (17a) et (17b).

Mais examinons maintenant la définition de COUSIN¹ donnée dans le *Nouveau Petit Robert* :

Descendant d'un frère ou d'une sœur par rapport aux descendants d'un frère, d'une sœur de l'un de ses parents.

Clairement, COUSIN¹ a un sens très complexe et cette lexie peut-être utilisée pour dénoter des individus de notre famille auxquels nous sommes liés de multiples façons. Mon cousin peut ainsi être le fils du frère de ma mère, le fils de la sœur de mon père, etc. Alors, pourquoi ne pas considérer qu'il existe plusieurs lexies COUSIN^{1a}, COUSIN^{1b}, etc., chacune dénotant un type de lien familial spécifique ? Tout simplement parce que *cousin* peut-être utilisé pour dénoter **à la fois** toutes ces relations familiales. On peut ainsi tout à fait dire :

(19) *Je te présente mes deux cousins : le fils de mon oncle Alfred et le fils de ma tante Léontine.*

On dira que la lexie COUSIN¹ est vague :

Le sens d'une lexie (ou d'une expression lexicale) est considéré comme **vague** s'il peut désigner des faits ou des entités qui nous semblent relativement distincts, par comparaison à d'autres sens considérés comme plus spécifiques. On dira indifféremment qu'une lexie ou que le sens d'une lexie est vague.

La définition ci-dessus rend explicite le fait que le caractère vague du sens d'une lexie est une propriété relative. On va considérer COUSIN¹ comme étant vague parce qu'il nous semble de cette lexie peut renvoyer à des relations familiales somme toute assez différentes. Mais pour un locuteur d'une langue dans laquelle il n'y aurait pas, par exemple, de terme distinct pour désigner cousins, cousines, frères et sœurs, la lexie française COUSIN¹ pourrait sembler au contraire relativement spécifique.

Il est très naturel de s'interroger sur le caractère vague de COUSIN¹ car la définition de cette lexie comporte des disjonctions (*frère ou sœur* et *frère, sœur*, dans la définition donnée plus haut). Cependant, pratiquement toute lexie pourrait, dans un contexte donné, sembler vague (ou, au contraire, spécifique). Une conséquence de cela est qu'il est généralement difficile de dire qu'une lexie est vague dans l'absolu. Par exemple, PARLER [*Il lui parle gentiment.*] est vague si on la compare à CHUCHOTER, MURMURER, SUSURRER, HURLER, etc. ; mais elle est à l'inverse relativement spécifique si on la compare à COMMUNIQUER (un hyperonyme).

Pour conclure sur ce point, il faut noter que si presque tout le monde s'accorde pour dire que la distinction entre ambiguïté et vague est fondamentale en sémantique, la modélisation de ces phénomènes peut varier considérablement. Il est notamment assez rare de voir explicitement présenté le caractère relatif de la notion de vague. De plus, on omet fréquemment d'indiquer de façon claire que l'ambiguïté est une propriétés de **formes** linguistiques alors que le vague est une propriété relative du **sens** des lexies.

Analyse componentielle

Nous examinerons maintenant brièvement une approche de la description des sens lexicaux qui se veut une alternative à la définition lexicographique.

L'*analyse componentielle* du sens des lexies est une « métaphore scientifique » de la caractérisation des phonèmes d'une langue par traits distinctifs. Je réfère ici au type d'analyse qui permet, par exemple, de contraster les deux phonèmes /b/ et /p/ de la façon suivante :

$$/b/ : \begin{bmatrix} - \text{vocalique} \\ + \text{occlusif} \\ + \text{bilabial} \\ + \text{sonore} \end{bmatrix} \quad /p/ : \begin{bmatrix} - \text{vocalique} \\ + \text{occlusif} \\ + \text{bilabial} \\ - \text{sonore} \end{bmatrix} .$$

Dans le cadre de l'analyse componentielle, les traits distinctifs sont appelés **sèmes** ; ils sont de nature sémantique (au lieu d'être de nature phonique ou articulatoire, comme c'est le cas pour les traits distinctifs des phonèmes). Le sens d'une lexie, appelé **sémème**, est modélisé comme une matrice de sèmes, qui caractérise cette lexie par rapport aux autres lexies appartenant à un même champ sémantique.

Voici un exemple d'analyse componentielle, qui permet de contraster le sens de BICYCLETTE et celui de MOTO :

$$\text{'bicyclette'} : \begin{bmatrix} + \text{véhicule} \\ + \text{deux-roues} \\ - \text{motorisé} \end{bmatrix} \quad \text{'moto'} : \begin{bmatrix} + \text{véhicule} \\ + \text{deux-roues} \\ + \text{motorisé} \end{bmatrix} .$$

L'analyse componentielle est une méthode de description des sens lexicaux qui n'a qu'une apparence de rigueur. En réalité, elle est le plus

souvent utilisée de façon *ad hoc* et ne permet pas de construire des descriptions rigoureuses, cohérentes et véritablement explicatives. Cela est dû essentiellement au caractère binaire des traits distinctifs utilisés. Parce que je savais qu'il me fallait contraster les lexies BICYCLETTE et MOTO, j'ai « inventé » les traits [\pm véhicule], [\pm deux-roues] et [\pm motorisé]. Or, si l'on peut aisément interpréter [+ deux-roues] comme signifiant 'qui possède deux roues', on voit mal quelle pourrait être une bonne interprétation de [- deux-roues] : 'qui ne possèdent pas deux roues' semble absurde. Le caractère binaire des traits distinctifs nous force à utiliser des descriptions qui n'ont pas d'interprétation intuitive immédiate.

On voit que l'analyse componentielle, loin de proposer une véritable analyse du sens lexical qui en dégagerait les **composantes**, ne fait souvent qu'associer à celui-ci une matrice de traits, choisis de façon plus ou moins arbitraire. C'est pourquoi cette technique de description, au lieu d'être nommée *analyse componentielle* (c'est-à-dire, une analyse en terme de composantes), devrait plutôt être désignée sous le nom d'*analyse sémantique par traits distinctifs*. En théorie, chaque trait pourrait être associé à une composante sémantique définitionnelle. Cependant, on ne peut pas tester la pertinence d'une analyse componentielle comme on le fait avec une définition analytique : comment faire un test de substitution en contexte avec une matrice de traits binaires ?

Je n'irai pas plus loin dans la présentation de l'analyse componentielle et je propose en lecture complémentaire un texte d'A. Wierzbicka, qui en fait une critique très bien articulée.

Structure sémantique des vocables

Lorsque l'on définit une lexie, il faut bien entendu définir en même temps ses synonymes approximatifs — nous venons de le voir au début de ce chapitre —, mais il faut aussi et surtout définir l'ensemble des lexies qui appartiennent avec elle au même vocable.



En effet, toutes les lexies d'un vocable partagent des composantes sémantiques et leurs définitions devront mettre clairement en évidence

les différents « glissements de sens » qui se produisent lorsque l'on passe d'une acception du vocable à une autre.

Les lexies d'un même vocable peuvent manifester des relations de sens très variées. Il est courant d'identifier dans chaque vocable une **lexie de base**, dont le sens est considéré comme « premier » : c'est la source dont sont « dérivées » sémantiquement les autres acceptions. La détermination de quelle est, parmi toutes les acceptions, la lexie de base d'un vocable est un problème très délicat dans lequel je n'entrerai pas ici. On notera simplement que le sens de la lexie de base est souvent désigné sous le nom de **sens propre**, par opposition aux sens qui en seraient dérivés, généralement appelés **sens figurés** lorsqu'ils correspondent à des métaphores ou à des métonymies (voir ci-dessous).

Les trois liens sémantiques suivants sont très fréquemment rencontrés entre lexies d'un même vocable :

1 Lien de **causativité**

Une lexie L_2 est liée par un lien de causativité à une lexie L_1 du même vocable si

$$\langle L_2 \rangle \cong \langle \text{causer } L_1 \rangle.$$

Par exemple, $\langle \text{casser}^1 \rangle$ [*La branche a cassé.*] cohabite en français avec son causatif $\langle \text{casser}^2 \rangle$ [*Jules a cassé la branche.*] :

$$\langle \text{Jules a cassé}^2 \text{ la branche} \rangle \cong \langle \text{Jules a causé que la branche casse}^1 \rangle.$$

Nous avons déjà examiné cette relation de sens entre lexies d'un même vocable dans le chapitre précédent, lorsque nous avons traité de la polysémie (voir page 128). Il est donc inutile de s'étendre plus ici sur la causativité. Remarquons simplement que cette relation se manifeste très fréquemment à l'intérieur du lexique des langues — pas nécessairement, d'ailleurs, entre lexies d'un même vocable — et qu'on lui a donc associé une fonction lexicale : **Caus**. Cette fonction lexicale peut encoder un lien de causativité « pure » entre lexies, mais elle est aussi souvent utilisée en combinaison avec d'autres fonctions lexicales pour former des liens lexicaux plus riches :



Attention ! Il ne faut pas confondre les métaphores/métonymies lexicalisées et les métaphores/métonymies libres, qui résultent de l'utilisation de procédés stylistiques. Ainsi, si je dis :

(20) *Ce type est une vieille poire blette.*

parce que je trouve que l'individu en question est mou et un peu dégoûtant, je fais preuve de créativité et le processus mis en jeu intéresse très peu l'étude de la sémantique lexicale. Par contre, il est clair qu'on ne peut décrire ce processus sans référer aux sens des lexies sur lesquelles il s'appuie. Notamment, il faut savoir que BLETTE exprime un type d'**AntiBon** de *poire*.

La phrase suivante relève d'un cas tout différent de (20) ; elle va nous permettre de contraster métaphore libre et métaphore lexicalisée :

(21) *Il a reçu un coup sur la poire.*

POIRE₃ (numérotation du *Nouveau Petit Robert*) est une lexie du langage familier, une métaphore lexicalisée basée sur une vague ressemblance de forme entre la poire₁ (le fruit) et la tête d'un individu.

Même si l'on pense qu'une tête a plus une forme de melon qu'une forme de poire, on ne dira certainement jamais (22) ci-dessous, à moins de vouloir produire un effet de style particulier :

(22) [#]*Il a reçu un coup sur le melon.*

Il existe bien entendu une lexie métaphorique MELON₃ (cf. *Nouveau Petit Robert*), qui pourrait avoir été utilisée dans (22) ; mais cette lexie signifie 'chapeau' et non 'tête' !

Pour démontrer le caractère lexicalisé vs non lexicalisé d'une métaphore ou d'une métonymie, on peut bien sûr s'en rapporter à l'« autorité suprême » du dictionnaire, comme je viens de le faire. Mais d'abord, on n'a pas toujours de bon dictionnaire sous la main au moment où on en a besoin ; ensuite, les dictionnaires sont des produits de l'activité humaine et, comme tels, ils sont nécessairement faillibles. On ne doit pas leur faire aveuglément confiance. Acquérir la maîtrise de la lexicologie, c'est se donner les moyens de trouver par soi-même des réponses à certaines questions concernant le lexique. Nous allons donc, comme exercice concluant ce chapitre, examiner un autre exemple de métaphore et voir

comment on peut déterminer s'il s'agit d'une métaphore libre ou lexicalisée.

Exercice pratique sur la métaphore

Présentation de l'exercice

Nous allons examiner une phrase, extraite d'un roman, où le personnage principal, que sa compagne a récemment quitté, décrit son état d'esprit. L'expression linguistique métaphorique qui va nous intéresser apparaît en gras :

Je n'éprouvai ni passion, ni jalousie, ni nostalgie. Je me sentais vide, lucide, propre et limpide **comme une casserole d'aluminium**.

Eco, Umberto (1990) *Le Pendule de Foucault*. « Le Livre de Poche », n° 4301, Paris: Grasset, p. 220.

Il ne s'agit pas ici de faire une analyse littéraire de cette métaphore, de l'évaluer du point de vue du style, des effets subtils qu'elle peut avoir sur le lecteur, etc. Nous voulons simplement savoir si elle appartient ou non au lexique du français. Si ce n'est pas le cas, on considérera qu'elle est une création de l'auteur (ou de quelqu'un à qui l'auteur aurait emprunté la métaphore en question). La distinction est essentielle. En effet, si cette métaphore appartient d'une façon ou d'une autre au lexique du français, il faut l'enseigner et l'apprendre en tant que telle. Si elle est une création, son étude relève de l'étude des figures de style. On peut regarder comment elle a été construite, comment elle fonctionne et comment on peut créer d'autres figures de style sur le même modèle. Dans ce dernier cas, son étude ne relève qu'indirectement de la lexicologie.

La métaphore à l'étude peut appartenir au lexique français de deux façons : en tant que lexie de la langue ou en tant qu'élément de la combinatoire d'une ou plusieurs lexies, c'est-à-dire en tant que valeur de fonction lexicale. Nous allons examiner à tour de rôle chacune de ces deux possibilités.

La métaphore est-elle lexicalisée ?

La question qui nous intéresse maintenant est de savoir si, dans la citation ci-dessus, nous sommes en présence d'une acception

métaphorique du vocable CASSEROLE. Les vocables désignant des objets peuvent avoir des acceptions servant à désigner des individus possédant certaines propriétés (physiques, de comportement, etc.) que l'on peut associer aux objets en question. C'est le cas par exemple du vocable ROC :

- (23) a. *Il a escaladé ce roc sans trop d'effort.*
b. *Umberto est un roc.*

On interprète immédiatement la phrase (23b) sans qu'il soit besoin de proposer un contexte pour cela car elle met en jeu une acception du vocable ROC. Cette dernière est une lexie à part entière de la langue, connue de la plupart des locuteurs du français. En revanche, si je veux que mon interlocuteur me comprenne lorsque j'énonce la phrase suivante :

- (24) *Umberto est une casserole.*

il va falloir que je m'assure que le contexte lui permettra de faire une certaine interprétation. Une casserole, c'est soit un ustensile de cuisine, soit un mauvais instrument de musique. La difficulté que l'on a à immédiatement interpréter (24) démontre qu'il n'existe pas d'acception de CASSEROLE dénotant un individu se trouvant dans un certain état d'esprit.

On n'a donc pas de lexie métaphorique CASSEROLE permettant d'expliquer le fonctionnement de la phrase de U. Eco. Mais peut-être est-ce parce que la métaphore en question s'appuie sur une lexie qui ne correspond pas au seul mot-forme *casserole*. Il pourrait s'agir d'une locution nominale CASSEROLE D'ALUMINIUM — sur le modèle de CHEMIN DE FER [*Le chemin de fer passe dans la vallée.*] — ou d'une locution adjectivale (ou adverbiale) COMME UNE CASSEROLE D'ALUMINIUM — sur le modèle de COMME UN POISSON DANS L'EAU [*Il était parmi nous comme un poisson dans l'eau.*]. Cependant, les phrases ci-dessous posent les mêmes difficultés d'interprétation que (24) ci-dessus :

- (25) a. *Umberto était une casserole d'aluminium.*
b. *Umberto se sentait comme une casserole d'aluminium.*

Nous avons donc démontré que la métaphore en question ne reposait pas sur l'existence d'une lexie française. Examinons maintenant la seconde possibilité.

La métaphore est-elle une valeur de fonction lexicale ?

Nous avons vu — Chapitre 3, Section *Étude des lexies et étude de la sémantique des langues*, page 43 — que les lexies devaient être décrites sur le modèle des signes linguistiques en fonction de leur signifié, leur signifiant et leurs propriétés de combinatoire. Les collocations contrôlées par une lexie donnée font partie de sa combinatoire.

Or justement, un grand nombre de lexies contrôlent l'usage d'expressions métaphoriques pour exprimer certains sens bien particuliers au sein de collocations. Ces expressions métaphoriques sont, en d'autres termes, des valeurs de fonctions lexicales. Ainsi, pour exprimer le **Magn** de BROUILLARD, on peut utiliser notamment *épais brouillard* ou *brouillard à couper au couteau*. Dire qu'un brouillard est épais ou est à couper au couteau relève clairement de la métaphore. Ces métaphores ne sont pas à proprement parler des unités lexicales en elles-mêmes car leur emploi est très restreint et relativement imprévisible ; elles fonctionnent comme collocatifs de BROUILLARD au sein de collocations. Par exemple, on ne dit ni **épaisse pluie* ni **pluie à couper au couteau* ; on dit *épaisse fumée* mais pas **fumée à couper au couteau* ; etc. On ne va donc pas décrire ces métaphores en tant que lexies ; cependant, en tant que valeurs de fonctions lexicales, elles relèvent tout de même de l'étude lexicologique et devront être enchâssées dans la description de la combinatoire des lexies qui en contrôlent l'usage.

Examinons donc maintenant si *comme une casserole d'aluminium* peut être dans notre exemple un collocatif, contrôlé notamment par les adjectifs PROPRE et LIMPIDE. Dit-on normalement en français *propre/limpide comme une casserole d'aluminium* pour exprimer par exemple 'très propre/limpide' ? Clairement non. Ainsi, l'expression *propre comme une casserole d'aluminium* n'a pas du tout le même statut que *propre comme un sou neuf*. Cette dernière expression est, elle, une belle métaphore d'intensification (un **Magn**) contrôlée par la lexie PROPRE.



Lorsque j'utilise *propre comme un sou neuf*, je ne fais que faire appel à ma connaissance lexicale du français, notamment à ma connaissance des propriétés de combinatoire de l'adjectif PROPRE. Il n'en va pas de même si j'énonce *propre comme une casserole d'aluminium*. Cette expression, si elle est comprise par mon interlocuteur, n'est pas **reconnue** par ce dernier : elle est **interprétée**. Nous avons donc démontré que nous étions ici en présence d'une métaphore libre, non lexicalisée.

Nous en avons maintenant terminé avec ce chapitre crucial sur l'analyse du sens linguistique. À ce stade, toutes les notions centrales de l'ouvrage ont été introduites. Il nous reste maintenant à commencer à prendre notre envol et à nous aventurer dans deux domaines qui sont à la périphérie de la sémantique et de la lexicologie « pures » :

- 1 la lexicographie (Chapitre 9), qui est en quelque sorte de la lexicologie appliquée ;
- 2 la pragmatique (Chapitre 10), qui concerne l'interface entre la sémantique et la situation de parole.

Lectures complémentaires

Mel'čuk, Igor A., André Clas et Alain Polguère (1995) Chapitre IV, Section 2.1. Champ sémantique. In : *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve : Duculot, pp. 173-175.

Un texte qui apportera un complément d'information sur la notion de champ sémantique et sur son importance en lexicologie.

Wierzbicka, Anna (1988) L'amour, la colère, la joie, l'ennui. La sémantique des émotions dans une perspective transculturelle. *Langages*, n° 89, pp. 97-107.

Ce court texte permet de comprendre comment l'analyse sémantique lexicale peut être faite dans une perspective plus « sociologique », afin de mettre en évidence certaines différences culturelles qui sont reflétées dans le lexique des langues.

Wierzbicka, Anna (1977) Mental language and semantic primitives. *Communication and Cognition*, Vol. 10, n° 3-4, pp. 155-179.

Ce texte est à lire notamment pour sa critique de l'analyse du sens lexical basée sur l'utilisation de traits sémantiques (l'analyse componentielle).

Aristote *Organon V : Les Topiques*. Paris : Librairie J. Vrin, 384 pages.

Il est impossible de recommander seulement un extrait de cet ouvrage, qui forme un tout cohérent. Aristote expose dans *Les Topiques* sa théorie de la « dialectique » et ce sont bien entendu les chapitres portant spécifiquement sur la notion de définition des concepts qui sont particulièrement intéressants pour nous. C'est à mon avis une lecture incontournable pour toute personne intéressée à comprendre les fondements épistémologiques du recours à la définition en linguistique et, plus généralement, en sciences.

Exercices

- 1 Soit la phrase ci-dessous qui, sans être absolument incohérente sémantiquement, semble un peu bizarre :

(26) [?]#Elle le dévisageait du coin de l'œil.

Utiliser cet exemple pour affiner la définition de DÉVISAGER qui a été proposée en (12) ci-dessus.

- 2 Pousser encore plus loin le travail sur cette définition en veillant maintenant à ce qu'elle prenne en compte non seulement le contraste entre DÉVISAGER et SCRUTER, mais aussi celui existant entre DÉVISAGER et FIXER [*Léo fixait Sylvain de son regard fou.*].
- 3 Corriger les définitions (14ab), page 158, pour en éliminer le cercle vicieux.
- 4 Écrire des définitions analytiques pour 'éplucher I' [*Il a épluché trois kilos de patates.*] et 'éplucher II' [*Je vais bien éplucher ton texte avant d'en faire la critique.*]. Il faudrait que ces définitions mettent en évidence le lien sémantique qui existe entre les deux lexies ÉPLUCHER I et ÉPLUCHER II au sein du vocable ÉPLUCHER.

- 5 Soit la définition suivante de la lexie AVALER [*Jean a avalé une arête.*] : ‘fait de manger quelque chose’. Pourquoi cette définition n’est-elle pas une définition analytique valide ?
- 6 Donner dix lexies du champ sémantique des phénomènes atmosphériques. Justifier la réponse.
- 7 Faire une description de la lexie DORMIR [*Il dormait à poings fermés depuis trois heures.*].
 - au moyen d’une définition lexicographique,
 - puis au moyen d’une analyse componentielle.Quelles conclusions peut-on en tirer ?
- 8 Identifier les différentes lexies du vocable BRÛLER. Quels liens sémantiques les unissent ?
- 9 Dans la phrase (27) ci-dessous, sommes-nous en présence d’une métaphore libre ou lexicalisée ?

(27) *Le beau Frédo dévorait Léontine du regard.*

- 10 Soit la phrase suivante, entendue à la radio :

Je suis allé voir un expert pour savoir si cette pointe de l’iceberg est bien ce qu’elle est, c’est-à-dire la pointe d’un iceberg qui se trouve en dessous d’elle.

La personne qui a produit cette phrase s’est sérieusement enlisée en tentant de décortiquer une locution du français. Identifier avec le plus d’exactitude possible cette locution et élaborer sa définition.



Chapitre 9 La lexicographie

Dictionnaire, lexicographie, dictionnaire commercial vs théorique, dictionnaire de langue (monolingue), dictionnaire encyclopédique, dictionnaire plurilingue/bilingue, dictionnaire d'enseignement/d'apprentissage, macrostructure, superarticle, article, entrée, nomenclature, microstructure, définition lexicographique, dictionnaire de synonymes.

I promise nothing complete; because any human thing supposed to be complete, must for that very reason infallibly be faulty.

Herman Melville, *Moby-Dick* (Chapter 32 on *Cetology*)

Nous allons maintenant examiner les dictionnaires, qui sont en quelque sorte des « produits dérivés » de la lexicologie, comme les grammaires scolaires sont des produits dérivés de l'étude de la grammaire des langues. La présentation des dictionnaires peut faire à elle seule l'objet de tout un ouvrage. Je me contente de donner ici des informations permettant de relier les dictionnaires aux notions de base de lexicologie qui ont été présentées dans les chapitres précédents.

Les dictionnaires sont, pour la plupart, des ouvrages écrits par des spécialistes de la langue pour un public de non spécialistes. De plus, les dictionnaires sont perçus par le grand public comme des **symboles** de la langue elle-même. S'acheter un dictionnaire revient en quelque sorte à s'approprier la connaissance véritable de ce qu'est la langue. La foi dans l'infailibilité des dictionnaires est totale ; on perd de vue le fait que ces ouvrages sont élaborés par des êtres humains, qui mangent, boivent, dorment, sont fatigués, font des erreurs, font des omissions, ignorent

certaines données, etc. Le mythe de l'exhaustivité des dictionnaires a la vie dure. Les gens vont admettre que « tous les mots de la langue » ne sont peut-être pas décrits dans leur dictionnaire favori, mais ils croiront dur comme fer que la description d'un « mot » donné — ce que nous appelons un vocable — est, elle, complète et juste.

C'est pourquoi j'ai débuté ce chapitre par cette belle citation de H. Melville : en tant que produits de l'activité humaine, les dictionnaires sont tous d'une façon ou d'une autre incomplets et erronés. Cela n'enlève d'ailleurs rien au fait qu'ils peuvent être d'irremplaçables outils de travail. Ce chapitre vise autant à connecter la lexicologie à la lexicographie qu'à démystifier le concept de dictionnaire.

Dictionnaires et lexicographie

Qu'est-ce qu'un dictionnaire ?

Un **dictionnaire** d'une langue donnée est un répertoire du lexique de cette langue qui fournit, pour chaque lexie, une description selon un patron relativement rigide (définition, étymologie, prononciation, exemples d'emploi, etc.).

La **lexicographie** est l'activité ou le domaine d'étude visant la construction de dictionnaires.

Le terme *dictionnaire* est habituellement utilisé pour désigner les **dictionnaires commerciaux** comme le *Petit Robert* ou le *Larousse illustré*, qui sont avant tout des produits destinés à la vente. Rédiger de tels dictionnaires revient en quelque sorte à faire de la « lexicologie appliquée ».

Il existe cependant aussi des **dictionnaires théoriques**, c'est-à-dire des dictionnaires conçus comme des outils de recherche en linguistique, que l'on développe en vue d'étudier le lexique des langues. Les dictionnaires théoriques peuvent cependant aussi être utilisés comme modèles expérimentaux pour améliorer la qualité (complétude, cohérence, etc.) des dictionnaires commerciaux. Voici deux exemples de dictionnaires théoriques :

Mel'čuk, Igor A. *et al.* (1984, 1988, 1992, 1999) *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexicosémantiques* (Volumes I à IV). Montréal : Les Presses de l'Université de Montréal.

Wierzbicka, Anna (1987) *English Speech Act Verbs: A Semantic Dictionary*. Sydney : Academic Press.

Dans ce chapitre, nous allons nous concentrer sur le cas des dictionnaires commerciaux, les dictionnaires théoriques ne pouvant être étudiés que dans le cadre d'un cours plus avancé de lexicologie ou de sémantique.

Les dictionnaires commerciaux ne sont pas des livres comme les autres, cela pour au moins deux raisons :

- 1 Les dictionnaires ont une importance sociale considérable. En effet, en décrivant le lexique d'une langue, le dictionnaire se présente en même temps comme un reflet de la société dans laquelle cette langue est parlée. On peut donc affirmer l'existence même d'une société, d'une culture, en entreprenant la rédaction d'un dictionnaire de sa langue. Les états sont souvent impliqués dans la rédaction de dictionnaires (*Dictionnaire de l'Académie française*, etc.), qu'ils peuvent encourager, financer ou même, pourquoi pas, décourager pour des raisons politiques. Cela est particulièrement évident dans le cas de la rédaction de dictionnaires décrivant une variante dialectale d'une langue. Par exemple, un dictionnaire comme *The Macquarie Dictionary* (rédigé et publié à l'Université Macquarie de Sydney)¹ ne fait pas que décrire une variante de l'anglais : il est aussi l'affirmation de l'existence d'une culture proprement australienne. On pourrait aussi mentionner ici le cas du projet de rédaction du *Dictionnaire canadien bilingue*, financé depuis plusieurs années par le gouvernement fédéral du Canada (publication prévue pour 2004).
- 2 Tout le monde ou presque possède au moins un dictionnaire, même si c'est souvent un ouvrage plutôt aride, destiné à la consultation, et même si nombreuses sont les personnes qui n'ouvrent jamais le dictionnaire qu'elles ont à la maison ou au bureau. Un dictionnaire est un peu comme un annuaire téléphonique : un « gros livre » que l'on se doit de posséder

1. Delbridge, A. *et al.* (1997) *The Macquarie dictionary*. 3^e éd., Macquarie University : Macquarie Library, 2504 pages.

au cas où. Les gens « ordinaires » entretiennent donc une relation un peu particulière avec leur dictionnaire. Parce qu'ils en possèdent tous un, parce que le dictionnaire est généralement présenté comme un outil pour le grand public, ils pensent bien connaître et comprendre l'information qu'il contient. Or, rares sont ceux qui ont véritablement pris le temps d'étudier quelle information contient leur dictionnaire et comment elle y est présentée. On ouvre bien souvent un dictionnaire uniquement pour vérifier l'orthographe d'un mot ou, au mieux, pour trouver la signification d'un mot rare ou technique que l'on a rencontré.

Types de dictionnaires

Il existe une grande variété de dictionnaires, selon le type de public et d'utilisation visé ; par exemple :

- 1 Les **dictionnaires de langue (monolingues)** présentent les lexies de la langue dans leur réalité linguistique : prononciation, partie du discours, sens, etc. Ils ne contiennent généralement pas de noms propres. Ils se distinguent des **dictionnaires encyclopédiques**, qui contiennent notamment de nombreux noms propres (noms de pays, de personnalités, etc.) et, surtout, donnent pour chaque unité décrite des informations non linguistiques sur les entités correspondantes. Ainsi, un dictionnaire encyclopédique ne va pas décrire la **lexie** VACHE, mais plutôt l'**animal** lui-même : ce que mange une vache, son poids moyen, la façon dont fonctionne son système digestif, etc.
Certains dictionnaires de langue peuvent se focaliser sur un aspect particulier de la description lexicale. Par exemple, le *BBI*² se concentre sur la description des collocations contrôlées par les lexies anglaises, information qui n'est généralement pas explicitement présentée dans les dictionnaires de langue traditionnels.
- 2 Les dictionnaires monolingues font une description qui fonctionne à l'intérieur même de la langue décrite, alors que les **dictionnaires plurilingues**, généralement **bilingues**, décrivent les lexies d'une langue soit par leur équivalent lexical dans une autre langue, voir (1a), soit par une véritable définition formulée dans cette autre langue, voir (1b) :

2. Benson, M., Benson, E. and R. Ilson (1997) *The BBI Dictionary of English Word Combinations*. Éd. révisée, Amsterdam/Philadelphie : Benjamins, 386 pages.

- (1) a. **PAIN** *bread*
b. **TARTINE** *slice of bread spread with something like butter or jam*

3 Les **dictionnaires d'enseignement**, ou **d'apprentissage**, sont conçus pour être utilisés par des personnes qui apprennent activement la langue. Ils sont plus ou moins riches (de quelques dizaines de vocables à plusieurs milliers) selon le niveau d'enseignement visé. Les dictionnaires destinés à un très jeune public contiennent souvent beaucoup plus d'illustrations que les dictionnaires de langue courants³. De plus, les dictionnaires d'enseignement peuvent avoir des structures très originales, regroupant par exemple les lexies par champs sémantiques plutôt que par vocables listés selon l'ordre alphabétique (cf., pour l'anglais, le *Language Activator*, un ouvrage remarquable publié par Longman Group⁴). Une telle organisation encourage un usage « actif » du dictionnaire : on s'en sert avant tout pour trouver les moyens d'encoder sa pensée (et non pour simplement trouver le sens d'un terme inconnu). Mon but n'étant pas d'entamer ici une véritable présentation de la lexicographie, mais plutôt de la présenter sous l'angle de ses connexions avec les notions de base de lexicologie, je me concentrerai dans ce qui suit sur le cas le plus simple : les dictionnaires de langue monolingues.

Macrostructure et microstructure des dictionnaires

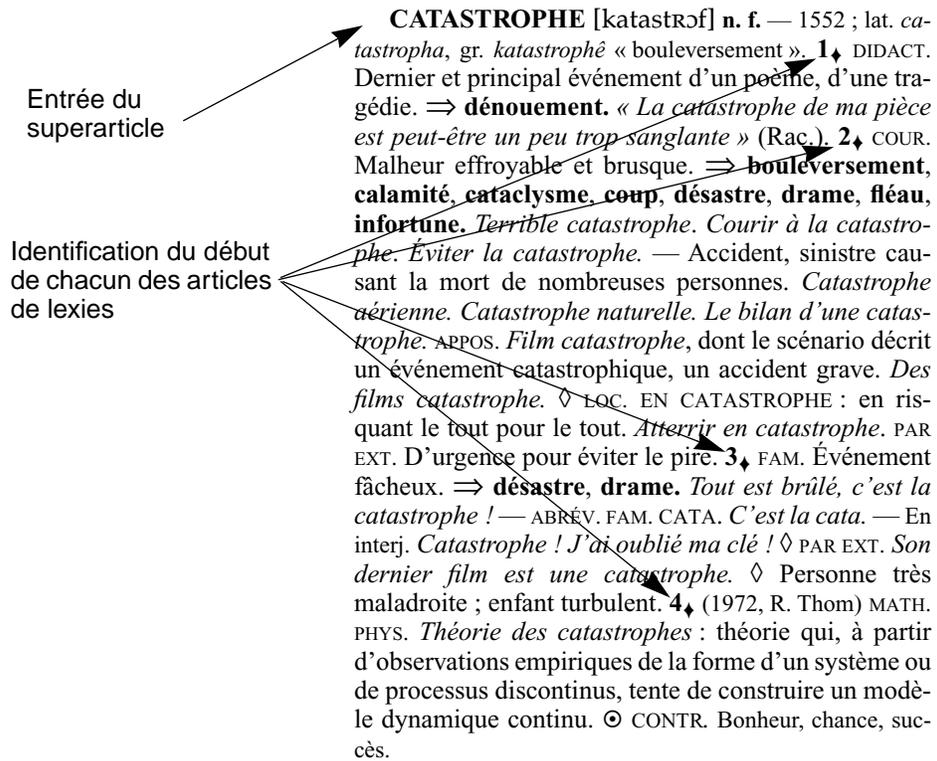
Les dictionnaires de langue courants possèdent une organisation interne caractéristique.

La **macrostructure** d'un dictionnaire — son ossature générale — est organisée autour d'une succession de **superarticles**, ordonnés alphabétiquement, dont chacun contient la présentation d'un vocable. Chaque lexie du vocable est présentée dans un **article** du superarticle. Le nom du vocable, qui identifie le superarticle, est généralement appelé **entrée** de dictionnaire. L'ensemble des vocables ayant une entrée dans le dictionnaire est la **nomenclature** du dictionnaire.

3. Certains dictionnaires de langue, comme le *Petit Robert*, ne contiennent **aucune** illustration.

4. Summers, D. (1993) *Longman language activator*. Harlow : Longman, 1 587 pages.

Voici, comme illustration, le superarticle du vocable CATASTROPHE, extrait du *Nouveau Petit Robert* (1993) :



Comme on le voit ici, chaque lexie (acception) d'un vocable est généralement numérotée à l'intérieur du superarticle. L'organisation interne des superarticles et des articles, en fonction du contenu de chaque article (étymologie, prononciation, définition, exemples, etc.) est appelée la **microstructure** du dictionnaire.

Il faut bien comprendre qu'il n'existe pas une façon unique de décrire les vocables de la langue dans un dictionnaire. Les dictionnaires diffèrent énormément non seulement quant au format de présentation qu'ils utilisent, mais aussi pour ce qui est de l'information même qu'ils

procurent sur les vocables. Comparons, par exemple, le superarticle ci-dessus avec celui donné, pour le même vocable, dans le *Lexis* (1979)⁵ :

CATASTROPHE [katastrɔf] n. f. (lat. *catastrophā*, gr. *katastrophē*, bouleversement; 1546). Événement subit qui cause un bouleversement, des destructions, des victimes : *Un avion s'est écrasé au sol; c'est la troisième catastrophe de ce genre en un mois* (syn. ↓ACCIDENT). *Le « captain » Lyttelton ne désirait manifestement pas commencer sa mission par une catastrophe* (de Gaulle). *L'incendie prend les proportions d'une catastrophe* (syn. CALAMITÉ, DÉASTRE). *Son échec à cet examen est pour lui une vraie catastrophe* (syn. MALHEUR). ● LOC. ADV. *En catastrophe*, se dit d'un avion qui atterrit dans des conditions particulièrement difficiles pour éviter un accident : *Le pilote a atterri en catastrophe* ; se dit d'une action qui est faite d'urgence et au dernier moment : *Un mariage en catastrophe*. ♦ **catastrophique** (...) ♦ **catastropher** (...) ♦ **catastrophisme** (...).

● CLASS. et LITT. **catastrophe** n. f. **1.** Dénouement d'une tragédie, d'un récit : *Qu'on se figure une salle de spectacle vide, après la catastrophe d'une tragédie* (Chateaubriand). — **2.** Issue malheureuse d'un événement : *La catastrophe de ce fracas fut la perte de deux chevaux* (La Fontaine).

Vous trouverez à la fin de ce chapitre plusieurs exercices portant sur l'analyse du contenu de ces superarticles.

Les définitions lexicographiques

Nous avons étudié en détail dans le chapitre précédent comment décrire les sens lexicaux au moyen de définitions, notamment de définitions analytiques. Le plat de résistance, dans un article de dictionnaire, est bien évidemment la définition, dite **définition lexicographique**. En théorie, une « bonne » définition lexicographique ne peut qu'être, en même temps, une définition analytique. Pourtant, il peut arriver que les dictionnaires utilisent, de façon ponctuelle ou quasiment systématique,

5. *Larousse de la langue française : lexis* (1979) Paris : Librairie Larousse, 2 109 pages.

Pour gagner de la place, j'ai enlevé de ce superarticle la description des vocables enchâssés CATASTROPHIQUE, CATASTROPHER et CATASTROPHISME.

une description sémantique des lexies basée sur une liste de synonymes approximatifs. Contrastons (2a) et (2b) ci-dessous ; de par leur structure, ce sont des équivalents monolingues des exemples (1a) et (1b) donnés page 179 à propos des dictionnaires bilingues⁶ :

- (2) a. **RÉPULSION** *aversion, dégoût, haine, répugnance*
- b. **RÉPULSION** *répugnance physique ou morale à l'égard d'une chose ou d'un être qu'on repousse*



Une liste d'un ou plusieurs synonymes peut suffire, dans certains contextes d'utilisation, comme moyen de cerner approximativement le sémantisme d'une lexie. Mais une telle liste ne devrait pas être appelée *définition*, dans la mesure où, justement, elle ne définit pas. L'usage de ce type de liste est tout à fait justifié dans le cas de dictionnaires bilingues — dans lesquels on cherche généralement une **traduction** plutôt qu'une définition, mais elle reflète nécessairement une faiblesse dans le cas des dictionnaires de langue monolingues, qui ne sont pas supposés être des **dictionnaires de synonymes**. Si le dictionnaire de langue doit vraiment être utilisé comme un outil pédagogique ou d'apprentissage, il faut qu'il ait recours aux définitions analytiques.

La seule contrainte qui s'applique aux définitions analytiques et qui peut parfois être assouplie dans le cas des dictionnaires commerciaux est celle voulant que la définition soit une paraphrase de la lexie définie. Pour des raisons pédagogiques, certains dictionnaires vont préférer utiliser des définitions aux allures plus « digestes ». Examinons, par exemple, la définition donnée pour le verbe anglais INCITE1 [*The media incited them to strike.*] dans le *Collins Cobuild English Language Dictionary* (1987)⁷ :

- (3) **INCITE1** *If you incite someone to do something, you encourage them to do it by making them excited or angry.*

Ce dictionnaire, qui se veut avant tout un outil pédagogique, a pris le parti de « parler » directement à son usager plutôt que d'utiliser une définition paraphrastique créant plus de distanciation.

6. La définition (2b) est empruntée au *Nouveau Petit Robert*.

7. *Collins COBUILD* (1987) London : Collins, 1 703 pages.

Le texte de J. Picoche donné en lecture ci-dessous examine en détail la problématique de la définition lexicographique.

Ceci termine ce bref aperçu du domaine de la lexicographie, que j'ai avant tout cherché à faire percevoir comme la branche appliquée de la lexicologie. Il me reste maintenant à présenter brièvement, dans le dernier chapitre, une autre discipline directement reliée au contenu de ce cours : l'étude des phénomènes dits « pragmatiques ».

Lectures complémentaires

Picoche, Jacqueline (1977) Chapitre V. La définition. In : *Précis de lexicologie française*. Paris : Nathan, pp. 133-148.

Ce texte complète ce qui a été dit sur la notion de définition dans le présent chapitre et dans le Chapitre 8. Noter que les définitions analytiques sont appelées ici *définitions substantielles*.

Rey-Debove, Josette et Alain Rey (1993) Introduction pour *Le Nouveau Petit Robert 1*, Paris : Le Robert, pp. IX-XIX.

On ne lit jamais les introductions des dictionnaires et c'est un tort...

Wierzbicka, Anna (1996) Chapter 9, Semantics and lexicography. In : *Semantics: Primes and Universals*, Oxford/New York : Oxford University Press, pp. 258-286.

Un texte de réflexion qui explicite les rapports entre la sémantique/lexicologie et la lexicographie, notamment celle visant la construction de dictionnaires théoriques.

Exercices

- 1 Comparer méthodiquement les deux superarticles de dictionnaire donnés dans ce chapitre (contenu et organisation de ce contenu).
- 2 En se basant sur la description du *Lexis*, déterminer quelle est la nature sémantique — prédicat ou objet sémantique — des deux acceptions marquées CLASS. et LITT. de CATASTROPHE (... *après la catastrophe*

d'une tragédie et La catastrophe de ce fracas...). Si ce sont des prédicats, combien ont-il d'arguments ? Il faut être capable de justifier les réponses.

- 3 Le *Petit Robert*, édition de 1981, définit l'adjectif ROUGE1.1 [*une rose rouge*] de la façon suivante :

(4) Qui est de la couleur du sang, du coquelicot, du rubis, etc.
(extrémité du spectre solaire).

Que peut-on dire de cette définition ?

- 4 Le *Nouveau Petit Robert*, édition de 1993, procède d'une autre façon :

(5) I_♦ Adj. 1_♦ Qui est de la couleur du sang, du coquelicot, du rubis, etc.
(cf. ci-dessous II, *le rouge*) (...)

II_♦ N. m. (XII^e) LE ROUGE. 1_♦ La couleur rouge. *Le vert est la couleur complémentaire du rouge. Le rouge, extrémité du spectre visible*
(...)

Comment expliquer ce changement ?

- 5 Toujours dans le *Nouveau Petit Robert*, SANG1 est défini de la façon suivante :

(6) Liquide visqueux, de couleur rouge, qui circule dans les vaisseaux, à travers tout l'organisme, où il joue des rôles essentiels et multiples (nutritif, respiratoire, régulateur, de défense, etc.).

Faire un examen détaillé du contenu de cette définition. Trouver le cercle vicieux. Peut-il être évité ?

- 6 Indiquer, en se basant sur le superarticle du *Nouveau Petit Robert* donné page 180, quelle lexie est sémantiquement plus simple : CATASTROPHE2 ou MALHEUR ? Combien de lexies CATASTROPHE ce superarticle décrit-il ?



Chapitre 10

La pragmatique

Sémantique d'une langue vs sémantique linguistique, pragmatique, message linguistique, référent, (théorie des) acte(s) de parole, acte locutoire vs illocutoire vs perlocutoire, énoncé et verbe performatifs, lexie à valeur pragmatique, fonctions du langage.

*Sa vie à elle était loin d'être belle
Mademoiselle madame veuve et mademoiselle
Voyez ce que je veux dire voyez peut-être pas
Ce que je veux dire je ne le dirai pas
Dick Annegarn, L'institutrice*

Ce dernier chapitre fait un survol des phénomènes liés à l'interaction entre la langue proprement dite et la situation dans laquelle un énoncé est produit. Ces phénomènes relèvent de la parole au sens saussurien. Nous allons donc sortir du champ d'étude étroit que j'ai assigné au début de ce cours à la lexicologie et à la sémantique.

Je vais tout d'abord définir la notion de pragmatique, puis je donnerai quelques précisions sur la nature des échanges linguistiques : que se passe-t-il lorsque deux individus communiquent au moyen de la langue ? J'introduirai ensuite une approche théorique qui vise la modélisation de certains phénomènes pragmatiques : la théorie des actes de parole. Je conclurai par l'examen de certaines lexies qui entretiennent une relation privilégiée avec le contexte d'énonciation dans lequel elles doivent être employées.

La pragmatique

La pragmatique peut se définir de la façon suivante :

La **pragmatique** est l'ensemble des phénomènes, dits *phénomènes pragmatiques*, qui mettent en relation la langue avec le contexte dans lequel les énoncés sont produits (c'est-à-dire, le contexte d'énonciation).

La prise en compte de la pragmatique permet notamment de comprendre les contraintes contextuelles s'appliquant à l'utilisation de la langue, qui font que cette dernière fonctionne ou non efficacement en tant que système sémiotique. Par exemple, il faut sortir de l'analyse strictement lexicale et grammaticale de la langue pour pouvoir expliquer pourquoi la réponse donnée dans le dialogue ci-dessous ne sera pas jugée adéquate dans de nombreux contextes d'énonciations :

- (1) — *Qu'est-ce que tu avais comme animal domestique quand tu étais petite ?*
— *Un mammifère.*

Bien entendu, l'analyse lexicologique et sémantique va être fort importante pour expliquer le dysfonctionnement du dialogue (1) — notamment, si l'on utilise les concepts d'hyponyme et de synonyme. Mais elle n'est pas suffisante¹.

Avant de poursuivre sur la pragmatique, revenons brièvement à la notion de sémantique, à propos de laquelle il convient de faire une mise au point terminologique.

1. Cela était la principale raison avancée par Saussure pour exclure l'étude de la parole du champ de la linguistique : rien de solide sur la parole ne peut être dit sans avoir préalablement fait une bonne analyse structurale des langues. On peut cependant considérer que l'exclusion de la parole du champ d'étude de la linguistique était justifiée au début du XX^e siècle, à une époque où la linguistique restait à construire en tant que science. Cette position peut fort légitimement être remise en question de nos jours.

Comme je l'ai mentionné en début d'ouvrage (Chapitre 1, page 10) le terme *sémantique* est ambigu ; on peut distinguer deux sémantiques :

- 1 La ***sémantique d'une langue*** donnée (sémantique du français, de l'anglais, etc.) est *grosso modo* l'ensemble des sens exprimables dans cette langue ainsi que l'ensemble des règles d'expression et de combinaison de ces sens. C'est une des composantes structurales du système sémiotique qu'est une langue.
- 2 La ***sémantique linguistique*** est l'étude scientifique de la sémantique des langues.

Ainsi, dans cet ouvrage, j'ai introduit ce qu'est la sémantique (au sens **2**) linguistique, en empruntant la plupart de mes exemples à la sémantique (au sens **1**) du français.

Cette mise au point étant faite, je peux maintenant entrer dans le vif du sujet et présenter la pragmatique.

On oppose souvent la sémantique linguistique (sémantique au sens **2**) et la pragmatique. Mais, dans ce cas, on emploie le terme *pragmatique* pour désigner une **discipline** qui s'attacherait à étudier les phénomènes pragmatiques. Je ne suis pas convaincu qu'une telle science existe, ni qu'elle puisse exister. En effet, dès que l'on sort de la description de la structure des langues proprement dite, on entre dans des champs d'étude mettant en jeu des phénomènes pragmatiques. Ainsi, la linguistique appliquée à l'enseignement ou à l'apprentissage, la sociolinguistique, etc. sont toutes des branches « légitimes » de la linguistique qui impliquent la prise en compte de phénomènes pragmatiques. Si on cherche à isoler la pragmatique comme une discipline à part, on va donc se retrouver avec un domaine d'étude vague et hybride, manquant de l'homogénéité requise pour définir une branche spécifique de la linguistique. C'est la raison pour laquelle je n'emploie le terme *pragmatique* que pour désigner un ensemble de phénomènes, et non un domaine d'étude.

Nature des échanges linguistiques

Le terme *échange linguistique* ou des expressions comme *échanger des propos* illustrent le fait que l'on conceptualise souvent la communi-

cation linguistique comme une sorte de transaction : un échange de **message** entre le locuteur et le destinataire.

De façon caricaturale, la communication linguistique pourrait se ramener à un processus subdivisé en au moins quatre étapes majeures :

- 1 construction et encodage linguistique du message par le locuteur ;
- 2 émission physique de ce message par le locuteur ;
- 3 réception du message physique par le destinataire ;
- 4 décodage du message par le destinataire.

Mais cette façon de modéliser la communication linguistique est clairement insuffisante : elle postule que la communication est réussie si le message décodé par le destinataire est le même que celui qui a été encodé par le locuteur. Or, la communication ne fonctionne pas de façon aussi simple.

Les principaux phénomènes dont rend difficilement compte une telle approche sont au moins au nombre de six :

- 1 Les phrases sont en général ambiguës. Pour que le message soit transmis, il faut que l'interlocuteur désambiguïse la phrase, ou à tout le moins qu'il soit à même d'identifier en priorité le sens approprié. Il y a donc des principes qui gouvernent la désambiguïsation et la compréhension. Ces principes ne sont cependant pas pris en compte dans les étapes mentionnées ci-dessus puisque l'on a envisagé la compréhension comme un simple décodage, alors qu'il s'agit plutôt d'une interprétation.
- 2 Une partie du message qu'on veut transmettre peut ne pas être encodée explicitement dans le message communiqué. Par exemple, l'identification du **réfèrent** exact (cf. Chapitre 6). Si je dis *Le premier homme à avoir marché sur la lune*, mon interlocuteur va peut-être comprendre la même chose que si j'avais dit *Neil Armstrong*, mais ce n'est pas ce que j'énonce de façon explicite. Après tout, ma phrase pourrait tout aussi bien faire référence à Tintin (cf. *Objectif lune* de Hergé).
- 3 Les intentions du locuteur ne sont pas nécessairement encodées directement dans le message. Seule la mise en contexte va permettre de les expliciter. Par exemple, si un professeur dit à ses étudiants :

(2) — *N'oubliez pas d'aller aux séances de monitorat !*

cela peut être pour leur rappeler un fait, pour les menacer (*Si vous n'y allez pas, ne vous plaignez pas d'échouer à l'examen*), etc. Or, il est essentiel, pour que l'échange d'information ait lieu, que le destinataire identifie précisément si l'énoncé s'interprète comme une simple transmission de données, une promesse, une menace, etc. Il y a donc des éléments informationnels mis en jeu dans la communication qui n'appartiennent pas véritablement au message **linguistique** transmis. Il faudrait au moins considérer une extension de ce message à ce qui n'est pas encodé au moyen de la langue pour modéliser comment fonctionne l'échange linguistique.

- 4 Il y a des cas où le message que l'on veut transmettre ne correspond pas du tout au sens littéral de l'énoncé produit. Il peut même être en totale opposition avec ce sens littéral. Par exemple :

(3) — *Ah ! Bravo !*

quand quelqu'un vient de renverser son café sur la table. En fait, on veut manifester notre mécontentement, mais on énonce une phrase qui **littéralement** exprime notre admiration pour ce qui a été fait.

- 5 On peut aussi exprimer plus que ce que l'on dit. L'exemple classique habituellement cité est celui de quelqu'un qui entre dans une pièce et dit :

(4) — *Il fait froid ici !*

pour en fait demander à la personne qui se trouve dans la pièce de fermer les fenêtres. Un exemple plus subtil et plus intéressant est l'extrait de chanson donné en exergue à ce chapitre, page 185.

- 6 Une énonciation peut avoir un autre but premier que la communication. Elle peut être l'accomplissement d'un certain acte social, selon les normes d'un rituel :

(5) a. — *Je vous déclare mari et femme.*

b. — *La séance est ouverte.*

Ou encore, on peut dire quelque chose comme :

(6) — *J'ai lu tout Proust.*

non pour communiquer le message exprimé mais, par exemple, pour impressionner son interlocuteur.

La communication linguistique est donc bien plus qu'un simple échange de messages, qu'une simple transaction. D'une façon ou d'une autre, pour rendre compte complètement de cette forme de communication, il faudra tenir compte des systèmes de conventions sociales et de croyances partagées par le locuteur et le destinataire.

La théorie des actes de parole

Nous n'allons pas entrer dans une étude poussée des différentes théories qui ont été proposées pour modéliser les phénomènes pragmatiques intervenant dans la communication linguistique. Nous nous contenterons d'examiner brièvement une des plus connues, qui présente l'avantage d'étendre ses ramifications en dehors de la seule étude des phénomènes pragmatiques, puisqu'elle concerne aussi la sémantique lexicale.

La **théorie des actes de parole** (angl. *Speech Act Theory*), proposée dans les années soixante par le philosophe du langage John Austin, postule un modèle de la communication qui va permettre de rendre compte de certains des phénomènes pragmatiques mentionnés ci-dessus.

Selon Austin, en énonçant une phrase, on accomplit trois actes de parole simultanés :

- 1 un **acte locutoire** — la production d'un énoncé linguistique, le fait d'énoncer ;
- 2 un **acte illocutoire** — la communication linguistique elle-même, le fait de transmettre une information donnée au moyen de la langue ;
- 3 un **acte perlocutoire** — qui concerne les fins plus lointaines de l'énonciation, qui peuvent échapper au destinataire même si celui-ci maîtrise et comprend parfaitement la langue ; cet acte comprend l'acte illocutoire plus ses effets.

Ainsi, lorsque Chimène lance à Don Rodrigue sa fameuse litote² ;

(7) *Va, je ne te hais point.*

elle accomplit l'acte d'énoncer, d'émettre cette phrase (acte locutoire). Elle transmet aussi un message à Don Rodrigue (acte illocutoire), message qu'on pourrait aussi exprimer avec plus de mots et moins d'élégance par quelque chose comme :

(8) *C'est OK maintenant. Tu peux aller faire ton devoir. Et puis, rappelle-toi, je ne t'en veux pas. En fait, je suis même franchement amoureuse de toi.*

Cet acte illocutoire va avoir des conséquences (acte perlocutoire). Il s'agit bien entendu du fait que Rodrigue est maintenant informé de tout ceci, mais aussi, peut-être, qu'il va se sentir plein de courage et de détermination pour accomplir ce qui doit être accompli (effet voulu par Chimène). Mais celui-ci peut tout aussi bien répondre dans un premier temps (comme il le fait dans la pièce) par un fataliste :

(9) *Tu le dois.*

On peut noter que les actes perlocutoires sont « moins linguistiques » que les actes illocutoires. Ces derniers ont en effet des liens directs avec les formes linguistiques correspondantes — au niveau du lexique (DEMANDER, PROMETTRE, etc.) et de la grammaire (phrase interrogative, exclamative, etc.). La modélisation de l'accomplissement d'actes perlocutoires, quant à elle, relève surtout d'une modélisation de la réalité extralinguistique ainsi que de certains processus cognitifs qui ne sont pas nécessairement liés au langage.

En relation avec la notion d'acte illocutoire, Austin a aussi proposé celle d'énoncé performatif :

Un **énoncé performatif** est tel que son énonciation constitue en même temps l'accomplissement de l'acte dénoté (c'est-à-dire, signifié) par l'énoncé lui-même.

On peut distinguer plusieurs types d'énoncés performatifs :

2. Corneille, *Le Cid*, Acte III, Scène IV.

- Performatifs explicites

- (10) a. — *Je vous déclare mari et femme.*
b. — *Je déclare la séance ouverte.*
c. — *Je vous ordonne de partir.*

- Performatifs implicites

- (11) — *Échec et mat !*

Ce qui est intéressant pour la sémantique et l'étude lexicale, c'est qu'il existe dans chaque langue des verbes, dits **verbes performatifs**, qui peuvent contrôler la construction d'énoncés performatifs. L'emploi performatif d'un verbe se fait nécessairement à la première personne du singulier du présent de l'indicatif³. Comparons :

- (12) a. — *Je te demande/te supplie/t'ordonne de partir.*
b. — *Je t'interroge sur ton âge.*

La phrase (12b) ne peut qu'être utilisée pour constater un fait, et non pour interroger. Elle peut par exemple apparaître dans le contexte suivant :

- (13) *Je t'interroge sur ton âge, et toi tu te fâches !*

La phrase (12a), cependant, est utilisée très naturellement pour exprimer une requête : le locuteur demande à son interlocuteur de partir.

En conséquence, on peut dire que DEMANDER, SUPPLIER et ORDONNER sont des verbes performatifs, **dans les acceptions considérées ci-dessus**. Le verbe INTERROGER, par contre, n'est pas un verbe performatif. On notera, à ce propos, que l'acception de DEMANDER qui est un synonyme approximatif d'INTERROGER n'est pas non plus performative. Un francophone ne demandera pas à quelqu'un son âge en disant (14a), mais plutôt au moyen d'une question comme (14b) :

3. Bien entendu, cette définition vaut ici pour le français. Pour d'autres langues (notamment, pour les langues sans temps grammatical), il faut considérer l'emploi de verbes dans des constructions qui sont équivalentes à cette construction française.

Le caractère arbitraire des lexies à valeur pragmatique apparaît encore mieux quand on cherche à les traduire ; seul (17a) est approprié si l'on veut un panneau anglais vraiment standard :

- (17) a. *No parking*
b. *Parking forbidden*

Prenons un autre exemple de panneau. Comment signaler qu'un mur a été repeint et qu'il faut faire attention de ne pas en toucher la peinture ? Bien entendu, c'est (18a) qui est le panneau véritablement « idiomatique ». Le panneau (18b) est certainement moins courant et (18c) est carrément bizarre :

- (18) a. *Peinture fraîche*
b. *Attention à la peinture*
c. *Vient d'être peint*

Finalement, pour sortir du cas des panneaux, prenons un exemple de lexie que nous employons continuellement dans la vie de tous les jours. Soit le contexte pragmatique suivant :

La personne **A** adresse la parole à **B** pour la première fois de la journée. (Elle peut d'ailleurs très bien ne pas lui avoir parlé depuis beaucoup plus longtemps que ça.) **A** cherche à établir le contact de façon amicale mais, tout de même, conventionnelle — c'est-à-dire que **A** ne cherche pas à faire quelque chose qui soit marqué ou original.

En fonction de toutes ces contraintes pragmatiques, **A** dira peut-être :

- (19) *Comment vas-tu ?*

La lexie COMMENT VAS-TU/ALLEZ-VOUS ? est une lexie à valeur pragmatique car on ne peut pas se contenter d'en faire une définition analytique pour expliciter sa valeur linguistique. Il faut aussi lui associer un contexte d'énonciation particulier, ce que tout bon dictionnaire fera.

Ici s'achève cette présentation des phénomènes pragmatiques. C'est aussi la fin du cours. Je propose, après les lectures et les exercices pour ce chapitre, une brève conclusion indiquant notamment comment le lecteur pourra poursuivre sa formation en lexicologie et sémantique.

Lectures complémentaires

Reboul, Anne et Jacques Moeschler (1998) *La pragmatique aujourd'hui*. Collection « Points Essais », n° 371, Paris : Éditions du Seuil, 209 pages.

Je recommande fortement la lecture de ce petit ouvrage à toute personne que mon introduction à la pragmatique aura laissée sur sa faim. Il a le mérite de proposer une perspective cohérente sur la pragmatique, plutôt que d'être un simple répertoire des travaux effectués dans le domaine (comme c'est malheureusement trop souvent le cas pour les textes d'introduction).

Ducrot, Oswald et Jean-Marie Schaeffer (1995) Langage et action. In : *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris : Éditions du Seuil, pp. 776-783.

À lire notamment pour une présentation de la théorie des actes de parole de Austin. Ce texte contient aussi une brève introduction aux **fonctions du langage**, identifiées par Roman Jakobson. Pour ne pas allonger indûment ce chapitre, je ne présente pas ici ces fonctions. Le lecteur intéressé à en savoir plus pourra consulter directement le texte très clair de Jakobson ci-dessous.

Jakobson, Roman (1973) Linguistique et poétique. In : *Essais de linguistique générale*, Vol. II : Rapports internes et externes du langage. Paris : Éditions de minuit, pp. 209-248.

Exercices

- 1 Monsieur X rencontre Monsieur Y dans la rue. Cela fait trois mois qu'il ne l'a pas vu, alors qu'auparavant ils se croisaient tous les jours. X dit à Y :

(20) — *Est-ce que vous avez été malade ?*

Décrire ce que fait ici Monsieur X, en terme d'actes locutoire, illocutoire et perlocutoire.

- 2 Parmi les verbes suivants, lesquels sont performatifs ?

- PRIER [*Jules prie Émile de se taire.*]
- SUPPLIER [*Jules supplie Émile de partir.*]
- VOULOIR [*Jules veut qu'Émile sorte.*]

3 Lequel des deux verbes RECONNAÎTRE¹ ou RECONNAÎTRE² est performatif ? Justifier.

(21) a. *Je reconnais¹ l'autorité de la Cour.*

b. *T'as tellement changé que je ne te reconnais² plus.*

4 Les phrases suivantes semblent être des paraphrases :

(22) a. *Ouvert 24 heures sur 24.*

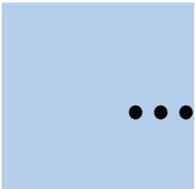
b. *Jamais fermé.*

Pourtant, ce n'est que (22a) que l'on voit normalement dans la vitrine des magasins. Expliquer.

5 Le panneau ci-dessous a été aperçu dans la vitrine d'un magasin de chaussures annonçant des soldes⁴. Analyser son contenu.

(23) **TOUT, TOUT JUSQU'À 50% ET PLUS**

4. Véridique !



Conclusion

Nous sommes maintenant arrivés au terme de cet ouvrage. Son étude aura permis au lecteur d'acquérir un ensemble assez important de notions de base en lexicologie et sémantique, de même qu'un bon éventail de notions connectant ces deux disciplines au reste de l'étude linguistique.

J'ai cherché à assurer au mieux la cohérence de l'ensemble des notions introduites tout en évitant de me situer à l'intérieur du cadre étroit d'une approche théorique linguistique donnée. Ce cours est cependant conçu comme une première étape avant de pouvoir se lancer dans un véritable travail de modélisation de la langue, travail qui doit toujours s'appuyer sur une ou plusieurs approches théoriques bien définies.

L'approche théorique qui est le plus directement compatible avec les bases notionnelles introduites ici est la *théorie Sens-Texte*, déjà mentionnée plusieurs fois dans cet ouvrage (voir l'Index des notions). Le texte suivant offre une brève introduction aux principes centraux de la théorie Sens-Texte :

Polguère, A. (1998) La théorie Sens-Texte. *Dialangue*, Vol. 8-9, Université du Québec à Chicoutimi, pp. 9-30.

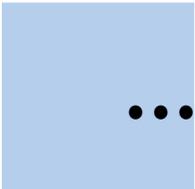
En lisant ce texte, on peut notamment voir comment certaines notions qui ont été introduites dans cet ouvrage — réseau sémantique, fonction lexicale, etc. — peuvent trouver leur place au sein d'une modélisation complète de la langue (grammaire et lexique). Cet article est peut-être difficile à se procurer sous sa forme publiée, mais il est accessible à l'adresse Internet suivante :

<http://www.fas.umontreal.ca/ling/olst/FrEng/PolgIntroTST.pdf>

Le lecteur qui voudrait pousser plus avant l'étude de la lexicologie et de la sémantique dans le cadre de la théorie Sens-Texte pourra utiliser l'ouvrage ci-dessous comme suite logique du présent cours :

Mel'čuk, Igor, André Clas et Alain Polguère (1995) *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve : Duculot, 256 pages.

Il importe de rappeler que tout cours de ce type, qui vise en priorité la cohérence plutôt qu'une couverture maximale de la discipline, est nécessairement incomplet. De nombreux aspects importants de l'étude du lexique des langues ont été volontairement laissés de côté, pour permettre un meilleur approfondissement de l'étude des notions retenues. Mon espoir est que ces dernières forment en quelque sorte le noyau notionnel de la discipline étudiée, à partir duquel le lecteur pourra élargir et solidifier ses connaissances.



••••• Annexe : Correction des exercices

[Dans cette version, je ne donne que les corrigés des exercices du Chapitre 1. Les corrigés de tous les exercices seront dans la version publiée.]

Cette annexe propose, pour la plupart des exercices donnés en fin de chapitres, une énumération des principaux éléments de réponse. Il est utile d'essayer de rédiger des réponses détaillées à partir des indications données ici. Pour cela, il convient de respecter les trois contraintes suivantes :

- 1 organisation logique de la réponse ;
- 2 expression claire ;
- 3 usage de la terminologie et des conventions d'écriture appropriées.

Chapitre 1

Les exercices pour ce chapitre sont volontairement faciles. Nous avons très peu de notions précisément définies à notre disposition pour parler de la langue et nous devons donc rester à un niveau très général. Les exercices deviendront plus « techniques » au fur et à mesure que nous progresserons dans le cours.

Exercice 1

- Un coup de poing au visage est, comme un sourire ou un signe de la main, un type de geste.

- On peut imaginer des situations où X donne un coup de poing à Y non vraiment pour lui faire mal, mais simplement pour lui montrer qu'il est en colère, etc. Dans ce cas, le coup de poing est bien la manifestation d'un acte de communication.
- Cependant, un coup de poing au visage est généralement donné uniquement pour assommer, pour porter atteinte physiquement, sans que ce geste ne vise aucune communication. C'est donc plutôt, par défaut, un geste qui n'est pas destiné à fonctionner comme un signe.
- Noter que si l'on voit X donner un coup de poing à Y et que l'on en déduit que Y est violent, ce coup de poing fonctionne pour nous comme un signe. Mais c'est alors un signe non intentionnel, comme on le verra au Chapitre 2. (X n'a pas posé ce geste pour exprimer quoi que ce soit. C'est nous qui l'interprétons d'une certaine façon.)

Exercice 2

Les éléments qui dans

(1) *Est-ce que tu peux me passer le sel ?*

identifient **explicitement** le locuteur et le destinataire sont :

- le pronom de « première personne » *me*, qui désigne le locuteur ;
- le pronom de « deuxième personne » *tu*, qui désigne le destinataire.

Ces pronoms, dont le sens ne se conçoit que par rapport aux participants de la situation de communication, sont appelés des **déictiques**. Nous examinerons cette notion au Chapitre 6.

Lorsque l'on se parle à soi-même, on est dans un cas un peu atypique où locuteur et destinataire coïncident. Il serait intéressant d'examiner comment s'emploient les pronoms de première et deuxième personne lorsque l'on se parle à soi-même. Emploie-t-on *je* ou *tu* ? Peut-on employer les deux ?

Exercice 3

J'insère directement dans le courrier électronique des commentaires sur les éléments qui me semblent relever de la langue parlée :

To: machin@LING.UMontreal.CA
 From: truc@pacific.net.sg
 Subject: Re: Je ne suis pas avare de vocables

Cc:
Bcc:
X-Attachments:

>Tout est reçu. On s'y met.

- Les chevrons nous indiquent que Truc, l'auteur du message, cite le message qu'il a reçu de Machin. C'est un peu comme s'il nous faisait entendre quelque chose qui a été enregistré. Cette façon de procéder est caractéristique du courrier électronique. Même si ce n'est pas une marque de l'oral, cela nous éloigne du texte écrit standard.
- Il arrive fréquemment à l'oral que l'on sous-entende les liens logiques entre phrases, comme c'est le cas ici. À l'écrit, on préfère en général utiliser des conjonctions ou des adverbes de phrase qui vont expliciter ces liens : *Tout est reçu. On peut **donc** commencer.*
- Noter de plus que *On s'y met* relève du langage familier, plus neutre à l'oral.

>J'arrive à Paris le 1 mai, et c'est un mercredi :

- Machin fait une transition brusque. Il change de sujet sans l'exprimer linguistiquement. Cela ne se fait pas dans un texte écrit standard : *Pour ce qui est de mes dates, j'arrive à Paris le 1^{er} mai.*

>Janine a dû se tromper.

>M.

Oui. Ça doit être l'impatience de te voir :-)

- Le *smiley* (aussi appelé *souriard* ou *binette*) est associé à la messagerie électronique. Il est vite devenu évident, lorsque l'usage du courrier électronique s'est généralisé, qu'il fallait avoir recours à certaines conventions pour palier l'absence de marquage gestuel ou intonatif de l'ironie, ou d'autres nuances expressives. Comme vous le savez certainement, il existe toute une batterie de conventions de ce type, cherchant à recréer les différentes expressions faciales (ou les intonations correspondantes) qui pourraient accompagner une phrase à l'oral :

: - / , :- (, : o) , etc.

À suivre...

Index des notions

Note : Le symbole ~ renvoie ici au terme indexé.

Symboles

- # [dièse] 155
- * [astérisque] 26, 155
- ≅ [quasi-équivalence] 96
- ~ [tilde] 37
- ≡ [équivalence exacte] 96

A

- abréviation 59
- acception 42
- acronyme 60
- acte de parole 190
 - ~ illocutoire 190
 - ~ locutoire 190
 - ~ perlocutoire 190
- adjectif 75
- adverbe 75
- affixe 52
 - ~ dérivationnel 55
 - ~ flexionnel 53
- allomorphe 51
- ambiguïté 161
 - ~ lexicale 161
 - ~ syntaxique 161
- analyse componentielle 164
- Anti** 133
- AntiBon** 139
- AntiMagn** 139

- antonymie 124
- argument (d'un prédicat) 107
- article (de dictionnaire) 179

B

- base d'une collocation 136
- Bon** 139

C

- catégorie flexionnelle 53, 73
- catégorie syntaxique [→ partie du discours] 73
- Caus** 166
- causatif 129
- causativité 129, 166
- CausOper₁** 167
- cercle vicieux 158
- champ sémantique 159
- circularité 158
- classe grammaticale
 - ~ fermée 76
 - ~ ouverte 74
- classe grammaticale [→ partie du discours] 73
- collocatif 136
- collocation 80, 135
- combinatoire
 - propriété de ~ 26, 112

composante sémantique 105, 118
composition 58
compositionnalité sémantique 40, 135
concordance 86
conjonction 76
connotation 105
conversivité 126
corpus linguistique 81

D

définition
~ analytique 153
~ circulaire 158
~ lexicale 145, 152
~ lexicographique 181
~ par genre prochain et différences
spécifiques 152
déictique 101, 200
dénotation 99
dénoter 100
dérivation 55
destinataire 7
déterminant 76
diachronie 11
dialecte 71
dictionnaire 176
~ bilingue 178
~ commercial 176
~ d'apprentissage 179
~ d'enseignement 179
~ de langue (monolingue) 178
~ de synonymes 182
~ encyclopédique 178
~ plurilingue 178
~ théorique 176
article (de ~) 179
entrée (de ~) 179
macrostructure (d'un ~) 179
microstructure (d'un ~) 180
nomenclature (d'un ~) 179
superarticle (de ~) 179
différences spécifiques 152
disjonction de sens 119

E

embrayeur [→ déictique] 114
énoncé
~ performatif 191
enquête linguistique 81
entrée (de dictionnaire) 179
évidence linguistique 106, 125
expression
~ agrammaticale 26
~ grammaticale 26
~ idiomatique 135
~ semi-idiomatique 135

F

flexion 37, 53
fonction lexicale 132
~ paradigmatic 133
~ syntagmatic 137
Anti 133
AntiBon 139
AntiMagn 139
Bon 139
Caus 166
CausOper₁ 167
Func_i 143
Magn 138
Oper_i 142
S₀ 134
S_i 134
Syn 133
V₀ 134
fonctions du langage 195
forme fléchie 54
fréquence d'emploi 82
Func_i 143

G

genre prochain 152
grammaire 10
graphe sémantique [→ réseau sémantique]
110

H

hapax 90

hiérarchie sémantique des lexies 121
homographie 128
homonymie 43, 127
homophonie 128
hyperonymie 120
hyponymie 120

I

icône 18
identité de sens 118
idiolecte 68
inclusion de sens 119
index (de signifiants lexicaux) 83
indice 19, 102
intersection de sens 118
introspection 81

K

KWIC [= *KeyWords In Context*] 86

L

langage 8
langue 4
 ~ écrite 7, 71
 ~ parlée 7, 71
lemmatisation 85
lexème 36, 37, 41
lexicographie 176
lexicologie 3, 10, 31
lexie 31, 41
 ~ à valeur pragmatique 193
 ~ antonyme 124
 ~ causative 129
 ~ contrastive 125
 ~ conversive 126
 ~ de base (d'un vocable) 166
 ~ homographe 128
 ~ homonyme 127
 ~ homophone 128
 ~ hyperonyme 120
 ~ hyponyme 120
 ~ synonyme 122
lexique 10, 64
lien lexical
 ~ paradigmatique 80, 132

 ~ syntagmatique 80, 134
linguistique 4
 ~ diachronique 11
 ~ quantitative 87
 ~ synchronique 11
locuteur 7
locution 38, 41, 135
 ~ adjectivale 39
 ~ adverbiale 39
 ~ nominale 38
 ~ prépositionnelle 39
 ~ verbale 39

M

macrostructure (d'un dictionnaire) 179
Magn 138
marque d'usage 69
message linguistique 188
métaphore 167
métonymie 167
microstructure (d'un dictionnaire) 180
morphème 49
morphème 51
morphologie 10, 48
mot 24, 33
mot grammatical 76, 106, 112
mot lexical 76
mot-forme 34

N

nœud (d'un réseau sémantique) 110
nom 75
nomenclature (d'un dictionnaire) 179

O

objet sémantique 108
occurrence 82
onomatopée 22, 102

Oper_i 142

P

paradigmatique [→ fonction lexicale] 133
paradigmatique [→ lien lexical] 80, 132
paraphrase 96, 122
parole 7, 185

partie du discours 55, 73
performatif 191
phonétique 10
polysémie 42, 128
pragmatique 186
prédicat sémantique 107
préfixe 53
préposition 76
pronom 76
propriété de combinatoire 26, 112

R

radical 52
référent 100, 101, 120, 188
relation sémantique lexicale 117
 antonymie 124
 causativité 129, 166
 conversivité 126
 disjonction de sens 119
 hyperonymie 120
 hyponymie 120
 identité de sens 118
 inclusion de sens 119
 intersection de sens 118
 métaphore 167
 métonymie 167
 synonymie (exacte/approximative)
 122
réseau sémantique 110
rhème 113

S

S₀ 134
S_i 134
sémantique
 ~ d'une langue 10, 187
 ~ linguistique (= étude sémantique)
 10, 187
sème 164
sémème 164
sémiologie [→ sémiotique] 17
sémiotique 17
sens
 ~ figuré 166
 ~ générique 152

 ~ grammatical 106, 153
 ~ lexical 106
 ~ linguistique 96, 97
 ~ logique 103
 ~ plus riche (qu'un autre) 120
 ~ plus simple (qu'un autre) 119
 ~ propre 166
disjonction de ~ 119
identité de ~ 118
inclusion de ~ 119
intersection de ~ 118
shifter [→ déictique] 114
siglaison 60
signe 16

 ~ complexe 25
 ~ élémentaire 25
 ~ grammatical 24
 ~ iconique 18
 ~ indiciel 19
 ~ intentionnel 16
 ~ lexical 24
 ~ linguistique 21
 ~ non intentionnel 16
 ~ segmental 50
 ~ suprasegmental 50
 ~ symbolique 19
 ~ zéro 54

signifiant 21
signifié 21
statistique lexicale 87
structuralisme européen 114
structure communicative 98, 111, 113
substantif [→ nom] 75
suffixe 53
superarticle (de dictionnaire) 179
symbole 19
Syn 133
synchronie 11
synonymie
 ~ approximative 122
 ~ exacte 122
syntactique du signe linguistique 25
syntagmatique [→ fonction lexicale] 137
syntagmatique [→ lien lexical] 80, 134
syntaxe 10

T

terminologie 71
test de substitution en contexte 122, 158
thème 113
théorie de la grammaire systémique fonctionnelle 114
théorie des actes de parole 190
théorie Sens-Texte 25, 130, 132, 197

U

unité lexicale [→ lexie] 31, 41

V

V₀ 134
vague 163
valeur (du signe linguistique) 99
valeur de vérité 103
verbe 75
verbe auxiliaire 76
verbe performatif 192
vocable 42
 ~ polysémique 128
vocabulaire
 ~ d'un individu 68
 ~ d'un texte 67





198
Melville, H. 175, 176
Michéa, R. 83
Moeschler, J. 195
Muller, C. 91
Murakami, H. 15

N

Nida, E. 60
Niklas-Salminen, A. 2
Nothomb, A. 147, 148

O

Oakes, M. 92

P

Palmer, F. R. 91, 146
Peirce, C. S. 114
Perec, G. 63, 64
Perrot, J. 90, 91
Picoche, J. 2, 99, 113, 183
Polguère, A. 172, 197, 198

Q

Queneau, R. 71

R

Reboul, A. 195
Rey, A. 183
Rey-Debove, J. 12, 65, 91, 183
Rivenc, P. 83

S

Saint-Exupéry, A. de 151
Saussure, F. de 9, 12, 21, 22, 23, 27, 50,
80, 99
Sauvageot, A. 83
Schaeffer, J.-M. 20, 28, 91, 195
Summers, D. 179

T

Touratier, C. 2
Tremblay, M. 115

W

Wierzbicka, A. 165, 172, 173, 177, 183