

ALISE LEHMANN
FRANÇOISE MARTIN-BERTHET

Lexicologie

Sémantique, morphologie,
lexicographie

5^e édition

ARMAND COLIN

■ Chapitre 1

Mot et lexique

1. Le mot
2. Le lexique

1. Le mot

Chaque domaine de la linguistique doit définir son unité : les unités de la phonétique et de la phonologie sont le *son* et le *phonème* ; l'unité de la syntaxe est la *phrase* ; l'unité du lexique est communément appelée *mot*. L'identité d'un mot est constituée de trois éléments : une forme, un sens et une classe syntaxique.

« Un mot résulte de l'association d'un sens donné à un ensemble de sons donnés susceptible d'un emploi grammatical donné. » (A. Meillet, *Linguistique historique et linguistique générale*, Paris-Genève, Champion-Slatkine, 1982 [1913], p. 30)

À ces trois aspects correspondent trois types de difficultés pour définir le mot et établir la liste des unités du lexique.

1.1 La classe syntaxique

Dans la définition du mot par Antoine Meillet citée ci-dessus, l'expression « emploi grammatical » désigne la classe syntaxique, par laquelle le mot s'intègre à la phrase. En français, la tradition grammaticale reconnaît huit *classes de mots*, appelées aussi *parties du discours* : nom, verbe, adjectif, déterminant, pronom, adverbe, préposition, conjonction.

L'expression traditionnelle **parties du discours** est le calque du latin *partes orationis*, où *partes* signifie « parties » par rapport au « tout » de l'énoncé (*oratio*) ; ce sont les unités constituantes de la phrase.

On peut parler aussi de *catégories morphosyntaxiques* (puisque des traits morphologiques interviennent dans leur définition, à savoir la variation en genre, nombre, personne) ou de *catégories grammaticales* ; mais ce dernier terme a aussi un autre sens : il désigne les catégories notionnelles qui ont une expression grammaticale, comme le temps, la personne, le genre, le nombre.

Dans la tradition de la grammaire scolaire, la classe du mot est sa **nature** (opposé à **fonction**).

Les mots comme *allo !, ouf !, chut !* sont recueillis dans une neuvième classe, celle des **interjections** ; mais il ne s'agit pas d'une classe de mots semblable aux autres : ce sont des mots qui sont des phrases (ou *mots-phrases*) et non des constituants de la phrase.

Certaines formes correspondent à plusieurs classes et ont des sens tout à fait distincts, comme *boucher* Nom et *boucher* Verbe. (Du point de vue étymologique, ils ont des origines différentes.) Ce sont donc deux mots différents, qui sont *homonymes* (cf. chap. 5, 1.3).

L'un des mots peut être dérivé de l'autre : seule la classe change, le sens étant globalement conservé ; par exemple, *déjeuner* Nom dérive de *déjeuner* Verbe. C'est une *conversion* (cf. chap. 8, 2).

La forme *faux* peut être a) un adjectif qualificatif (*une fausse note, ce passeport est faux*) ; b) un adverbe (*il chante faux* ; on parle aussi d'adjectif employé comme adverbe) ; c) un nom (*ce document est un faux*), qui vient de l'adjectif par conversion ; d) un autre nom homonyme (*couper l'herbe à la faux*, en relation avec le verbe *faucher*), qui n'a aucun rapport de sens avec le précédent (et qui a une autre origine). Selon les analyses, on distinguera deux mots (a/b/c + d), ou quatre (a + b + c + d).

De plus, certains mots ont plusieurs **emplois** pour une même classe, c'est-à-dire qu'ils entrent dans plusieurs constructions et contextes sans changement de classe. Par exemple, le verbe *casser* a un emploi transitif (*casser une vitre*) et un emploi intransitif (*la branche casse*) ; le *café* que l'on boit n'est pas celui où l'on va ; un *pauvre* homme n'est pas un homme *pauvre* ; le dérivé *pauvreté* est lié à ce second emploi. La différence de sens liée à la différence d'emploi est

plus ou moins importante (cf. ci-dessous 2.1) ; mais chaque emploi est isolable comme une unité du lexique.

1.2 Le sens

1.2.1 Mots homonymes et mot polysémique

À certaines formes de même classe correspondent plusieurs sens. Ainsi, *bise* signifie « vent du nord » ou « baiser » ; *perle* signifie « petite boule de nacre », « personne remarquable (dans sa partie : *une perle rare*) », « erreur ridicule (dans un texte) ». On devrait alors, pour s'en tenir à la définition donnée, considérer qu'il y a autant de mots différents que de sens différents (*homonymie*) ; dans le deuxième cas pourtant (*perle*), la notion d'un lien métaphorique entre les différents sens pourra suggérer que c'est un mot unique qui a plusieurs acceptions (*polysémie*). Il est souvent difficile de trancher avec la seule intuition sémantique (cf. chap. 5). Il y a des contextes différents (*cette personne est une perle, un collier de perles, perles relevées dans une copie* ; cf. ci-dessus *emplois*) mais aussi des contextes communs, non discriminants (*trouver une perle*).

1.2.2 Mots pleins et mots vides

Certains mots dits *vides* n'ont pas de sens par eux-mêmes : il s'agit surtout des **mots grammaticaux** (prépositions, conjonctions, déterminants, pronoms), opposés aux **mots lexicaux** ou *mots pleins* (noms, adjectifs, verbes, adverbes ; on parle aussi de *catégories majeures*). Ceux-ci ont un sens plus autonome et une référence : ils renvoient à un objet ou une notion définissable (cf. chap. 2).

On appelle **grammaticalisation** le « passage d'un mot autonome au rôle d'élément grammatical¹ » c'est-à-dire le fait qu'un mot lexical devienne grammatical ou acquière un emploi grammatical, avec une perte de sens. Par exemple, les formes verbales de participe présent *pendant* (verbe *pendre*) et *durant* (verbe *durer*) donnent des prépositions : *pendant ce temps, durant cette période* ; le verbe de mouvement *aller* devient un auxiliaire de futur proche : *je vais partir dans cinq minutes* ; les noms *pas, point* deviennent éléments d'adverbe négatif avec *ne*, à

1. A. Meillet, *Linguistique historique et linguistique générale*, Paris-Genève, Champion-Slatkine, 1982 [1912], p. 130-133.

partir de contextes comme : *je ne marche pas, je ne vois point!*. Le nom *question* « a donné naissance à une nouvelle préposition : *Question chômage, c'est catastrophique.* » : la grammaticalisation conjuguée à des degrés divers recatégorisation (morphologique), réanalyse (syntaxique) et réinterprétation² (sémantique).

B. Pottier oppose *lexèmes* et *grammèmes* et appelle *grammémisation* la transformation d'un lexème en grammème³.

Cependant, les mots grammaticaux sont plus ou moins vides : par exemple, certaines prépositions ont un sens spatial particulier (*sur, devant*). D'autre part, le statut grammatical implique moins l'absence de sens qu'un mode de signification : le mot grammatical a un sens très abstrait, il concourt à l'interprétation de la phrase en mettant en relation les mots lexicaux. On opposera alors *sens grammatical* et *sens lexical*.

Dans une terminologie plus récente, on oppose *sens instructionnel* (consistant plutôt en « instructions » guidant la mise au point du sens des mots et de la phrase) et *sens référentiel* ou *descriptif* (consistant à décrire les propriétés d'un référent ; cf. chap. 2, 1 ; chap. 3, 4).

1.3 La forme

1.3.1 Mot fléchi

Les mots variables ont plusieurs formes : ils ont une **flexion** (conjugaison, déclinaison, variation en genre et en nombre). Les différentes formes du mot variable sont aussi des mots : on peut parler du mot *grand* (adjectif), en tant qu'unité lexicale abstraite transcendant ses formes observables dans les phrases (*grand, grande, grands, grandes*), et du mot *grands* dans la phrase : *Les enfants sont grands* ; ou du mot *partir* (verbe) et du mot *partirons* dans la phrase : *Nous partirons à midi*.

On utilise conventionnellement l'une des formes du mot variable pour le nommer en tant qu'unité lexicale : le masculin singulier des mots variables en

1. G. Gougenheim, *Les Mots français dans l'histoire et dans la vie*, Omnibus, 2018 [Picard, 1966], p. 636, *Mots à sens plein devenus mots grammaticaux*.

2. Ch. Marchello-Nizia, *Grammaticalisation et changement linguistique*, p. 15, p. 42 sq.

3. *Linguistique générale, Théorie et description*, Klincksieck, 1974, p. 272.

genre et en nombre, le singulier des mots variables en nombre, l'infinif présent des verbes, le cas sujet des mots à déclinaison.

Cette forme est celle qui sert d'**adresse** (ou **entrée**) dans les dictionnaires (sauf indication de la variation en genre et en nombre par l'entrée : *blanc, blanche* ; *œil, yeux* ; cf. chap. 12, 1.3), ou de **lemme** dans le domaine du traitement automatique du langage. La **lemmatisation** consiste à indexer les différentes formes fléchies sous cette forme unique.

Si une des formes ou si chaque forme a un sens particulier, elle acquiert le statut d'unité lexicale différente. Cela arrive souvent avec la variation en genre : par exemple, *cuisinière* « appareil » n'est pas la forme féminine de *cuisinier* « humain » ; *loup* « demi-masque noir » n'est pas la forme masculine de *louve* « animal » ; *couturier* et *couturière* désignent des métiers différents. Le nombre peut aussi être concerné : on distinguera *la vacance* (d'un poste) et *les vacances* (les congés).

On utilise souvent le terme **lexème** (par analogie avec *phonème* et *morphème*) pour référer à l'unité lexicale abstraite. Ce terme est opposé à *mot-forme* ou *lexe* (cf. l'opposition *morphème / morphe*, voir chap. 7, 1) chez I. Mel'čuk¹.

1.3.2 Mot graphique

Fur, guingois, guise, plupart sont des mots graphiques : ils sont délimités à l'écrit par les blancs². Ce sont d'anciens noms. Aujourd'hui, ce ne sont plus des unités lexicales : les unités intégrantes de la phrase sont les expressions complexes dont ils font partie : locutions adverbiales (*de guingois, au fur et à mesure, à ma, ta, sa... guise*) ; locution prépositionnelle (*en guise de*) ; locution conjonctive (*au fur et à mesure que*) ; déterminant ou pronom indéfinis (*la plupart de Article N, la plupart*).

Pomme, de et *terre* sont trois unités lexicales du français ; *pomme de terre* en constitue une quatrième, composée des précédentes.

L'orthographe ne délimite pas toujours les mots (unités lexicales) ; l'identification et la délimitation de ceux qui sont composés de plusieurs mots

1. *Cours de morphologie générale*, Presses de l'Université de Montréal-CNRS Éditions, 1993, volume 1.

2. Le terme de typographie est *espace*, nom féminin.

graphiques, ou mots **polylexicaux**, doit se faire d'après des critères linguistiques (cf. chap. 9, 1).

2. Le lexique

L'ensemble des mots d'une langue constitue son lexique.

Cet ensemble est structuré par des relations entre ses unités ; il se diversifie selon un certain nombre de variables ; il n'est pas clos, et ses contours ne sont pas fixés de manière absolue.

2.1 Les relations

2.1.1 Relations paradigmatiques et syntagmatiques

Les unités de la langue sont reliées entre elles par deux grands types de relations :

- Les *relations* (ou *rappports*) **syntagmatiques** sont des relations d'enchaînement (ou concaténation) et de contiguïté : *le + livre, lire + un + livre* ; un *syntagme* est un groupement d'unités (*le livre, lire un livre, le livre de ton ami*).
- Les *relations* (ou *rappports*) **paradigmatiques** sont des relations d'équivalence ; un *paradigme* est l'ensemble des unités substituables dans un même enchaînement syntagmatique : par exemple, *le / un / ce / mon / chaque* /etc. (*livre doit être lu*) ; *lire un(e) livre / texte / roman / note / affiche* /etc.

Chaque point d'une séquence d'unités est le croisement d'un *axe vertical* (*paradigmatique*), où est sélectionnée l'unité, et d'un *axe horizontal* (*syntagmatique*), où les unités se placent les unes à la suite des autres, en raison de la *linéarité* du langage.

« Le signifiant, étant de nature auditive, se déroule dans le temps seul et a les caractères qu'il emprunte au temps : a) il représente une étendue, et b) cette étendue est mesurable dans une seule dimension : c'est une ligne. » L'écriture « substitue la ligne spatiale des signes graphiques à leur succession dans le temps. » Ferdinand de Saussure, *Cours de linguistique générale*, Payot, 1995 [1916], Première partie, Chapitre 1, p. 103.

L'opposition entre *paradigme* et *syntagme* est issue d'une « dichotomie saussurienne » voisine mais différente, où les rapports « verticaux » sont dits **associatifs** et sont indépendants des rapports horizontaux syntagmatiques : « les mots offrant quelque chose de commun s'associent dans la mémoire, et il se forme ainsi des groupes au sein desquels règnent des rapports très divers ». Par exemple, *enseignement* entre dans les *séries associatives* a) *enseignement, enseigner, enseignants*, etc. b) *enseignement, armement, changement*, etc. c) *enseignement, instruction, apprentissage, éducation*, etc. Le *rapport associatif* unit des termes *in absentia* dans une série mnémorique virtuelle. « Le *rapport syntagmatique* est *in praesentia* : il repose sur deux ou plusieurs termes également présents dans une série effective » ; son support est l'étendue. Ces deux ordres « correspondent à deux formes de notre activité mentale, toutes deux indispensables à la vie de la langue. » *Cours de linguistique générale*, Deuxième partie, Chapitre V, p. 170-175)

Le terme *syntagme* désigne chez Saussure toute combinaison d'unités consécutives : *re-lire, contre tous, la vie humaine, Dieu est bon, s'il fait beau, nous sortions*, etc. (p. 170). Il désigne plutôt aujourd'hui un groupe de mots syntaxique à l'intérieur de la phrase.

Le terme **paradigme**, emprunté au grec ancien, vient de la grammaire, où il signifie :

- 1) « exemple », « modèle » (c'est le sens du mot en grec) ; c'est le mot type, celui qui sert de modèle pour présenter une conjugaison ou une déclinaison : par exemple, *chanter* est le paradigme des verbes du premier groupe ; en latin *rosa* est le paradigme de la première déclinaison des noms.
- 2) « liste close de formes fléchies » (conjugaison, déclinaison), par exemple les formes du verbe *chanter* au présent de l'indicatif ; on passe de l'exemple à ce qu'il exemplifie, par métonymie.

Le point commun entre ces sens en grammaire et le sens linguistique évoqué ci-dessus est la notion de rapports « verticaux » et « *in absentia* » par opposition aux rapports syntagmatiques, « horizontaux » et « *in praesentia* ».

2.1.2 Relations sémantiques, morphologiques, syntaxiques

Le terme **relations lexicales** (relations entre les mots) désigne communément les **relations de sens** entre mots de même classe syntaxique et sans lien morphologique : synonymie, antonymie, hyperonymie et hyponymie, méronymie et holonymie (voir chap. 4). Une relation sémantique comme l'antonymie est lexicale (*grand / petit*) ou morphologique (*juste / injuste*). Une certaine synonymie peut exister aussi entre mots de classe syntaxique différente, comme

entre un verbe et un nom d'action ; elle est lexicale (*tomber* sujet humain et *chute*) ou morphologique (*chuter* et *chute*) (voir chap. 8, 1.2.2).

Les **relations morphologiques** sont des relations de forme et de sens : elles ont lieu entre un mot simple et les mots construits qu'il produit (dérivés, composés) et entre les mots construits de même structure (voir Deuxième partie, « Morphologie lexicale »).

Les **relations syntaxiques** sont des relations combinatoires (syntagmatiques) de construction (comme entre un verbe et ses objets : par exemple, écrire + lettre, livre, message, etc.) et de collocation (une peur intense est une peur bleue ; voir chap. 4, 2.1.2). Ce sont aussi les relations paradigmatiques de classe et sous-classe (nom, nom d'animé, verbe, verbe intransitif, etc.).

Ces trois types de relations sont intimement liés. La combinatoire syntaxique est associée à des compatibilités sémantiques. Une relation de forme sans relation de sens est purement phonétique (et graphique) et relève de la paronymie ou de l'homonymie (voir chap. 5, 1). Les relations morphologiques portent sur des emplois définis syntaxiquement (voir *supra* 1.1 ; chap. 8, 1.2.1). Une relation comme celle entre un verbe et le nom d'action synonyme concerne à la fois la morphologie, le sens et la syntaxe.

Fonctions lexicales. Le modèle Sens-Texte (qui cherche à représenter comment on va du sens au texte qui l'exprime) regroupe les liens entre les mots en deux grands types : les *dérivations sémantiques* et les *collocations*.

Les *dérivations sémantiques* comprennent non seulement les relations de sens classiquement répertoriées (telles que synonymie, antonymie) mais aussi le changement de classe grammaticale (*marchander / marchandage, vendre / vente, frapper / coup ; terre / terrestre, ville / urbain*), et la relation situation / élément de situation, participant (actant) ou circonstant (*nager / nageur, dormir / lit*). Elles sont ou non morphologiques. (Cette notion est reprise à J. Picoche, voir chap. 8, 1.2.2). Les *collocations* associent une base à ses *collocatifs* : ex. *dormir : comme une souche*.

Ces liens sont représentés comme des *fonctions* (au sens mathématique) appliquées à des *arguments* : « instrument » (*couper*) > *couteau, hache*, etc. ; « intensificateur » (*méchant*) > *comme la gale, comme une teigne*¹.

1. I. Mef'čuk, A. Polguère, *Lexique actif du français*, Bruxelles, 2007, p. 18-23.

2.2 Les variables

2.2.1 Les domaines

Le lexique général est commun à tous les locuteurs ; les *lexiques de spécialité* sont liés à un *domaine* : science (chimie, astronomie), science et technique (informatique), métier (menuiserie), activité (jardinage), métier et activité (politique), etc. L'étude des lexiques de spécialité est la *terminologie*.

2.2.2 Les usages

Une langue vit longtemps et peut être parlée dans un espace très étendu : l'usage du lexique varie selon deux facteurs principaux, qui sont le temps et le lieu. Il se diversifie aussi selon les groupes sociaux et la situation de communication.

La variation **diachronique** a lieu dans le temps. Si un fonds stable demeure depuis les débuts du français, le lexique du français contemporain n'est pas identique à celui des périodes précédentes : certains mots anciens ne sont plus utilisés ni compris (ex. *malandrin*) ; des mots nouveaux, ou *néologismes*, apparaissent (ex. *zapper*). Cette variation est sensible entre les générations.

La variation **diatopique** a lieu dans l'espace. Les régions de France et les pays de la francophonie présentent des mots qui leur sont propres : ainsi *cheni* « saleté, désordre », *panosse* « serpillère » sont employés en Savoie et en Suisse ; *magasiner* est un québécoisme qui signifie « aller faire des achats dans les magasins ».

La variation **diastratique** est liée aux groupes sociaux (représentés comme des *strates*) : certaines façons de parler sont considérées comme « populaires », d'autres au contraire comme marquant l'appartenance à un milieu social favorisé et cultivé. Certains groupes ont des argots spécifiques (« argot des malfaiteurs », argot scolaire, verlan des « jeunes des cités »). Cependant, la détermination sociale impliquée par cette notion est moins facile à appliquer au lexique qu'à d'autres aspects du langage, comme la prononciation ou la syntaxe.

La variation **diaphasique** (-*phas*- « parler », cf. *aphasie*) est liée à la situation de communication, qui sera plus ou moins familière ou formelle, induisant un vocabulaire plus ou moins familier ou recherché chez un même locuteur ou entre des locuteurs par ailleurs non différenciés par les autres facteurs.

Les dimensions *diastatique* et *diaphasique* ont des expressions voisines ; elles sont souvent confondues et exprimées en termes de *registres* ou de *niveaux de langue*. La *diastatie* est aussi *diatopie* dès lors que les groupes sociaux sont séparés dans l'espace : c'est ce que montre une expression comme « langage des cités ». Les classes d'âge (« langage des jeunes ») concernent la *diachronie* et la *diaphasie*. Les notions de *strates* et de *niveaux* peuvent véhiculer une dimension hiérarchique et une attitude normative, au détriment de leur vocation descriptive.

La description des usages est complexe, comme le montrent les marques lexicographiques courantes, qui se fondent sur des critères divers : situation de communication (*familier*), attitude de l'énonciateur (*péjoratif*), jugement moral ou esthétique (*vulgaire*), etc. La marque *littéraire* est tantôt une marque de domaine (elle réfère à un type de texte), tantôt une marque de registre (signifiant « soutenu »). Un terme comme *populaire* signifie plutôt « très familier » que « utilisé par les classes populaires ». (Voir chap. 11, 1.2 ; chap. 12, 2)

Le terme *diachronique* est, chez Saussure, opposé à *synchronique*. Les autres termes ont été forgés pour rendre compte de la variation synchronique. Les termes *diatopique* et *diastatique* sont dus au norvégien Leiv Flydal. Eugenio Coseriu les reprend et introduit le terme *diaphasique* pour traiter plus spécifiquement des registres liés aux situations de communication, tels que « usuel », « solennel », « familier ». Toutefois, les aspects sociaux et les niveaux de langue étant très mêlés, on utilise couramment aujourd'hui *diastatique* de façon extensive¹.

Les mots attachés à l'usage d'un temps, d'un lieu ou d'un registre sont dits *marqués*, par opposition à un fonds commun *non marqué*.

2.2.3 La fréquence

Un petit nombre de mots très fréquents constituent un noyau autour duquel se superposent des couches de fréquence moindre, jusqu'aux mots très rares et aux *hapax*, qui sont les mots dont on a une seule attestation. La fréquence est évaluée intuitivement par les locuteurs d'après leur expérience : par ex. chacun sent que *quand* est plus fréquent que *lorsque*. La fréquence est une donnée

1. E. Coseriu, « Structure lexicale et enseignement du vocabulaire », *Les théories linguistiques et leurs applications*, Conseil de la coopération culturelle, AIDELA, Nancy, 1967, p. 33. Pour un développement de cette nomenclature par F.-J. Haussman, voir l'*Encyclopédie internationale de lexicographie*, Berlin / New-York, Walter de Gruyter, 1989-1991, t. 1, p. 649-657.

objective quand elle est tirée de l'observation des occurrences du mot dans un *corpus*, ou ensemble de textes ; mais elle est alors dépendante de ce corpus. Les moyens informatiques permettent aujourd'hui de réunir des corpus suffisamment importants pour fournir une bonne image de la fréquence des mots. Pour le français, on dispose notamment, grâce au laboratoire ATILF (Analyse et traitement informatique de la langue française, CNRS – Université Nancy 2, www.atilf.fr), de la base FRANTEXT, qui rassemble un très grand nombre de textes à dominante littéraire.

Les expressions *mot courant* ou *sens courant* sont ambiguës, parce qu'elles peuvent renvoyer aux trois variables évoquées ci-dessus : *courant* peut s'opposer 1) à *spécialisé* 2) à *marqué* 3) à *rare*. La rareté est d'ailleurs souvent liée au caractère spécialisé ou marqué.

Le terme *registre* est parfois utilisé de manière extensive, renvoyant à tous les aspects indiqués ici (domaine, usage, fréquence).

La *fréquence absolue* est le nombre d'occurrences dans un corpus. La *fréquence relative* est rapportée à la longueur du corpus.

Le TLF, *Trésor de la langue française, Dictionnaire de la langue du XIX^e et du XX^e siècle* (www.atilf.fr/tlfi), donne à la fin de certains articles la fréquence absolue de l'entrée dans les textes littéraires et les fréquences relatives dans les quatre demi-siècles, d'après le *Dictionnaire des fréquences, Vocabulaire littéraire des XIX^e et XX^e siècles* d'É. Brunet¹.

Dans *Le vocabulaire français de 1789 à nos jours, d'après les données du Trésor de la langue française*², É. Brunet donne les fréquences des 6 700 mots de fréquence absolue supérieure à 500 dans FRANTEXT, par genre et par période.

J. Picoche et J.-C. Rolland, pour établir la nomenclature de leur *Dictionnaire du français usuel*³, sont partis des 907 mots de fréquence supérieure à 7 000 dans la liste d'É. Brunet, pour sélectionner 442 entrées consistant en mots lexicaux hyperfréquents et polysémiques qui conduisent, par leurs relations sémantiques et morphologiques, à un lexique de 15 000 mots.

Il faut noter que des mots usuels dans les échanges quotidiens apparaissent peu dans les listes statistiques, parce que leur apparition est liée à des thèmes

1. Paris, Didier, 1971.

2. Tome II, Slatkine, 1981.

3. De Boeck-Duculot, 2002.

ou à des situations : c'est le cas notamment des noms d'objet comme *livre*, *chaise*. Aussi, lorsqu'on a cherché dans les années 1950 à établir un « français fondamental » pour l'enseignement, a-t-on corrigé la fréquence par la *disponibilité*, évaluée par des enquêtes auprès d'un échantillon de locuteurs¹.

L'étude des phénomènes de fréquence lexicale dans les corpus est la *lexicométrie*, ou *statistique lexicale* (encore appelée *statistique linguistique* ou *linguistique quantitative*, dans la mesure où les comptages peuvent porter aussi sur des aspects morphosyntaxiques, comme les temps des verbes). Depuis P. Guiraud², elle a porté notamment en France sur des œuvres littéraires³ et sur le vocabulaire politique⁴. Elle entre aujourd'hui dans le cadre plus général de la *linguistique de corpus*.

2.3 Les marges

Le lexique comporte à ses marges un ensemble flou de mots plus ou moins intégrables dont le statut est incertain.

2.3.1 Noms propres

On considère que, à la différence des noms communs et des autres mots de la langue, ils n'ont pas véritablement de sens, mais seulement un référent (cf. chap. 2), qui est une entité individuelle : *Paris*, *la Loire*, *Zola*.

Ce point a été débattu par la philosophie : pour certains, il y a un sens des noms propres, qui se constitue des caractéristiques de leur référent et se dégage de leurs emplois dans les discours. De plus, il existe des classes sémantiques de noms propres : noms de pays, prénoms, noms d'animaux domestiques, etc.

1. G. Gougenheim, R. Michéa, P. Rivenc, A. Sauvageot, *L'élaboration du français fondamental (1^{er} degré)*, Didier, 1964.
2. *Les caractères statistiques du vocabulaire*, PUF, 1954, *Problèmes et méthodes de la statistique linguistique*, PUF, 1960.
3. Ch. Müller, *Étude de statistique lexicale, Le vocabulaire du théâtre de Pierre Corneille*, Slatkine, 1993 [Larousse, 1967] ; É. Brunet, *Le vocabulaire de Jean Giraudoux, Structure et évolution*, Slatkine, 1978 ; D. Mayaffre éd., *Etienne Brunet - Écrits choisis Volume 1 : Comptes d'auteurs. Études statistiques. De Rabelais à Gracq*, Champion, 2009.
4. M. Tournier et « groupe de Saint-Cloud », *La parole syndicale, Étude du vocabulaire confédéral des centrales ouvrières françaises, 1971-1976*, PUF, 1982 ; D. Mayaffre, *Le poids des mots, Le discours de droite et de gauche dans l'entre-deux-guerres*, Champion, 2000.

Certains noms propres entrent dans le lexique par leurs dérivés, qu'il s'agisse de noms de lieux (*français*, *parisien*, *savoyard*) ou de noms d'auteurs (*cornélien*, *rabelaisien*, *marxisme*) et de personnages (*gargantuesque*, *ubuesque*).

D'autres passent dans la langue en devenant des noms communs. C'est le cas des noms de personnages emblématiques désignant un type d'individu : un *don Juan* ou *don juan* est un séducteur, un *tartufe* ou *tartuffe* est un hypocrite, une *Lolita* ou *lolita* (du nom de l'héroïne éponyme du roman de Nabokov) est une « nymphette », un *Machiavel* est un « homme d'État sans scrupule » (définitions du *Petit Robert*) ; Mécène, Mentor (*un mécène*, *un mentor*) sont des personnages de l'histoire et de la littérature antiques. La rhétorique classique a répertorié ce fait dans les « figures de mots », sous le terme *antonomase*. Ils peuvent prendre la marque du pluriel (*les don Juan* ou *les dons juans*, *les Lolita* ou *les lolitas*, *les tartuffes*, *les Machiavels*, *les mécènes*) et donner des dérivés : *donjuanesque* et *donjuanisme*, *tartuf(f)erie*, *machiavélique* et *machiavélisme*, *mécénat*. Le passage à l'initiale minuscule ne se fait pas toujours (*Machiavel*) : la lexicalisation comporte des degrés ; l'allusion culturelle est plus ou moins présente, le registre est souvent soutenu.

Certains auteurs parlent de *nom propre métaphorique*, par opposition aux *noms propres métonymiques* tels que : un Picasso = « un tableau de Picasso ». Les noms de personnages emblématiques entrent aussi à titre de parangon dans des comparaisons figées à valeur intensive, appartenant là encore à un registre plutôt soutenu ; ils gardent alors la majuscule et la référence particulière : *fier comme Artaban*, *riche comme Crésus*, *vieux comme Hérode*.

Des noms d'inventeurs ou de responsables, généralement oubliés, deviennent par métonymie le nom de leur invention ou de l'objet qu'ils ont promu : *godillot* (Alexis Godillot, fournisseur de l'armée), *poubelle* (Eugène Poubelle, préfet de la Seine), *guillotine* (le docteur Guillotin).

Les noms de marque (noms déposés) s'intègrent au lexique ou aux phrases de deux manières différentes :

- ils servent à désigner un type de produit : *frigidaire*, *goretex*, *sopalin* ; cet emploi est critiqué par les puristes ;
- ils fonctionnent comme noms de sous-classes d'objets (*hyponymes*, cf. chap. 4, 1.1) : *rouler en Renault* (sorte de voiture), *mettre ses Nike* (sorte de chaussures), *être en Chanel* (sorte de vêtements) : dans ce cas, ils gardent

leur majuscule, restent invariables, et les dictionnaires de langue ne les répertorient pas.

Inversement, des noms communs deviennent des noms propres : c'est le cas des noms d'événements ou de périodes (*la Renaissance*), ou encore d'institutions (*l'Assemblée nationale*).

2.3.2 Mots étrangers

Toute langue emprunte à d'autres une partie de son lexique. Les mots étrangers sont « hors système » : ils ont leurs particularités phonétiques et morphologiques. Les *emprunts* sont assimilés quand ils sont conformes aux structures du français (*sentimental*, par exemple, est un anglicisme daté de 1769), ou quand ils se coulent dans ses moules phonétiques, orthographiques et morphologiques : *beefsteak* devient *bifteck* ; *blog*, *look* donnent *blogueur*, *relooker*. En revanche, *apartheid* (afrikaans) ou *tchador* (persan) comportent des sons qui ne sont pas dans le phonétisme français. Nombre de mots étrangers peu utilisés restent aux marges du lexique.

Emprunt, calque, xénisme. Le terme *emprunt* désigne tout élément provenant d'une autre langue, il a une valeur générique. Un *calque* est un emprunt qui résulte de la traduction littérale d'une expression (*col blanc* : anglais *white-collar* ; *gratte-ciel* : *skyscraper*), d'un dérivé (*surpêche*, *surréagir* : anglais *overfishing*, *overreact*) ou d'une acception figurée (*souris* au sens de « boîtier connecté à un ordinateur » : anglais *mouse*). Dans ce dernier cas, on parle aussi d'*emprunt sémantique*. Le terme *xénisme* est réservé à l'emprunt qui correspond à une réalité étrangère (afrikaans *apartheid*, russe *toundra*).

L'emprunt aux langues anciennes (grec et latin). Un mot peut être emprunté au grec ancien ou au latin de la même façon qu'aux langues modernes : *humus* est latin ; *agora*, *chaos*, *misanthrope* sont grecs. Deux autres types d'emprunt aux langues anciennes ont un statut différent :

- le latin étant à la fois langue mère et langue d'emprunt, le français a des *mots savants* empruntés au latin, qui parfois doublent un *mot populaire* de même étymon : le latin *auscultare* donne en français *écouter* et par emprunt *ausculter* (voir chap. 6, 1) ;
- le français emprunte au latin et au grec des *éléments* pour construire des *composés savants* (*omnivore* ; *xénophobe*), des *dérivés savants* dont le radical a la forme d'un mot latin (*épiler* de *pilus* « poil », *carcéral* de *carcer* « prison ») ou pour servir de préfixe ou de suffixe (*hypersensible*, *insecticide*).

Il s'agit alors de modes de formation internes au français. Voir chap. 6, 2.1 ; chap. 8, 1.2.2 ; chap. 9, 3.

2.3.3 Mots nouveaux

L'apparition d'un mot nouveau, ou *néologisme*, est un fait historique (qu'il n'est pas toujours facile de dater d'ailleurs) qui a lieu dans un discours individuel, par nécessité (nommer une réalité ou une notion nouvelle), par recherche d'expressivité, par jeu, etc. Il s'intègre ou non dans la langue comme bien commun de la collectivité¹.

La nouveauté est dans la forme, le sens ou la syntaxe :

- Les néologismes de forme créent une nouvelle forme ; pour la plupart, ils sont produits par les procédés de formation que sont la dérivation (cf. chap. 8) et la composition (cf. chap. 9) : par ex. *démondialisation*, *écoquartier*, *monte-escalier* ; ce sont aussi des emprunts : *selfie*, *sudoku* ;
- Les néologismes de sens créent un nouveau sens pour une forme donnée, selon certains mécanismes sémantiques, tels que les métaphores, les métonymies, les extensions ou restrictions (cf. chap. 5, 3) : par ex., on peut parler de *polluer l'atmosphère* dans un sens moral et psychologique, par métaphore.
- Les néologismes syntaxiques créent un nouvel emploi, soit par changement de classe grammaticale (ou *conversion*, cf. chap. 8, 2), soit par changement de construction : par ex. à *angoisser* transitif à objet humain (*cet endroit m'angoisse*) s'ajoute depuis peu une construction intransitive à sujet humain : *j'angoisse* « j'éprouve de l'angoisse ».

Un nouveau sens est souvent aussi un nouvel emploi : dans le « langage des cités », l'adjectif *moelleux* s'applique à une personne (nouvel emploi) avec le sens « mou, sans énergie² » ; *être (un) décroissant* (Adj et N) quand il s'agit d'une personne, c'est remettre en cause la croissance économique.

On appelle parfois mots *théoriques* ou *possibles* les formes produites par le système virtuel de la langue et non (encore) attestées, dans lesquelles puisent les néologismes de forme.

1. J.-F. Sablayrolles, *La néologie en français contemporain, Examen du concept et analyses de productions néologiques récentes* ; J. Pruvost, J.-F. Sablayrolles, *Les néologismes*.

2. *Lexik des cités*, Éditions Fleuve, 2007.

« Un mot que j'improvise, comme *in-décor-able*, existe déjà en puissance dans la langue ; on retrouve tous ses éléments dans les syntagmes tels que *décor-er, décor-ation : pardonn-able, mani-able : in-connu, in-sensé*, etc., et sa réalisation dans la parole est un fait insignifiant en comparaison de la possibilité de le former. » F. de Saussure, *Cours de linguistique générale*, Payot, 1996 [1967], Troisième partie, Chapitre IV, p. 227

Le lexique attesté s'augmente ainsi de la masse des mots « en puissance » dans la langue. Cependant, on ne peut pas dire quels sont les mots « théoriques » sans avoir élucidé tous les mécanismes de la formation des mots ; or, ces mécanismes ne sont pas totalement décrits. Il est encore plus difficile de prédire les sens ou les emplois possibles. Il faut noter aussi que l'usage n'est pas indifférent : savoir que *irrécupérable* est lexicalisé, mais non *indécorable*, ou *imbattable* et moins *battable* (cf. chap. 6, 2.3 ; chap. 7, 2.3.3), fait partie de la compétence lexicale.

Partie 1

Sémantique lexicale

■ Chapitre 2

Le signe linguistique

1. Le signe linguistique et le référent
2. Extension vs intension
3. Dénotation vs connotation(s)

La sémantique lexicale a pour objet l'étude du sens des unités lexicales. Elle se sert des concepts liés au signe, hérités souvent de la philosophie. On définira brièvement ici les concepts fondamentaux.

1. Le signe linguistique et le référent

1.1 Le signe

Ferdinand de Saussure (1857-1913) définit le signe linguistique comme une entité à double face, l'une sensible qui est son **signifiant** (l'image acoustique), l'autre, abstraite, qui est son **signifié** (le concept).

Signifiant et signifié s'impliquent réciproquement. Au signifiant oral [flœR] ou au signifiant graphique (*fleur*) est associé le signifié « production de certains végétaux, colorée, souvent parfumée ».

« Le signe linguistique unit non une chose et un nom, mais un concept et une image acoustique. Cette dernière n'est pas le son matériel, chose purement physique, mais l'empreinte psychique de ce son, la représentation que nous en donne le témoignage de nos sens [...] »

F. de Saussure, *Cours de linguistique générale*, première partie, chap. 1, p. 98

Le lien unissant le signifiant au signifié est arbitraire.

« Ainsi l'idée de "sœur" n'est liée par aucun rapport intérieur avec la suite de sons *s-ø-r* qui lui sert de signifiant ; il pourrait aussi bien être représenté par n'importe quelle autre... » (*Ibid.*, p. 100)

1.2 Le référent

Les signes linguistiques permettent au locuteur de parler de la réalité qui l'entoure. Ils ont, en effet, la propriété de pouvoir renvoyer aux objets du monde, extérieurs à la langue ; ces objets sont les référents. Signes linguistiques et référents ne doivent pas être confondus. Ainsi c'est bien la fleur (la plante concrète) qui embaume et non pas le mot *fleur* ni le signifié de *fleur*. La confusion entre le signe et le référent désigné par le signe s'observe dans l'apprentissage du langage. Dans sa thèse, I. Berthoud-Papandropoulou rapporte différents tests qu'elle présente à des enfants âgés de 4 à 6 ans¹. L'un des tests consiste à leur soumettre des mots de longueur différente en demandant de juger de leur dimension. À la question : « Pourquoi *armoire* est un mot long ? », l'enfant répond : « parce qu'il y a beaucoup d'affaires dedans ». Les enfants répondent à la question **métalinguistique** – c'est-à-dire portant sur la langue – par des réponses liées aux caractéristiques du référent. Ces réponses traduisent le refus de l'arbitraire du signe et le besoin de motivation (*cf.* chap. 6, 3).

Les référents sont des entités matérielles ou conceptuelles (êtres, objets, lieux, processus, propriétés, événements, etc.). Ils relèvent de l'univers extralinguistique réel ou fictif (par exemple, la sirène). Les référents ne sont pas tous des données immédiates du réel. Ils sont médiatisés par la langue. Les rapports entre l'ordre de la langue et l'ordre du monde sont complexes ; ils ont nourri les débats philosophiques dès l'Antiquité.

On distingue l'emploi référentiel du signe de son emploi autonymique. Lorsque le signe ne renvoie pas à un référent (externe) mais à lui-même, il est dit **autonyme** (il se désigne lui-même comme signe). Dans l'exemple : *Chat rime avec rat, chat est autonome* (on dit que le locuteur fait mention du signe). En revanche, dans l'exemple : *Le chat miaule sur le balcon, le chat renvoie au référent* (on dit que le locuteur fait usage du signe).

1. I. Berthoud-Papandropoulou, *La Réflexion métalinguistique chez l'enfant*, Université de Genève, 1980.

Ce n'est que par le passage de la langue au discours, par le biais d'un acte d'énonciation, que le signe réfère, c'est-à-dire qu'il permet au locuteur de désigner un objet. Dans l'énoncé : *Mon chat miaule sur le balcon, mon chat vise un référent particulier appartenant à l'univers du locuteur*

1.3 Le sens référentiel

Le sens référentiel d'une unité lexicale (aussi appelé **sens désignatif** ou **sens dénotatif**) est le signifié stable du signe qui correspond à la relation de désignation entre le signe et le référent. Dans cette conception, le sens d'un mot détermine sa référence ; il est constitué des propriétés qui permettent d'identifier et d'isoler une catégorie d'objets extralinguistiques par rapport à d'autres objets, même si le signifié ne prend pas en compte toutes les caractéristiques du référent. On peut donc dire que le sens référentiel rassemble les critères ou les informations que la langue a retenus pour référer à un objet extralinguistique. Ainsi, pour que le nom *chat* puisse être attribué à un individu particulier, il faut que ce dernier satisfasse à un ensemble de conditions comme « animal », « quadrupède », « petit », « domestique », « miaule », etc. (*cf.* chap. 3, 1.1 ; 3.1).

À la suite de J.-C. Milner¹, on distingue parfois **référence actuelle** (référence en discours) et **référence virtuelle** (hors discours). Le sens référentiel équivaut alors à la notion de référence virtuelle².

1.4 Le sens différentiel

Pour Saussure, le sens d'une unité lexicale est différentiel. La langue étant un système de signes dont tous les termes sont solidaires, le signifié du signe linguistique est déterminé par sa position à l'intérieur du système linguistique, c'est-à-dire par les rapports qu'il entretient avec les autres signifiés voisins qui lui sont opposables.

« Dans l'intérieur d'une même langue, tous les mots qui expriment des idées voisines se limitent réciproquement : des synonymes comme *redouter, craindre, avoir peur* n'ont de valeur propre que par leur opposition ; si

1. J.-C. Milner, *Introduction à une science du langage*, Paris, Seuil, 1989, p. 336.

2. Sur les rapports entre sens et référence, *cf.* G. Kleiber, *Problèmes de sémantique. La polysémie en questions*, p. 30-52.

redouter n'existait pas, tout son contenu irait à ses concurrents. » (F. de Saussure, *op. cit.*, deuxième partie, chap. 4, p. 160)

La conception différentielle du sens est à la base de l'analyse sémique. Ce point est développé dans le chapitre 3, II.

2. Extension vs intension

L'extension et l'intension (ou *compréhension*) sont des concepts issus de la logique, d'un emploi courant en lexicologie. Le terme *compréhension* est plus ancien mais il prête à confusion ; l'anglicisme *intension* est plus clair en ce qu'il s'oppose morphologiquement à *extension*.

Lorsque l'on définit une classe d'objets, on peut opter entre deux solutions :

- énumérer les éléments dont se compose la classe (définition en extension) ;
- définir la classe à l'aide des propriétés communes aux objets de la classe (définition en intension).

Sur le plan linguistique, l'extension d'un signe est l'ensemble des référents auxquels il s'applique ; l'intension d'un signe est l'ensemble des traits qui constituent son signifié (autrement dit son sens dénotatif). L'extension de *fleur* est l'ensemble des fleurs (tulipes, roses, marguerites, etc.) ; l'intension de *fleur* est le sens de fleur composé des traits sémantiques /production/, /venant de végétaux/, /colorée/, /souvent parfumée/. Extension et intension sont en relation complémentaire (cf. chap. 4, 1.1). L'intension d'un signe détermine son extension.

Certains mots, comme les termes de couleur, ne sont guère susceptibles d'une analyse en traits sémantiques : on peut décomposer *rouge* en /coloré/ mais comment poursuivre ? Ils requièrent donc une définition en extension, qui consiste à énumérer les objets de la couleur à définir (*rouge* : couleur du sang, du coquelicot, du feu, du rubis, etc.).

3. Dénotation vs connotation(s)

En logique, la dénotation est l'extension d'un signe, et la connotation son intension. On écartera ici l'acception logique du terme *connotation*.

En linguistique, le **sens** ou **signifié dénotatif** s'oppose au **sens** ou **signifié connotatif**.

La connotation d'un signe est constituée des valeurs sémantiques secondes qui viennent se greffer sur le sens dénotatif. Dans le domaine du lexique, la connotation recouvre différents faits : registres de langue (ainsi *policier* et *fic* n'ont pas la même connotation), contenus affectifs propres à un individu ou à un groupe d'individus (par exemple, le mot *paysan* a, selon les cas, une valeur positive ou négative), représentations culturelles et idéologiques liées aux contextes d'utilisation de l'unité lexicale ou en rapport avec les référents (par exemple, le mot *peuple*).

Clochard et *SDF* (*sans domicile fixe*) se différencient par leurs connotations : *clochard* peut avoir, dans certains contextes, une connotation affective que n'a pas l'euphémisme *SDF*. Le mot *sans-papiers* s'est diffusé à partir des années 1980 parce qu'il n'avait pas la charge négative de *clandestin*.

Les valeurs connotatives sont hétérogènes et variables selon les locuteurs. Bien que commode et d'un usage généralisé, la notion de connotation reste floue ; de là le choix du pluriel qui traduit la diversité des faits.

De plus, les critères de démarcation entre traits dénotatifs et traits connotatifs ne sont pas aisés (cf. chap. 3, 2.2.2). Selon C. Kerbrat-Orecchioni¹, ce qui fonde le statut spécial des valeurs sémantiques de la connotation tient à leur nature et/ou à leur modalité d'affirmation :

- Leur nature est spécifique parce que « les informations qu'elles fournissent portent sur autre chose que le référent du discours » et qu'elles renseignent sur le locuteur, la situation de communication, le genre de discours, etc.
- Leur modalité d'affirmation est spécifique : « véhiculées par un matériel signifiant beaucoup plus diversifié que celui dont relève la dénotation », puisque la connotation exploite la totalité du matériel linguistique, signifiant compris, « ces valeurs sont suggérées plus que véritablement assertées, et secondaires par rapport aux contenus dénotatifs auxquels elles sont subordonnées ». On ne saurait donc s'étonner de la difficulté que pose l'intégration des connotations dans une théorie sémantique.

1. C. Kerbrat-Orecchioni, *La Connotation*, Lyon, Presses Universitaires de Lyon, 1977, p. 18.

Un nom propre peut connoter. R. Barthes, dans son analyse de la publicité pour les pâtes *Panzani*, remarque que le signe *Panzani* n'est pas seulement le nom de la marque, mais « qu'il livre aussi, par son assonance, un signifié supplémentaire qui est, si l'on veut, "l'italianité"¹ ».

1. R. Barthes, « Rhétorique de l'image », *Communications* 4, Seuil, 1964, p. 41.

■ Chapitre 3

Les analyses du sens lexical

1. La définition par inclusion
2. L'analyse sémique ou componentielle
3. Prototypes et stéréotypes
4. Exercices (1, 2, 3, 4)

Pour la sémantique lexicale, les mots ont un sens en langue : il y a, sous les différentes occurrences (apparitions) d'un mot en discours, un invariant sémantique, un noyau stable inhérent au mot que l'on peut décrire en relation avec ses emplois et hors emploi. Comment rendre compte de ce sens codé, conventionnel, des unités lexicales ?

Les analyses du sens lexical diffèrent selon les modèles théoriques. On en retiendra trois qui forment le soubassement théorique de la sémantique lexicale. Les modèles présentés en 1. (la définition par inclusion) et en 3. (la théorie des prototypes et des stéréotypes) se situent dans une perspective de sémantique référentielle ou sémantique de la désignation (*cf.* chap. 2, 1.3) : le sens de l'unité lexicale est conçu en termes de traits référentiels, qu'il s'agisse des traits de la définition ou des traits typiques. Le modèle de l'analyse sémique (2.), en revanche, est orienté vers la sémantique de la signification (*cf.* chap. 2, 1.4) : le sens de l'unité lexicale est défini de manière différentielle par les rapports qu'il entretient avec les autres unités du système linguistique, en dehors de la référence. Le cadre théorique de l'analyse sémique est le structuralisme, celui de la sémantique du prototype se rattache au courant cognitif qui se développe à partir des années quatre-vingt.

■ Chapitre 4

Les relations sémantiques

1. Relations de hiérarchie et d'inclusion
2. Relations d'équivalence et d'opposition
3. Exercice (5)

Les relations sémantiques entre les unités lexicales sont de deux types :

- Relations hiérarchiques et d'inclusion lorsqu'elles concernent des unités qui n'ont pas le même rang (hyponymes et hyperonymes, relation partie-tout).
- Relations d'équivalence et d'opposition lorsqu'elles concernent des unités de même rang (synonymes, antonymes, co-hyponymes).

Ces relations, exception faite de la relation partie-tout, structurent le lexique sur le plan paradigmatique.

Ces termes ont en commun l'élément *-onyme* signifiant « appellation, nom » ; *hyponyme* et *hyperonyme* ont été créés sur le modèle de *synonyme* et *antonyme*, tout comme *méronyme* plus récent (il désigne la relation partie-tout) auquel répond *holonyme* (le tout par rapport à la partie).

1. Relations de hiérarchie et d'inclusion

1.1 L'hyponymie et l'hyperonymie

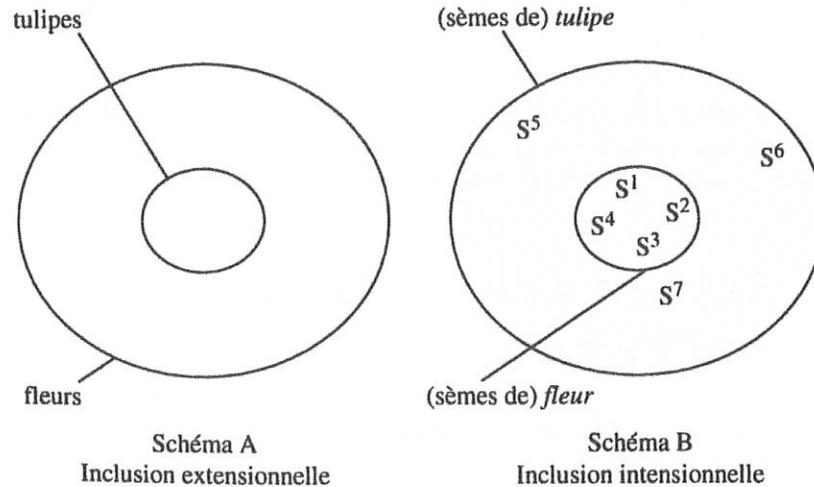
La relation d'hyponymie est une relation hiérarchique qui unit un mot spécifique (*sous-ordonné*), l'hyponyme, à un mot plus général (*superordonné*) nommé l'hyperonyme. Ainsi *tulipe* est l'hyponyme de *fleur*, *fleur* est

l'hyperonyme de *tulipe*, *morille* est l'hyponyme de *champignon*, *champignon* est l'hyperonyme de *morille*. Les co-hyponymes ont le même hyperonyme (cf. *infra*, 2.3). Ces rapports sont exploités, on l'a vu, dans les définitions lexicographiques : dans la définition de fonte « alliage de fer et de carbone [...] », *alliage* est l'hyperonyme de *fonte*, *fonte* étant l'hyponyme de *alliage*.

L'hyperonyme est aussi appelé *incluant*, *terme* (ou *mot*) *générique* et *archilexème* (dans le cadre de l'analyse sémique).

1.1.1 Inclusion extensionnelle et intensionnelle

Le rapport qui lie un hyponyme (*tulipe*) à un hyperonyme (*fleur*) est un rapport d'inclusion. Mais cette formulation est équivoque et il faut distinguer deux points de vue. Du point de vue de la référence, la classe des référents qui sont des tulipes est incluse dans la classe des référents qui sont des fleurs : l'inclusion est extensionnelle (schéma A). Du point de vue du sens, le sens de *fleur* est inclus dans le sens de *tulipe* : l'inclusion est intensionnelle (schéma B). Les sèmes de *fleur* : s^1 /production/, s^2 /venant de végétaux/, s^3 /colorée/, s^4 /souvent parfumée/ sont inclus dans le sémème de *tulipe* qui comprend, en outre, s^5 /racine bulbeuse/, s^6 /fleur évasée/, s^7 /couleurs vives/.



En conformité avec le principe selon lequel l'extension et l'intension (la compréhension) d'un terme sont en rapport inverse, l'inclusion extensionnelle est l'inverse de l'inclusion intensionnelle. L'hyponyme a une extension plus réduite que celle de son hyperonyme (tulipe dénote un sous-ensemble de fleurs) mais son intension est plus grande : il comporte un nombre plus élevé de sèmes. L'inclusion intensionnelle est en porte-à-faux avec les préfixes des dénominations retenues (*hypo-* et *hyper-*) qui se placent sur le terrain de la référence. La version extensionnelle, intuitivement plus satisfaisante, offre, de plus, l'avantage de convenir à la relation d'implication.

1.1.2 Relation d'implication

L'hyponymie établit un rapport d'implication unilatérale entre deux entités : *si x est une tulipe, alors x est une fleur* ; mais on ne peut pas dire : *si x est une fleur, alors x est une tulipe*. Cela explique que la relation s'établisse, dans le discours, de l'hyponyme à l'hyperonyme. L'hyperonyme, parce qu'il désigne ce que désigne l'hyponyme, peut reprendre – c'est-à-dire servir d'anaphorique à – l'hyponyme :

• *Un chat entra. L'animal était malade.*

et non

• **Un animal entra. Le chat était malade.*

De même la construction coordonnée par et d'autres n'admet que l'ordre suivant :

• *Paul a demandé des tulipes et d'autres fleurs.*

et non

• **Paul a demandé des fleurs et d'autres tulipes.*

1.1.3 Structures hiérarchiques

Un mot donné peut entrer dans une série d'inclusions successives qui dessinent des relations hiérarchiques dans le lexique.

Exemples :

- sapin/conifère/arbre/végétal
- tulipe précoce/tulipe/fleur/plante
- redingote/manteau/vêtement

Certains mots sont tour à tour hyponymes et hyperonymes : *manteau*, par exemple, est l'hyperonyme de *redingote* et l'hyponyme de *vêtement*. Les séries lexicales, contrairement aux inventaires taxinomiques, n'excèdent guère trois à quatre degrés. La hiérarchisation est bloquée vers le haut par la présence de noms très généraux (*chose*, *truc*) et vers le bas par des périphrases développées (par exemple, *redingote à double boutonnage et à parements de velours*). Les structures lexicales hiérarchiques diffèrent d'une langue à l'autre et peuvent présenter ce qu'il est convenu d'appeler des **trous lexicaux**. Aussi arrive-t-il que le besoin d'un hyperonyme se fasse sentir dans un domaine donné. *Deux-roues* a été créé en 1960 pour englober *scooter*, *vélocycle*, *bicyclette* et la série est devenue complète : *scooter/deux-roues/véhicule*.

Les rapports d'un terme à un autre à l'intérieur d'une série ne sont toutefois pas identiques. D'un côté, les hyperonymes qui se trouvent au sommet de la hiérarchie, tel *végétal* ou *animal*, dominent des sous-classes qui sont très disparates (pour *animal* : *oiseau*, *reptile*, *poisson*, *insecte*, etc.). J. Rey-Debove¹ suggère de réserver le terme de *superordonné* (calqué sur l'usage anglais) à cette catégorie d'hyperonymes pour laquelle il est impossible de choisir un prototype (cf. chap. 3, 3.2). D'un autre côté, la relation qui unit le mot composé au mot simple (*tulipe précoce* et *tulipe*) est plus étroite et détermine une subdivision interne à une classe, les hyponymes partageant les principales caractéristiques sémantiques de l'hyperonyme².

Certains mots relèvent d'une double classification : structures lexicales et catégories taxinomiques liées à l'organisation des connaissances. Les deux structures ne se recouvrent pas, même si elles peuvent parfois se superposer.

1. J. Rey-Debove, « Prototypes et définitions », *DRLAV* 41, 1989, p. 158.

2. Cf. G. Kleiber et I. Tamba, « L'hyponymie revisitée : inclusion et hiérarchie », *Langages* 98, 1990.

Ainsi l'hyperonyme de *chat* est soit *mammifère carnivore* ou *félin* (catégories zoologiques), soit *félin* ou *animal domestique*. Les définitions des dictionnaires portent la trace de ces hésitations (cf. exercice 1). Il peut, en effet, y avoir divergence entre le savoir non spécialisé et la connaissance spécialisée ; en botanique, l'hyperonyme de *courgette* n'est pas *légume* mais *fruit* et la baleine a été classée comme *poisson*, au XVII^e siècle, avant d'être cataloguée comme *mammifère*.

La relation hyper-/hyponymique touche différentes catégories syntaxiques : des verbes (*couper/cisailler*, *manger/grignoter*), des adjectifs (*rouge/pourpre*, *gail guilleret*) et surtout des noms, cela pour une double raison. D'une part, le nom est à la fois catégorie syntaxique (substantif) et outil de dénomination (or les relations de hiérarchie mettent en jeu des classes de référents) ; d'autre part, le nom se prête tout particulièrement à deux types de formation d'hyperonymes et d'hyponymes : changement de signe lexical (*tulipe fleur*), passage du mot simple au mot composé (voir l'exemple déjà cité *tulipe/tulipe précoce* ou *chat/chat siamois*, *table/table de bridge*).

À cela s'ajoute la catégorie des noms propres de marque. Cette formation a été exploitée dans les slogans publicitaires dans les années 1990. En voici un exemple : *Une Lancia c'est plus moderne qu'une voiture*. *Lancia* est hyponyme de *voiture* mais le slogan joue sur la transgression de la relation d'inclusion pour susciter l'interprétation selon laquelle une *Lancia* n'est pas une voiture (ordinaire).

Les relations hyperonymiques et hyponymiques ont un rôle central dans l'apprentissage du lexique. Le locuteur peut parler des objets dont il ne connaît pas le nom spécifique en recourant à l'hyperonyme, ou bien ne retenir que la seule relation liant l'hyponyme à l'hyperonyme sans rien savoir du sens de l'hyponyme (il saura, par exemple, qu'un varan est un lézard sans savoir en quoi il se différencie des autres lézards).

1.2 La relation partie-tout

La relation partie-tout est une relation hiérarchique qui existe entre un couple de termes dont l'un dénote une partie et l'autre dénote le tout (relatif à cette partie) : *guidon/bicyclette*, *poignée/valise*, *bras/corps*, *ongle/doigt*, *porte/maison*, *voile/bateau*. *Guidon* est une partie de *bicyclette* ou le **méronyme** (« nom de la partie ») de *bicyclette* ; inversement, *bicyclette* désigne le tout ou l'**holonyme** (« nom du tout ») de *guidon*. La relation partie-tout concerne les noms qui

renvoient à des référents divisibles et discrets, soit principalement les noms comptables.

Les noms communs sont classés en différentes sous-catégories syntactico-sémantiques : noms comptables/noms massifs, noms concrets/noms abstraits, noms animés/noms inanimés¹. Les noms comptables se prêtent au dénombrement, ce qui suppose la discontinuité, c'est-à-dire des entités distinctes ; ils se combinent avec les déterminants numéraux (*une, deux bicyclettes*). Les référents qu'ils dénotent présentent, pour la plupart, une substance hétérogène, susceptible d'être divisée en parties qui ne sont pas de même nature que le tout : le guidon par rapport à la bicyclette. De là leur aptitude à se prêter à la relation partie-tout. En revanche, les noms massifs sont indénumérables et se combinent avec l'article partitif (*de l'eau* et non **une, deux eaux*) ; ils supposent une substance homogène continue, dont la partition provoque une entité du même type : une partie de l'eau, c'est toujours de l'eau².

Les méronymes présentent quelques similitudes avec les hyponymes. Ils ont une composante de sens relationnel qui exige leur liaison à un autre terme : tandis qu'un N hyponyme est *une sorte de N hyperonyme* (*une tulipe est une sorte de fleur*), un N méronyme est *une partie du N holonyme*³. La relation de dépendance est orientée et récurrente, comme l'illustrent ces deux séries méronymiques : *ongle* partie du *doigt*, *doigt* partie de la *main*, *main* partie du *bras*, *bras* partie du *corps humain* ; et *ongle* (en tant que synonyme de *griffe*) partie de *patte*, *patte* partie de *animal*. La méronymie est, comme l'hyponymie, une relation logique d'implication (*doigt* implique *main*).

Mais la relation partie-tout se différencie de la relation d'hyponymie.

Elle exprime une appartenance que l'on peut paraphraser avec le verbe *avoir*. Soit *guidon/bicyclette*. *Le guidon est une partie de la bicyclette* se convertit en *La bicyclette a un guidon*. L'holonyme domine le méronyme en position d'objet. Au contraire, dans la relation hyponymique liée à l'opération d'inclusion, l'hyponyme est dominé par un nom attribut : *La tulipe est une fleur*.

D'autre part, et ce trait est essentiel, alors qu'un hyperonyme impose ses propriétés à ses hyponymes, les propriétés du tout (holonyme) ne sont pas

obligatoirement transmises à ses parties (méronymes). Les co-hyponymes de *fleur* : *tulipe, œillet, pivoine* (cf. *infra*, 2.3) possèdent les propriétés de la classe *fleur* (pétales, parfum, etc.), mais les méronymes de *bicyclette* comme *guidon, roues* n'ont pas en commun les propriétés de la classe *bicyclette*, puisque les parties des noms comptables ne sont pas homogènes (cf. *supra*).

Enfin, les reprises anaphoriques diffèrent ; les méronymes peuvent figurer en position d'anaphores associatives dans le cadre de deux phrases consécutives.

- *Il contemple cet arbre ; le tronc est tout craquelé.*
- *Il prend le stylo ; la plume est cassée.*

Les deux expressions anaphoriques (*le tronc, la plume*) correspondent à la partie et renvoient à des référents identifiés de façon indirecte par l'intermédiaire des antécédents (*cet arbre, le stylo*) auxquels ils sont associés par la relation partie-tout.

Cependant toutes les relations partie-tout ne sont pas susceptibles de donner lieu à une telle anaphore. Les noms de partie de corps des animés en sont, en particulier, exclus ; on ne dira pas : **Le garçon a couru sous la pluie. Les pieds étaient mouillés* mais on utilisera le possessif : *Le garçon a couru sous la pluie. Ses pieds étaient mouillés*.

Les rapports de dépendance méronymique sont variés et complexes : « au lieu d'y avoir une seule relation clairement différenciée, il y a en réalité une famille nombreuse de relations plus ou moins similaires¹ ». Les catégories habituellement retenues pour rendre compte des relations méronymiques sont les suivantes : membre/ensemble (*arbre/forêt, musicien/orchestre*), composant/assemblage (*anse/tasse, roue/voiture*), portion/masse (*part/gâteau, flocon/neige*), matière/objet (*cuir/valise, acier/bicyclette*), activité/phase (*discours/péroraison*), zone/lieu (*oasis/désert*). Mais il y aurait lieu d'établir une typologie plus rigoureuse.

1. Cf. N. Flaux et D. Van de Velde, *Les Noms en français, esquisse de classement*, Paris, Ophrys, 2000.

2. Cf. G. Kleiber, « Massif/comptable et partie/tout », *Verbum*, XIX, 3, 1997.

3. Cf. I. Tamba, « Organisation hiérarchique et relations de dépendance dans le lexique », *L'Information grammaticale* 50, 1991.

1. D. A. Cruse, *Lexical Semantics*, Cambridge University Press, 1986, cité par I. Tamba, « Un puzzle sémantique : le couplage des relations de tout à partie et de partie à tout », *Le Gré des langues* 7, 1994, p. 67.

2. Relations d'équivalence et d'opposition

2.1 La synonymie

2.1.1 Critères

La notion de synonymie est problématique bien qu'elle renvoie à une pratique intuitive largement reconnue.

La synonymie est la relation d'équivalence sémantique entre deux ou plusieurs unités lexicales dont la forme diffère. Les synonymes ont un même signifié et des signifiants différents ; ils s'opposent, en ce sens, aux homonymes définis par un même signifiant et des signifiés différents (cf. chap. 5, 1.2) ; dans les deux cas, il n'y a pas de symétrie entre le plan du signifié et le plan du signifiant. La **synonymie lexicale** (ou synonymie de mots) se manifeste entre mots et/ou syntagmes de même catégorie grammaticale : *pédicure/podologue, policier/agent de police*. La synonymie de phrases porte sur des unités supérieures (phrases, énoncés) ; les phrases représentant les formulations différentes d'un même contenu sémantique constituent des **paraphrases** : par ex. *un chien a attaqué le petit garçon/le petit garçon a été attaqué par un chien*. Sur le plan métalinguistique, on observe dans le dictionnaire la relation de synonymie entre le mot-entrée et la périphrase définitionnelle (cf. chap. 12, 3.1).

La synonymie ne doit pas être confondue avec la dénomination multiple ; les expressions définies (groupes nominaux pourvus d'un déterminant défini et qui renvoient à une entité singulière) telles *la Ville lumière* et *la capitale de la France* désignent le même référent (Paris) sans avoir le même sens.

La synonymie n'est pas, à proprement parler, une identité de sens. Lorsque la forme est différente, les locuteurs s'attendent à une différence de sens : les exemples qui viennent d'être proposés tels *pédicure* et *podologue* peuvent être perçus comme non synonymes. L'identité de sens (ou synonymie absolue) est un leurre. Tout au plus pourrait-on signaler quelques rares cas de synonymes « absolus » dans les lexiques spécialisés (par exemple, en phonétique *voisé* et *sonore* ou en lexicographie *entrée* et *adresse*, bien que ces deux termes n'aient pas la même fréquence). C'est la raison pour laquelle on utilise les termes de

paronymie ou de **quasi-synonymie** qui soulignent le caractère approximatif de la synonymie.

Les grammairiens du xvii^e siècle et du xviii^e siècle avaient déjà noté ce fait. En 1718, l'abbé Girard intitule son ouvrage *La Justesse de la langue française, ou les différentes significations des mots qui passent pour synonymes*. Dumarsais, de son côté, écrit : « S'il y avait des synonymes parfaits, il y aurait deux langues dans une même langue¹. »

Sur le plan théorique, la synonymie lexicale ne se conçoit que dans une théorie de la désignation qui envisage les relations entre les signes et les choses : si un objet a plusieurs noms (*vélo, bicyclette, bécane*), on peut obtenir cet objet en utilisant indifféremment un de ces noms. Il s'agit d'une approche onomasiologique. En revanche, dans une théorie de la signification (perspective sémasiologique), il ne peut y avoir de synonymes dans la mesure où tout mot possède « une partie désignative qu'il peut avoir en commun avec d'autres mots et une partie connotative propre qui ne se retrouve dans aucun autre mot² ». C'est l'approche sémasiologique qui domine dans les études et les définitions de la synonymie.

La synonymie doit être envisagée dans ses rapports avec la polysémie. Lorsque des synonymes concernent des unités polysémiques, la synonymie ne porte que sur une acception. Ainsi le mot *sommet* commute avec *cime* dans les exemples : *le sommet d'un arbre, la cime d'un arbre* (« endroit le plus élevé d'un objet vertical ») mais non pas dans l'exemple : *il est au sommet de sa gloire* où *sommet* (ayant le sens de « point le plus élevé ») aura pour synonyme *apogée*. Les synonymes permettent de distinguer les acceptions d'un mot polysémique ou de dégrouper les homonymes (cf. chap. 5, 2).

Le *Dictionnaire électronique des synonymes* du CRISCO de l'université de Caen, consultable en ligne (www.crisco.unicaen.fr/des/), fournit la liste des synonymes des unités lexicales (42 synonymes pour *sommet*, 14 pour *cime*) en liaison avec les diverses acceptions des unités lexicales. Plus précisément, il propose une représentation géométrique de la synonymie qui visualise la proximité plus ou moins

1. Dumarsais, *Des tropes ou des différents sens*, 1730, Paris, Flammarion, éd. 1988, p. 236.

2. J. Rey-Debove, « La synonymie ou échange de signes comme fondement de la sémantique », *Langages* 128, 1997, p. 95.

grande entre les diverses acceptions de l'unité lexicale et de ses sous-ensembles particuliers de synonymes¹.

2.1.2 La différenciation des synonymes

Les différences entre les paronymes se manifestent sur les trois plans, syntaxique, sémantique et pragmatique, bien souvent de façon simultanée.

Les différences syntaxiques

Les différences d'emploi entre les unités ont pour effet de restreindre la synonymie à un sous-ensemble de contextes communs : deux mots sont synonymes dans certains environnements et non dans d'autres.

Ce phénomène, en grande partie dû à la polysémie, est appelé **synonymie partielle** ou **contextuelle**, le terme *contexte* désignant ici non pas le contexte situationnel mais le contexte linguistique. Pour éviter l'ambiguïté du terme *contexte*, on peut utiliser le terme *cotexte* pour désigner les environnements linguistiques d'une unité.

Les substantifs *écrivain* et *auteur* forment un couple de synonymes lorsqu'ils ont le sens de « personne qui compose des ouvrages littéraires ». Mais ils ne commutent pas dans tous les emplois correspondant à cette acception : *écrivain classique* / *auteur classique* mais *auteur de romans* et non **écrivain de romans* (on trouve cependant sur le web quelques occurrences de *écrivain de romans*).

Le verbe *relever* (cf. chap. 5, 2.1.3) est synonyme de *dépendre* dans la construction syntaxique *relever de* (*la décision relève du ministre*) mais non dans les autres emplois.

Il en résulte que, sur le plan méthodologique, une étude des synonymes doit être d'abord contextuelle. La méthode de l'**analyse distributionnelle** est ici primordiale. Elle consiste à préciser les environnements possibles de chaque mot, du point de vue syntaxique et sémantique. À partir d'une étude des phrases où l'unité apparaît, elle dégage les propriétés distributionnelles qui appartiennent en propre à cette unité en spécifiant les constructions syntaxiques (**sous-catégorisation syntaxique**) et en indiquant la nature sémantique des environnements (traits de **sous-catégorisation sémantique** des substantifs dits

1. Cf. B. Victorri, F. Venant, « Représentation géométrique de la synonymie », *Le Français moderne* 1, 2007.

aussi **traits sémantiques de sélection**, tels que *humain, non humain, concret, abstrait, comptable, massif*). En voici une brève illustration :

Grave et *sérieux* commutent notamment lorsqu'ils sont utilisés avec un nom abstrait : *la situation (la question, l'affaire, le problème, la crise) est grave/est sérieux(se)* ; leur signification est voisine : « qui a une très grande importance et peut avoir des conséquences fâcheuses » ; mais *grave* comprend, en outre, le sème /imminence du préjudice/. En revanche, lorsqu'ils s'appliquent à un nom concret / audible/, la commutation est impossible : *Le son de cet instrument est grave (*sérieux)*.

L'analyse des environnements syntaxiques doit également prendre en compte les collocations. Les **collocations** désignent les groupes de mots fréquemment associés dans le discours et semi-figés. Exemples : on utilise habituellement avec le mot *discours* le verbe *prononcer* (*prononcer un discours*), on dit d'un *récit* qu'il est *fidèle*, on parle d'un *célibataire endurci*, d'une *peur bleue*. Les collocations, restreignant la combinatoire des mots sur le plan syntagmatique, limitent la variation des synonymes.

Ouvrir et *entamer* sont substituables dans certains contextes : *ouvrir/entamer un dialogue, une discussion* mais on dira *ouvrir le débat* plutôt que *entamer le débat, entamer des pourparlers* plutôt que *ouvrir des pourparlers*¹. De même, *gravement* et *sérieusement* diffèrent par leurs collocations : on dira (*un homme*) *gravement blessé* plutôt que (*un homme*) *sérieusement blessé*.

Le *Dictionnaire des combinaisons de mots. Les synonymes en contexte*, (dir. par D. Le Fur) Les Usuels des Robert, 2007 recense autour de 2 600 mots-clés (substantifs) les mots et les expressions avec lesquels ils se combinent : groupements libres, collocations et expressions idiomatiques. Voir aussi le *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*, de I. Mel'čuk, A. Polguère, Bruxelles, de Boeck, 2007.

Ce n'est donc qu'après avoir montré les conditions locales dans lesquelles la synonymie est possible que l'on pourra procéder à une analyse de ce qui les différencie (cf. exercice 5).

1. Cf. S. Verlinde, J. Binon, Th. Selva, « Corpus, collocations, et dictionnaires d'apprentissage », *Langue française* 150, 2006.

Les différences sémantiques

Les différences entre les sémèmes des synonymes portent sur les sèmes spécifiques.

Cime se distingue de *sommet* par le sème /pointu/ ; *pourpre* de *rouge* par le sème /foncé/ ; *lassitude*, *épuisement* désignent le même état physique que *fatigue* mais avec une différence de degré, ce sont des synonymes intensifs ; *rabais* et *remise* présentent une paronymie qui est analysée différemment selon les locuteurs (cf. chap. 3, 2.3.2).

Ces synonymes peuvent se trouver dans les mêmes contextes (le *sommet d'un arbre*, la *cime d'un arbre*, une *étoffe rouge*, une *étoffe pourpre*). La relation synonymique peut parfois se confondre avec la relation d'hyponymie : *pourpre* est à la fois un hyponyme et un synonyme de *rouge*. Mais du point de vue logique, la synonymie se distingue de l'hyponymie : implication bilatérale pour la synonymie, implication unilatérale pour l'hyponymie.

Les différences pragmatiques

Lorsque les mots synonymes ont le même sens dénotatif, ils diffèrent par leurs composantes pragmatiques ou **signifié connotatif** (cf. exercice 5). Le lexique, en effet, est composé de plusieurs sous-systèmes et les locuteurs peuvent choisir entre plusieurs variantes lexicales. C'est le phénomène général de la variation intralinguistique (cf. chap. 1, 2.2). Les différences pragmatiques des synonymes renvoient à différents aspects de la variation lexicale, traitée dans le dictionnaire sous la forme des marques d'usage et de domaine (cf. chap. 11, 1.2 et 1.3 et chap. 12, 2.1 et 2.2) :

- Variations diachroniques : *bru/belle-fille* ; *épatant/super* ; *clinquant/bling-bling* (néol.).
- Variations géographiques (diatopiques) : *wassingue* (nord de la France), *panosse* (dans le Midi), *serpillière*.
- Variations liées aux registres de langue (diastratiques et diaphasiques) : familier ou populaire/standard (*futal/pantalon*, *tronchel tête*), littéraire/standard (*croiséel/fenêtre*), argot (verlan)/langue commune (*meufl femme*, *keufl flic*).
- Variations liées à l'opposition langues de spécialité/langue commune : *rhinite/rhume* ; *préposéfacteur* ; *encéphalite spongiforme bovin/maladie de la vache folle*.
- Connotations, qu'il s'agisse de péjoration (*nègre* pour un être humain par rapport à *noir* ou *black*) ou au contraire d'euphémismes : *longue maladie*

(pour *cancer*), *demandeur d'emploi* (pour *chômeur*), *plan social* (pour *plan de licenciements*), *technicien de surface* (pour *balayeur*), *hôtesse de caisse* (pour *caissière*), *SDF* (pour *sans-abri*).

Comme ces caractérisations sont, par nature, fluctuantes et qu'elles se combinent (*bru* est à la fois *vieilli* et *rural*), on peut préférer avoir recours à la notion de **marque**, empruntée à la phonologie. Un terme marqué possède par rapport au terme non marqué une particularité supplémentaire que l'autre terme ne possède pas ; le terme marqué dans un couple de synonymes est le mot qui n'appartient pas au registre standard : *bru* est marqué par rapport à *belle-fille*. Le terme marqué a une fréquence moins grande que le terme non marqué (*bru* est moins fréquent que *belle-fille*). Cependant l'évaluation de la fréquence doit être rapportée aux types de locuteurs. Ainsi le mot *black* sera le terme marqué dans un groupe donné de locuteurs par rapport à *noir* ; dans un autre groupe c'est le mot *black* (habituellement utilisé) qui sera non marqué et *noir* sera le terme marqué.

En règle générale, les synonymes qui se distinguent par leurs composantes pragmatiques sont substituables dans les mêmes énoncés (par exemple, *Je vous présente ma belle-fille/ma bru, Il a une drôle de têtél/une drôle de tronche*). Mais ils ne sont pas utilisés dans les mêmes situations de discours ni par les mêmes énonciateurs. C'est dire que, quoiqu'ils aient la même référence, ils ne sont pas toujours réellement interchangeables (par exemple, *préposéfacteur*).

L'argot et les lexiques de spécialité – en particulier lorsque la taxinomie scientifique concurrence la taxinomie populaire (par exemple, *centauréel/bleuet*) – procurent des synonymes en abondance. En outre, les facteurs extralinguistiques favorisent dans certains domaines référentiels une affluence de synonymes. Là où s'exerce le poids des tabous, l'interdit stimule en retour une créativité linguistique. Tel est le cas du vocabulaire de la sexualité ; P. Guiraud a relevé près de 600 termes désignant l'organe sexuel masculin et un nombre équivalent pour le sexe de la femme¹. Un autre exemple de créativité linguistique est fourni par la « langue des cités », fortement imprégnée de verlan².

1. P. Guiraud, *Dictionnaire érotique*, Paris, Payot, 1978, p. 13.

2. Cf. B. Seguin, F. Teillard, *Les Céfrans parlent aux Français*, Calmann-Lévy, 1996 ; J.-P. Goudaillier, *Comment tu t'habilles ! Dictionnaire du français contemporain des cités*, Maisonneuve et Larose, 1997 ; *Lexic des cités*, Paris, Fleuve noir, 2007.

La synonymie se distingue des autres relations sémantiques par deux points : la forte dépendance par rapport au contexte syntaxique et ses liens avec les contraintes d'ordre stylistique (l'utilisation de synonymes est recommandée pour éviter l'abus de répétitions).

2.2 L'antonymie

Dans la tradition lexicographique, les antonymes sont définis comme des mots de sens contraire et, comme tels, ils paraissent opposés aux synonymes. Cette vue permet certes de souligner l'analogie de fonctionnement des antonymes avec les synonymes ; synonymie partielle et antonymie partielle participent du même processus puisqu'un terme polysémique a, selon ses acceptions et ses emplois, des antonymes différents ; par exemple pour l'adjectif *clair* : *trouble* (*eau claire*), *foncé* (*couleur claire*), *obscur* (*idée claire*). Mais cette définition de l'antonymie est trop vague. D'une part, l'antonymie implique une dimension de ressemblance entre les termes ; plus précisément, les sèmes antonymiques comportent toujours des sèmes communs : ainsi *frère* et *sœur* partagent les sèmes /être humain/ /né des mêmes parents que la personne considérée/ et s'opposent par le sème /relatif au sexe/. La relation d'antonymie unit donc deux mots de même catégorie grammaticale ayant une partie de leur sémème en commun. D'autre part, la notion d'antonymie n'est pas unitaire ; elle recouvre, de fait, différents types d'oppositions, principalement binaires. Sur la base de critères logiques, on distingue généralement trois types d'antonymes.

2.2.1 Antonymes contradictoires ou complémentaires

Ils sont en relation de disjonction exclusive : la négation de l'un des mots entraîne l'assertion de l'autre, les deux mots ne peuvent être niés simultanément. Soit le couple de mots *vivant/mort*. La proposition *x n'est pas vivant* implique que *x est mort* ; la proposition *x est mort* implique que *x n'est pas vivant* ; tout *x* (à condition qu'il s'agisse d'un être animé) ne saurait être que *vivant* ou *mort* (*À demi mort*, c'est être *vivant*). *Mâle/femelle*, *présent/absent*, *ouvert/fermé*, *marié/célibataire* illustrent cette même relation.

L'application des tests logico-sémantiques tient compte des présuppositions et croyances propres à une culture. Ainsi, la complémentarité de *marié/célibataire* exige que l'on pose qu'il s'agit d'un individu adulte, que le mariage est reconnu dans

la civilisation, que *être veuf* ou *divorcé* équivaut à *marié*, *avoir été marié* et que *conubin* est situé dans la même classe que *célibataire*. Mais le nouveau statut juridique défini par le Pacs (Pacte civil de solidarité) en 1998 a changé les données de telle sorte qu'il y a désormais trois antonymes mutuellement exclusifs : *marié/célibataire/pacsé*.

On voit que les oppositions exprimant l'incompatibilité peuvent comprendre plusieurs termes.

J. Lyons, recensant les oppositions non binaires dans les ensembles à plusieurs éléments composés de termes incompatibles, distingue deux types d'ordre¹ :

- Les ensembles qui sont ordonnés sériellement, c'est-à-dire qui présentent deux éléments extrêmes ; c'est le cas, par exemple, des hiérarchies comme les grades militaires {*général... caporal*}.
- Les ensembles qui sont ordonnés cycliquement, c'est-à-dire qui n'ont pas d'éléments extrêmes, chaque élément de l'ensemble étant ordonné entre deux éléments {*printemps, été, automne, hiver*} ou {*lundi... dimanche*}. Les éléments des ensembles cycliques, mutuellement exclusifs au sein d'un paradigme, peuvent être considérés comme des co-hyponymes (*cf. infra*).

2.2.2 Antonymes contraires ou « gradables »

Ces mots définissent les extrêmes d'une échelle de gradation implicite et autorisent l'existence de degrés intermédiaires. *Grand/petit*, *large/étroit*, *riche/pauvre*, *beau/laid*, *bon/mauvais*, *chaud/froid*, *amour/haine* relèvent de cette catégorie. Deux propriétés les distinguent des contradictoires :

- Ils sont sujets à la gradation (de là le terme de « gradable »). On peut dire *il est passablement (assez, moins, plus, très) grand* mais on ne peut pas dire *il est passablement (assez, moins, plus, très) marié*. La gradation repose en effet sur une comparaison. Celle-ci peut être explicite (formes du comparatif, par exemple : *x est plus grand que y*) ou implicite (*x est grand*). Le choix de ces antonymes par le locuteur dépendra de la norme socioculturelle en vigueur. Selon que l'on habite dans une petite ville ou à Paris, la propriété *grand* ou *petit* ne sera pas attribuée de la même façon à un logement.

1. J. Lyons, *Éléments de sémantique*, Paris, Larousse, 1978, p. 23.

- La négation de l'un des deux mots n'entraîne pas obligatoirement l'affirmation de l'autre. La proposition *x n'est pas grand* n'implique pas nécessairement *x est petit*, de même que *x n'est pas petit* n'implique pas *x est grand* ; *x* peut être ni *petit* ni *grand*, mais *moyen*. Les termes intermédiaires peuvent être lexicalisés : *moyen*, *médiocre* (pour l'opposition *bon/mauvais*) ou ne pas l'être : par exemple, dans le cas de l'opposition *riche/pauvre*, on dira : *il n'est ni riche ni pauvre*. Des séries existent dans lesquelles les antonymes s'opposent terme à terme : *glacial/brûlant*, *froid/chaud*, *frais/tiède*.

Dans les cas de neutralisation (effacement) de l'antonymie, ce sont les termes non marqués de l'opposition qui sont utilisés pour les noms et les adjectifs de dimension (*large, grand, long, haut*, etc.). Ainsi, *largeur, large*, perçus comme non péjoratifs, apparaissent dans les contextes neutralisants de ce type :

- Phrase interrogative : *Quelle est la largeur de ce ruban* et non **Quelle est l'étroitesse de ce ruban ?* ;
- Phrase déclarative : *Je demande la largeur de ce ruban* et non **l'étroitesse de ce ruban* ;
- Expression de mesure : *Ce ruban est large de trois centimètres* et non **étroit de trois centimètres*.

Les antonymes « gradables » sont fort nombreux. Sans doute est-ce la raison pour laquelle les ouvrages de sémantique réservent généralement le terme d'antonymes à cette seule catégorie. Les antonymes lexicaux (comme *intelligent/sot*, *bon/mauvais*) y sont plus nombreux que les antonymes morphologiques (*intelligent/inintelligent*, *hypertension/hypotension*). Pour J. Lyons, les antonymes lexicaux, parce qu'ils ne sont pas reliés du point de vue morphologique, « reflètent une lexicalisation plus poussée des contrastes entre deux pôles¹ ».

2.2.3 Antonymes converses ou réciproques

La relation d'opposition dans ces couples de mots (*mari/femme*, *posséder/appartenir*, *donner/recevoir*, etc.) se révèle par une permutation des actants (participants au procès du verbe) :

1. *Ibid.*, p. 225.

- *Jean est le mari de Julie. Julie est la femme de Jean.*

Mari et *femme* entretiennent une relation converse car la substitution de l'un par l'autre dans un énoncé contraint à inverser les actants (*Jean, Julie*) pour que soit conservée la relation de paraphrase. Ce test de permutation (que l'on peut rapprocher de l'actif et du passif) permet de vérifier la réciprocité de la relation :

- *Jean possède cette maison. Cette maison appartient à Jean.*
- *Jean est le professeur de Marie. Marie est l'élève de Jean.*

Les antonymes converses se rencontrent dans le domaine des relations de parenté et d'échanges sociaux (*père/fils*, *médecin/malade*, *patron/employé*) et dans le domaine des relations temporelles et spatiales (*avant/après*, *devant/derrière*).

Le principe de dichotomisation est un principe essentiel au fonctionnement des langues ; il correspond à des schèmes cognitifs. Il y a une tendance générale de l'homme à catégoriser l'expérience en termes d'oppositions binaires.

Le lexique des mots de l'espace, riche en antonymes (*haut/bas*, *devant/derrière*, *dedans/dehors*, *dessus/dessous*, etc.) est, à cet égard, révélateur ; structurant notre perception du monde, il donne lieu à une évaluation positive et négative, à la source de nombreuses métaphores¹. Ainsi les mots de la dimension verticale (tels *haut/bas*, *sommet/base*) sont particulièrement aptes à l'expression métaphorique de la hiérarchie sociale et de la conflictualité sociale².

2.3 La co-hyponymie

Sont co-hyponymes les mots partageant une même relation hiérarchique avec un hyperonyme (cf. 1.1). La co-hyponymie est donc une relation sémantique triangulaire³. *Tulipe, œillet* sont co-hyponymes de *fleur* ; *cèpe, morille, girolle* sont co-hyponymes de *champignon* ; *printemps, été* sont co-hyponymes

1. Cf. G. Lakoff et M. Johnson, *Les Métaphores dans la vie quotidienne*, chap. 4 ; cf. pour la présentation de cette approche chap. 5, 3.5.6.

2. Cf. S. Rémi-Giraud, « De haut en bas et de la base au sommet. Les parcours de la verticalité » in *Mots de l'espace et conflictualité sociale*, dir. par P. Bacot et S. Rémi-Giraud, Paris, L'Harmattan, 2007.

3. J. Rey-Debove, « La synonymie ou les échanges de signes comme fondement de la sémantique », *La Synonymie*. *Langages* 128, 1997, p. 92.

de *saison*. Les co-hyponymes sont des unités de même rang car ils sont situés au même niveau de la relation qui les rattache à l'hyperonyme. Ils diffèrent entre eux par un ou plusieurs traits spécifiques. Contrairement à ce qui se passe pour la relation d'antonymie fondée principalement sur une opposition binaire, la négation d'un des cohyponymes n'implique pas nécessairement l'affirmation d'un autre co-hyponyme, le choix restant ouvert : *si x n'est pas une tulipe, x peut être un œillet, une pivoine...* Les co-hyponymes sont mutuellement exclusifs : *une fleur est ou une tulipe ou un œillet ou une pivoine ; une saison est ou le printemps ou l'été.*

Les co-hyponymes peuvent entretenir entre eux des relations de synonymie ou des relations d'antonymie. *Guilleret/jovial*, que l'on peut considérer comme des cohyponymes de *gai* (*jovial* signifiant une gaieté franche, *guilleret* une gaieté vive), peuvent passer pour synonymes ; *acheter/voler*, co-hyponymes de *se procurer*, sont antonymes.

3. Exercice

Exercice 5

Étudier la paronymie des couples suivants en caractérisant les contextes dans lesquels ils peuvent ou non commuter et en spécifiant les différences qui les séparent :

- *filie mère/mère célibataire*
- *battre/frapper*

Corrigé

Fille mère/mère célibataire

Les noms composés *filie mère* et *mère célibataire* sont, à première vue, substituables dans un grand nombre d'énoncés. Ils ont le même signifié dénotatif (« femme non mariée qui a un ou plusieurs enfants ») et diffèrent par leur signifié connotatif : *filie mère* est péjoratif, *mère célibataire* est le terme plus neutre. La structure morphologique manifeste ces différences : dans *mère célibataire*, le déterminé est *mère*, mot à connotations positives, dans *filie mère* le déterminé est *filie*, terme à connotations péjoratives. *Mère célibataire* (qui date

de 1961) a remplacé *filie mère*, terme vieilli qui n'est plus guère employé de nos jours, ce qui limite de fait la substituabilité.

Toutefois, en raison de l'évolution des mœurs et de la diversification des structures familiales, le terme *mère célibataire* tend à disparaître de l'usage au profit de l'expression *famille monoparentale* apparue dans les années 1980 et désignant la famille où il n'y a qu'un seul parent, père ou mère.

Battre/frapper

Les verbes *battre* et *frapper* sont polysémiques. Ils présentent une proximité sémantique dans l'intersection de leurs polysémies respectives en relation avec leur distribution.

a) Ils sont substituables dans deux types d'environnements :

- Premier contexte : Humain + Verbe + Animé.

- *Le père bat son enfant. L'enfant bat son chien.*
- *Le père frappe son enfant. L'enfant frappe son chien.*

La différence entre *battre* et *frapper* est sémantique : *battre* « donner des coups à plusieurs reprises », *frapper* : « donner un ou plusieurs coups ». *Battre* présente obligatoirement le sème /plusieurs coups/ (cf. le nom de l'association *SOS Femmes battues*).

Au passif, si la commutation reste possible (*L'enfant a été battu par son père, l'enfant a été frappé par son père*), *frapper* est ambigu (il peut avoir le sens de « impressionner »).

Frapper admet, contrairement à *battre*, un complément qui précise le lieu du coup.

- *Le père frappe son enfant au visage.*
- **Le père bat son enfant au visage.*

- Deuxième contexte (fortement restreint) : le sujet désigne la pluie, la grêle..., l'objet est concret.

- *La pluie battait les vitres.*
- *La pluie frappait les vitres.*

Battre et *frapper* partagent les sèmes /toucher/, /avec une certaine force/, *battre* peut se différencier par le sème déjà observé /répétition de l'action/, *frapper* par le sème /rapidité de l'action/.

b) *Battre* et *frapper* ne sont pas substituables l'un à l'autre. Vu la polysémie des deux verbes, ces contextes sont trop nombreux pour pouvoir être décrits ; on proposera deux exemples concernant *battre*.

Emploi intransitif : le sujet désigne une partie du corps ou un objet susceptible de mouvements, *battre* signifie : « être animé de mouvements ».

- *Le pouls bat lentement.*
- **Le pouls frappe lentement.*

Emploi transitif : le sujet est humain et l'objet, humain, désigne un adversaire ; *battre* a le sens de « l'emporter sur ».

- *En 2002, la droite a battu la gauche aux élections législatives.*
- **En 2002, la droite a frappé la gauche aux élections législatives.*

c) Emplois figés : *battre* et *frapper* entrent dans de nombreux syntagmes figés dans lesquels toute commutation est bloquée (cf. chap. 9, 1.3).

- *Battre la campagne.*
- **Frapper la campagne.*

Toutefois la commutation peut être observée dans un cas :

- *L'enfant bat des mains.*
- *L'enfant frappe des mains (frappe dans ses mains).*

Battre équivaut à « applaudir », *frapper* a une extension plus grande (accompagner un rythme dans le but ou non d'applaudir).

Les deux locutions techniques *battre monnaie* et *frapper la monnaie* se distinguent par l'emploi de l'article et par le sens, *frapper* précise le processus de fabrication des pièces de monnaie (« reproduire des empreintes sur les deux faces d'une pièce de monnaie »).

Conclusion

Les synonymes lexicaux doivent être étudiés en contexte. Aux différences distributionnelles s'ajoute la codification des collocations qui limite considérablement la synonymie réelle dans l'usage.