

UNIVERSITÉS FRANCOPHONES

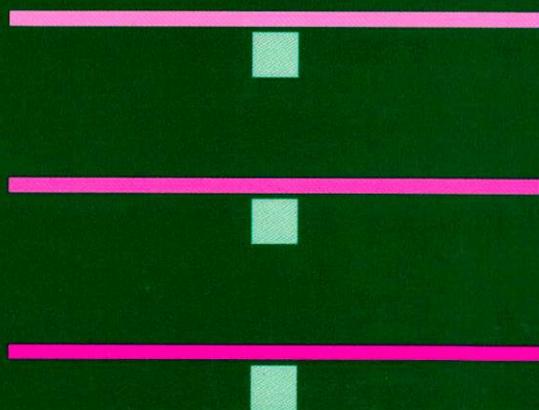


# LES DICTIONNAIRES BILINGUES

Henri Béjoint / Philippe Thoiron

Avec des contributions

de Claude Boisson, Teresa Cabré, André Clas, Cristina Gelpí,  
Valerie Grundy, Frank Knowles, Marcel Lemmens, Brigitte Lépinette,  
Eliott Macklovitch, Carla Marengo, Roda Roberts et de Thomas Szende



Aupelf-Uref - Editions Duculot



LES  
DICTIONNAIRES  
BILINGUES

# Champs Linguistiques

Collection dirigée par

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

## *Recherches*

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

Brès J., *La narrativité.*

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

Gosselin L., *Sémantique de la temporalité en français.*

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

Jonasson K., *Le nom propre.*

Kleiber G., *Anaphores et pronoms.*

Léard J.-M., *Les gallicismes.*

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

## *Manuels*

Chiss J.-L., Puech C., *Fondations de la linguistique. Études d'histoire et d'épistémologie.*

Furukawa N., *Grammaire de la prédication seconde. Forme, sens et contraintes.*

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

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

## *Recueils*

Francard M., Latin D. (Éds), *Le régionalisme lexical.*

Fuchs C., , *La place du sujet en français contemporain.*

Kleiber G., Riegel M. (Éds), *Les formes du sens.*

Thoiron P., Bejoint H. (Éds), *Les dictionnaires bilingues.*

UNIVERSITÉS FRANCOPHONES

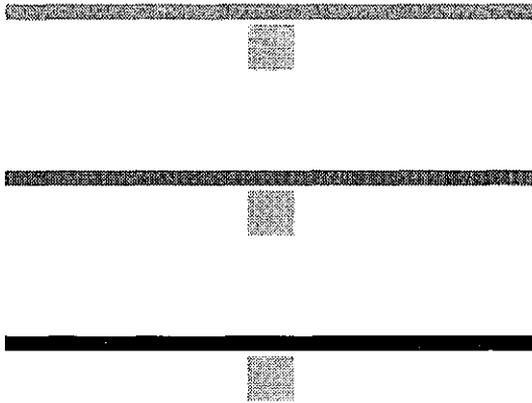


# LES DICTIONNAIRES BILINGUES

Henri Béjoint / Philippe Thoiron

Avec des contributions

de Claude Boisson, Teresa Cabré, André Clas, Cristina Gelpí,  
Valerie Grundy, Frank Knowles, Marcel Lemmens, Brigitte Lépinette,  
Eliott Macklovitch, Carla Marelllo, Roda Roberts et de Thomas Szende



**Aupelf-Uref - Editions Duculot**

© Duculot s.a. 1996  
Fond Jean-Pâques, 4 B - 1348 Louvain-la-Neuve

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

Printed in Belgium

D 1996/0035/20

ISBN 2-8011-1138-4

# INTRODUCTION

Henri BÉJOINT et Philippe THOIRON

Depuis le début des années 60, la métalexigraphie a connu une remarquable expansion, en France comme dans le reste du monde. Dire qu'il n'existait aucune activité métalexigraphique avant cette période serait exagéré, mais il est vrai que les spécialistes s'accordent à voir dans le congrès de Bloomington (1960; voir Householder et Saporta : 1962), pour la langue anglaise par exemple, ou dans la thèse de Bernard Quemada (1968) pour la France, des dates importantes qui marquent le début d'une nouvelle ère, qui a été, et est encore, marquée par un grand nombre de congrès, colloques et publications de tous ordres.

En France, et dans les autres pays occidentaux, on a surtout écrit sur les dictionnaires monolingues généraux : les travaux les plus marquants publiés en français, c'est-à-dire les ouvrages de Dubois (1971), Rey-Debove (1971), Rey (1977), Collignon et Glatigny (1978), etc. ne font guère mention des autres types de dictionnaires. Si des progrès sensibles ont été accomplis dans la rédaction des monolingues, c'est peut-être à cause de cette attention particulière que leur ont accordée les spécialistes – ce qui voudrait dire que l'activité métalexigraphique a une influence bénéfique sur la pratique lexicographique.

Deux types de dictionnaires importants ont été relativement délaissés dans ces recherches : les dictionnaires de langues de spécialité, et les dictionnaires bilingues<sup>1</sup> – les deux catégories se recouvrant d'ailleurs en grande partie<sup>2</sup>. Dans ces deux domaines, les dictionnaires publiés jusqu'à une date récente, et dans une certaine mesure encore actuellement, ne sont guère différents, du point de vue méthodologique, de ceux qui étaient publiés il y a cinquante ans. Cela est d'autant plus surprenant que les dictionnaires terminologiques et les dictionnaires bilingues sont sans doute les deux types dont la production est la plus importante. Tout se passe – se passait – comme si ces ouvrages étaient vus comme des outils, indispensables certes, mais pas plus dignes d'une réflexion méthodologique qu'un tournevis ou un tire-bouchon.

Ce décalage entre l'importance quantitative et commerciale (somme toute, sociale) de la production et la relative pauvreté de la théorisation justifierait à lui

---

1 Il y en a d'autres : les dictionnaires pour enfants, par exemple, qui offrent pourtant de vastes espaces au chercheur, dans la mesure où leur portée ouvertement pédagogique leur confère un statut quasi prototypique.

2 Il n'y a aucun ouvrage en langue française qui fasse autorité dans le domaine des bilingues actuellement. Pour la langue anglaise, il faut signaler Householder and Saporta (1962), Zgusta (1971; publié en anglais, mais abondant de nombreuses autres langues) et Svensén (1993; d'abord publié en suédois); pour l'italien, Marcellio (1989), etc. Evidemment, la bibliographie des travaux plus courts (articles, chapitres de livres, conférences, interventions dans des congrès, etc.) est beaucoup plus étoffée.

seul la parution du présent volume. Mais il semble bien qu'un certain nombre de faits convergent désormais pour conférer à ces deux types de dictionnaires une importance encore plus grande, parmi lesquels l'évidente intensification des échanges de tous ordres entre les différentes communautés humaines. Si la francophonie, en particulier, pouvait il y a peu encore survivre et prospérer dans son splendide isolement, une telle attitude devient de plus en plus difficile à maintenir, surtout dans le domaine des publications scientifiques.

Il y a, il est vrai, ceux qui considèrent que la communication interlinguistique des spécialistes peut, ou doit, être assurée par une seule langue, et pour qui l'anglais, dans ce que l'on appelle parfois sa variante internationale, semble tout désigné pour jouer ce rôle : la communication techno-scientifique assurée par le seul vecteur de l'anglais conduirait, suppose-t-on, à des économies de temps et de moyens. Beaucoup de scientifiques francophones font ainsi observer que l'édition dans les domaines spécialisés se faisant essentiellement en langue anglaise, il est vain de s'obstiner à rédiger dans une autre langue. Leur attitude ne milite guère en faveur de la lexicographie bilingue, et encore moins en faveur de la lexicographie bilingue spécialisée : on aboutirait, si cette logique était conduite à son terme, à un paradoxe, puisque la nécessaire communication entre communautés de langues différentes conduirait à l'uniformisation linguistique et donc à l'affaiblissement de chaque langue.

Mais le monolinguisme institué n'est pas accepté par tous. C'est ainsi que l'Union européenne souhaite préserver les richesses de son plurilinguisme et se préoccupe de mettre en œuvre les initiatives susceptibles de l'aider à atteindre cet objectif.<sup>3</sup> Il semble bien que l'Européen d'aujourd'hui prenne de plus en plus conscience qu'il est vain de chercher à imposer sa langue, quelle qu'elle soit, qu'il faut accepter que chaque langue n'est qu'une parmi les autres, et qu'il importe donc de favoriser toutes les activités qui consistent à jeter des passerelles entre les langues, en particulier bien sûr la traduction. Il convient alors de donner à chaque langue les moyens de son adéquation aux exigences de la communication d'aujourd'hui, notamment dans les domaines scientifiques et techniques, de la doter de terminologies appropriées et de mettre en place des outils lexicographiques et terminographiques adéquats pour que les transferts entre langues soient plus faciles et plus efficaces.

L'élément déterminant dans le renouveau de l'activité lexicographique, ce qui a permis de réaliser, et d'envisager, des progrès dans la façon de réaliser des dictionnaires et dans la qualité de leur contenu, c'est, bien sûr, le développement de l'informatique. Pour ce qui concerne la lexicographie bilingue, l'informatique a permis récemment l'utilisation de corpus (conformément à ce qui se pratique depuis assez longtemps pour les monolingues) et la gestion du contenu, en particulier l'harmonisation des deux parties du dictionnaire. A terme, on peut

3 C'était le thème du congrès tenu à Besançon en juin 1995 : *Plurilinguisme : Quelles initiatives dans l'Union européenne ?*

espérer qu'elle permettra d'exaucer les vœux depuis longtemps formulés par les métalexigraphes de voir fabriquer au moins quatre dictionnaires par couple de langues, selon la LI de l'utilisateur et selon l'utilisation visée, encodage ou décodage. Ce nombre peut d'ailleurs être augmenté, si l'on prend en compte d'autres variables, comme l'âge ou la compétence linguistique des utilisateurs. Ces dictionnaires différents peuvent désormais, grâce à l'informatique, être tirés d'une base lexicographique commune.

Les auteurs des différents chapitres de ce livre sont toutes et tous des métalexigraphes appartenant aux deux communautés dont la coopération conditionne les progrès de la discipline : celle des universitaires, qui peuvent prendre du recul par rapport au quotidien et échafauder des théories qui ne s'embarassent ni de détails de réalisation ni de considérations commerciales, et celle des lexicographes, qui bénéficient d'une expérience vécue de la compilation des dictionnaires sous tous ses aspects et peuvent donc mesurer la portée des théories à l'aune de la faisabilité. La plupart des auteurs sont spécialistes de linguistique – signe de l'importance (quoique non exclusive) de cette discipline dans le domaine de la lexicographie.

L'ouvrage a été conçu comme une initiation progressive au monde des dictionnaires bilingues. Il va du passé au futur, en passant par le présent et ses diverses facettes. Conçu pour être un ouvrage de référence, il met l'accent sur ce qui se fait, et ouvre quelques perspectives sur l'avenir. Manuel d'enseignement plutôt qu'ouvrage réservé aux spécialistes, il n'a pas pour ambition de rendre compte de recherches de pointe, mais propose plutôt des bilans de ce qui a été réalisé et des considérations sur ce qui est envisageable.

Le dictionnaire bilingue sera d'abord replacé dans son contexte historique et social, puis différents aspects de sa préparation seront examinés, en partant de la compilation de la liste des «mots» vedettes (en gros, la macrostructure) pour aller vers le contenu des articles (la microstructure). Une importance particulière sera accordée au dictionnaire bilingue de langues de spécialité, pour les raisons évoquées ci-dessus. Enfin, plusieurs chapitres seront consacrés au rôle de l'informatique dans la fabrication des dictionnaires bilingues de toutes sortes.

Le premier chapitre (**Claude Boisson**) examine différentes traditions lexicographiques pour tenter de discerner lequel, du monolingue ou du bilingue, a précédé l'autre dans l'histoire. Les conclusions bousculent les idées reçues, puisque la prétendue antériorité du bilingue n'est avérée que dans quelques civilisations, dont celles de l'Europe occidentale. L'auteur revient également sur l'idée, sans doute plus révolutionnaire qu'il n'y paraît, des «dictionnaires oraux», c'est-à-dire des listes de mots, monolingues ou bilingues, qui existent dans le répertoire des «conteurs» de certaines civilisations sans écriture.<sup>4</sup> Cette décou-

---

4 L'idée avait été introduite, semble-t-il pour la première fois en métalexigraphie, dans Boisson, Kirtchuk et Béjoint : 1991.

verte l'amène à réfléchir sur l'importance de la rupture épistémologique de l'apparition de l'écriture, que d'aucuns ont trop vite assimilée à un bouleversement des modes de pensée.

Un regard diachronique de cette sorte pose inévitablement le problème des frontières entre les genres : qu'est-ce qui différencie le dictionnaire de l'encyclopédie ? Qu'est-ce qui différencie le dictionnaire monolingue du bilingue ? A partir de quel moment, chronologiquement et méthodologiquement, peut-on dire qu'une liste de mots devient un dictionnaire ? A partir de quel moment peut-on dire qu'un dictionnaire est bilingue plutôt que «bi-dialectal» ou «bi-variétal» ? On verra que le dernier chapitre fait suite au premier, d'une certaine manière, puisqu'il traite de la place du dictionnaire bilingue dans l'affirmation identitaire d'une culture et d'une langue, en l'occurrence le catalan.

Ces questions de définitions sont reprises dans le deuxième chapitre, dont l'auteur (**Carla Marello**) a accepté de se livrer une fois de plus au périlleux exercice de la typologie des dictionnaires. Elle y présente les différents types de dictionnaires bilingues qui ont été réalisés dans les différentes sociétés. On y verra que le monde des dictionnaires bilingues évolue, malgré la résistance naturelle des utilisateurs aux innovations. Les deux premiers chapitres posent également la question du rôle de langue pivot que joue le latin dans les premiers dictionnaires bilingues de nos civilisations occidentales. On ne peut s'empêcher de se demander ce qu'on pourrait tirer d'une telle conception si on la transposait à la situation actuelle.

Un certain nombre de chapitres traitent ensuite des aspects macrostructurels et microstructurels des dictionnaires bilingues de langue générale. La conception d'ensemble de la macrostructure de ces dictionnaires fait l'objet d'un ré-examen dans le chapitre de **Brigitte Lépinette**. Elle y pose la question du rôle des modèles linguistiques dans l'élaboration des dictionnaires : à quoi peuvent servir, par exemple, la théorie Sens → Texte d'Igor Mel'čuk ou les travaux de Maurice Gross dans le cadre du L.A.D.L. lorsqu'il s'agit de confectionner un dictionnaire bilingue de langue générale ? L'auteur examine en particulier les conséquences qu'aurait une utilisation systématique du comportement syntaxique des éléments lexicaux pour distinguer entre les différentes acceptions. Ayant constaté que les dictionnaires conventionnels utilisent préférentiellement des critères de nature sémantique et que cette utilisation n'est pas basée sur une analyse aussi homogène qu'on le souhaiterait des items qui constituent la macrostructure, elle propose un type de dictionnaire bilingue dont le principe d'organisation, homogène, serait fondamentalement syntaxique.

Les aspects microstructurels des dictionnaires bilingues généraux sont abordés ensuite. La recherche des équivalents, fondamentale en lexicographie bilingue, constitue un difficile exercice d'équilibre auquel se livrent, qu'ils le veuillent ou non, et quelles que soient les difficultés, tous les auteurs de dictionnaires (**Thomas Szende**) : il s'agit de mettre en regard deux «suites syntagmatiques» issues de deux langues différentes, à partir d'une équivalence qui tombe plus ou

moins sous le sens mais qui résiste à la démonstration. Il est permis en effet de se demander comment on peut montrer que deux unités linguistiques sont équivalentes, si l'on ne sait pas décrire le sens autrement que par des moyens linguistiques (sauf à avoir recours à une vision bloomfieldienne du sens que la communauté des linguistes n'accepte plus, et qui n'a pas d'opérationnalité en lexicographie), puisqu'il s'agit, au fond, de mettre en regard des définitions, forcément en partie subjectives, au moins dans leur formulation, et dans deux langues différentes.

Ce chapitre soulève également la vaste question de la nature de l'information contenue dans un bilingue. S'agit-il d'un simple «index» de formes qui ne peut fonctionner qu'avec le concours de deux dictionnaires monolingues implicites (un pour chacune des langues), qui contiendraient le sens ? Peut-on dire en conséquence que la sémantique est absente du bilingue ? La distinction entre monolingue et bilingue est en tout cas moins claire que ce que l'on pourrait penser, à la fois historiquement (voir le premier chapitre) et en synchronie (voir le second). La place de la définition dans les dictionnaires monolingues est souvent – le fait est bien connu – occupée par des synonymes, et d'autre part les dictionnaires bilingues de langue générale contiennent un nombre variable de gloses tenant lieu d'équivalences, dans les articles consacrés aux particularismes culturels, aux mots de la cuisine, du costume, aux noms d'espèces animales ou végétales autochtones, d'artefacts non exportés, et également pour les idiotismes et phraséologismes de tout poil.

Il était indispensable de consacrer un chapitre à la place de la phraséologie et des collocations dans les dictionnaires bilingues généraux, et c'est **Roda Roberts** qui a accepté de s'en charger, forte de l'expérience acquise au cours des dernières années dans la rédaction du dictionnaire bilingue canadien réalisé conjointement par les universités de Montréal et d'Ottawa. L'importance de l'inclusion de ces suites syntagmatiques est apparue depuis peu à beaucoup de spécialistes, si l'on en juge par le nombre de publications récentes dans la littérature métalexigraphique. Curieusement, les dictionnaires anciens, ceux du XVI<sup>e</sup> siècle en Angleterre par exemple, étaient d'une certaine manière en avance sur nous, puisqu'ils avaient tendance à traiter comme des entrées des «morceaux de discours» aussi bien que des mots graphiques; mais l'habitude a ensuite été abandonnée, au nom d'une fidélité à l'idée qu'on se faisait de l'unité lexicale en tant qu'élément constitutif de la langue et représentatif de l'univers.

L'importance des expressions idiomatiques dans les dictionnaires bilingues ne fait aucun doute si l'on accepte l'idée qu'une langue utilise deux types de suites syntagmatiques : certaines sont structurales et d'autres idiomatiques. Alors que les premières sont assez faciles à comprendre ou à composer à partir des structures de la langue, les secondes, dont le sens ne se déduit pas simplement de la somme des sens de leurs éléments constitutifs, posent bien des difficultés aux locuteurs non natifs, qui doivent les apprendre consciemment comme les vocables «ordinaires». Pour faciliter ce processus d'apprentissage, il est indispensable que ce

type d'unités figure dans les dictionnaires bilingues. On s'accorde maintenant sur ces nécessités, mais il n'est pas encore possible de faire l'unanimité sur les solutions à proposer pour résoudre tous les problèmes posés. La définition et la typologie des unités idiomatiques étant mal établies, leur repérage reste problématique. Parce que leur statut lexicographique est encore vague, il est parfois difficile de les trier et de mettre au point des critères d'inclusion dans un dictionnaire bilingue.<sup>5</sup> L'ajout de commentaires métalinguistiques et un système rigoureux de classement dans les rubriques du dictionnaire bilingue peuvent constituer des moyens supplémentaires pour éclairer les usagers.

La question de l'inclusion d'informations d'ordre grammatical dans les dictionnaires bilingues est évoquée par **Marcel Lemmens** d'une manière originale, puisqu'il fait part des résultats d'une enquête menée auprès d'un certain nombre d'utilisateurs de dictionnaires. Sa contribution soulève le problème des rapports entre ce que les dictionnaires proposent, ce que les utilisateurs en attendent, et ce qu'ils font réellement, trois choses dont l'article fameux de Robert Galisson (1983) avait montré qu'elles pouvaient être sensiblement différentes. Une telle enquête constitue une amorce de réponse à une question moins triviale qu'il n'y paraît : à quoi sert le dictionnaire ?

La nature des exemples, les critères qui doivent présider à leur choix, leur authenticité, ont toujours préoccupé les lexicographes. Ces problèmes revêtent une importance particulière pour les dictionnaires bilingues, dans la mesure où il s'agit, entre autres choses, de fournir aux utilisateurs des modèles productifs plus que des exemples de bel usage. **Xavier Blanco**, en s'appuyant sur la quasi totalité des dictionnaires généraux français-espagnol et espagnol-français, montre que, bien que l'utilisation d'exemples y soient prévue et que l'on en vante la multifonctionnalité et la capacité illustrative, l'exemple n'est pas toujours conçu de manière à remplir la tâche qui lui est assignée. L'auteur propose d'une part l'établissement de critères susceptibles d'orienter la réexploitation de corpus d'exemples empruntés aux dictionnaires existants, et d'autre part la mise en place, grâce à l'outil informatique, de clés d'accès qui permettraient d'optimiser la rentabilité d'un tel corpus au sein même du dictionnaire qui le contient.

Les questions qui se posent au rédacteur de dictionnaires bilingues (ou multilingues) de langues de spécialités sont abordées par **André Clas**. Parmi celles-ci figure, encore une fois mais ce n'est pas surprenant, celle de la détermination des équivalences. Dans beaucoup de dictionnaires de langues de spécialité, on semble avoir admis le principe de l'identité conceptuelle, hérité du classicisme terminologique. Or même en langue de spécialité les concepts de la langue source et ceux de la langue cible peuvent être voisins sans être identiques, et il serait souvent plus prudent de parler de concepts homologues plutôt

5 Il est intéressant de noter à ce sujet qu'une meilleure prise en compte des finalités des bilingues, qui conduirait à la fabrication d'outils distincts en fonction de ces différents objectifs réglerait des problèmes difficiles, tels que l'inclusion de collocations ou d'expressions idiomatiques transparentes en décodage.

qu'identiques – tout en se gardant évidemment de nier qu'il puisse exister des identités. En acceptant ces éventuelles différences, on en vient à davantage de précautions dans la gestion des équivalences, et à mettre en œuvre lorsque c'est nécessaire l'idée de faux-amis terminologiques. La simple mise en équivalence d'un terme en langue source et d'un terme en langue cible peut se révéler dangereuse pour l'utilisateur dans certains cas, surtout s'il n'est pas un spécialiste averti du domaine de spécialité en question.

Mais quelles que soient les différences d'appréhension de la réalité (d'ailleurs variables selon les domaines et selon les paires de langues mises en regard), les lexicographes sont contraints, bon gré mal gré, de proposer des équivalences. Dans certaines activités humaines, telles que le droit, la politique ou la structuration sociale, les équivalences ont été éprouvées par une longue pratique, et les auteurs de dictionnaires pourront avertir le lecteur des divergences éventuelles de vision du monde ou d'usage linguistique en fournissant des explications, des gloses, des commentaires. Ce type d'information ne s'inscrit pas dans le cadre de la définition traditionnelle et donnera donc lieu à la création de rubriques inédites. Dans les domaines technoscientifiques, les équivalents sont plus souvent raisonnablement opératoires, même s'ils ne sont pas toujours parfaitement exacts, et il faudra souvent envisager de les compléter par des informations de nature encyclopédique.

Comment montrer que deux termes peuvent être considérés comme équivalents ? Il s'agit toujours, implicitement ou explicitement, de vérifier l'équivalence des définitions dans les deux langues, mais une telle opération est loin d'être simple si l'on veut y apporter une certaine rigueur. Des outils terminographiques dans lesquels on ne pose pas en principe l'identité transculturelle et translinguistique des concepts, et qui proposent donc une définition pour chaque terme et pour chaque langue, systématiquement, peuvent aider l'utilisateur à évaluer lui-même le degré d'équivalence pour les solutions qu'on lui propose. Il existe d'ores et déjà des modèles pour ce type de dictionnaires, par exemple dans des bases de données terminologiques bilingues. Dans les dictionnaires existants, les définitions ne sont le plus souvent rédigées que dans la langue source, solution qui ne résoud qu'une partie du problème. On commence toutefois à trouver des ouvrages dans lesquels les définitions sont données à la fois en langue source et en langue cible, comme dans *TERMIUM*. Là encore, l'outil informatique modifie les fondements mêmes du concept de dictionnaire bilingue spécialisé, qui se rapproche de la base de données et n'a plus guère de rapport avec la liste d'équivalents imprimée sur papier.<sup>6</sup>

6 On observe une évolution semblable dans le domaine des dictionnaires bilingues de langue générale. Certains dictionnaires «papier», d'ores et déjà, proposent à la fois des définitions et des équivalences, ce qui permet de faire passer de l'information qui sans cela ne serait pas transmise, ou serait transmise de manière ambiguë pour l'utilisateur étranger à la culture de la langue cible. Ainsi, la définition de «tiro-bouchon» («... qui débouche des bouteilles de vin») laisse-t-elle nombre de questions en suspens : tous les vins ? Même le champagne ? On touche ici aux frontières du linguistique et du pragmatique. Si ce n'est

Dans ce type de dictionnaire, la question de la synonymie se pose de manière particulière, puisqu'en terminologie elle est basée sur l'identité des référents. Il peut exister plusieurs termes renvoyant au même référent, et donc à la même définition. Le lexicographe devra donc sélectionner celui des synonymes qui fera l'objet d'un article complet, avec sa définition, les autres n'ayant qu'un article réduit soit à un renvoi soit à une liste d'équivalents en langue cible. C'est le critère de fréquence qui sera le plus naturellement utilisé, mais on sait bien à quel point ce critère est parfois difficile à mettre en œuvre.

Le problème du continuum entre terme simple, terme composé, assemblage de cooccurrents et assemblage libre est responsable de beaucoup de difficultés pour les terminographes, surtout s'ils préparent un dictionnaire «papier». Il est parfois difficile de décider si l'on a affaire à un terme pluriélémentaire (qui justifierait une entrée correspondant à l'ensemble) ou à une collocation (qui donnerait lieu à une entrée correspondant à une partie seulement de l'ensemble). On peut se demander par exemple s'il faut traiter **variable aléatoire** comme une vedette ou proposer une rubrique pour **variable** et indiquer **aléatoire** parmi les collocatifs (voir ci-dessous). Outre le problème théorique, les implications matérielles, et donc financières, ne sont pas négligeables, les dernières prenant même parfois le pas sur le premier (voir le chapitre de Valerie Grundy). Avec des dictionnaires bilingues informatisés l'environnement est totalement différent et la question ne se pose plus vraiment puisque les relations entre un générique et son ou ses spécifiques peuvent être gérées dans le système sous forme de mise en liaison instantanée. Observons que, fort commodément, le problème théorique n'a plus à être posé comme préalable à la mise en place d'une solution pratique. Le processus a été renversé.

Un autre point sur lequel on voit actuellement des progrès se dessiner en matière de dictionnaires bilingues de langues de spécialité est celui des collocations. On commence à trouver des dictionnaires bilingues de langues de spécialités «traditionnels» qui incluent les principales collocations. Il est vrai que leur traitement pose de nombreux problèmes. Quelles collocations faut-il retenir ? Faut-il inclure celles qui ont une très forte «prévisibilité», par exemple **dimension** comme collocatif de *calculer, estimer, évaluer, mesurer* ? Si les termes sont définis dans l'article, on peut espérer que l'utilisateur pourra inférer de ces définitions les propriétés fondamentales des concepts en cause, et donc les types de collocatifs au moins potentiels. Mais la difficulté est que la collocabilité n'est pas un phénomène seulement sémantique : elle est aussi lexicale, et par conséquent tout système visant à amener l'utilisateur à prévoir dans toute la mesure du

---

(suite note 6)

guère problématique en langue de spécialité, il est par contre permis de se poser des questions pour ce qui est des dictionnaires de langue générale : faut-il qu'un dictionnaire bilingue s'en tienne aux mots et à leur comportement, ou bien doit-il aussi parler des choses ? Question souvent débattue pour les dictionnaires monolingues...

possible quels sont les types de collocations qui fonctionnent ne peut pas être totalement opérationnel s'il est fondé sur la seule sémantique.

Dans les langues de spécialité comme pour la langue générale, il est très important de mettre au point des méthodes simples et efficaces (surtout du point de vue de l'utilisateur) permettant de classer les collocatifs à l'intérieur d'une même rubrique. Même si le système le plus simple, et sans doute encore le plus utilisé, est tout simplement alphabétique, les méthodes de classement selon des critères notionnels sont infiniment plus intéressantes – et infiniment plus problématiques. Les travaux de Benson, Benson et Ison, de Cohen, et de Pavel et Boileau entre autres, ont le mérite d'exister, et ceux d'Igor Mel'čuk constituent sans doute l'une des bases solides sur lesquelles pourraient être fondés des modèles encore plus systématiques. A titre d'exemple des inévitables bizarreries créées par un système purement formel – demandé d'ailleurs souvent par les utilisateurs eux-mêmes – on verra que, dans un dictionnaire (Pavel et Boileau : 1994), deux antonymes, *casser* et *coller*, se suivent dans une rubrique (**agrégat**), tandis que dans une autre (**courbe**) *dessiner* et *tracer* sont séparés par plusieurs autres collocatifs.

Le problème de l'organisation des informations phraséologiques se pose en langue de spécialité comme dans le lexique général (voir aussi le chapitre de Roda Roberts). Pour éviter de devoir répéter toute l'information concernant la phraséologie d'un terme dans les articles consacrés à chacun de ses synonymes, la solution économique la plus souvent retenue consiste à utiliser un système de renvois. Mais encore faut-il être sûr que les propriétés phraséologiques de toute la série synonymique sont bien les mêmes. Le contraire n'aurait rien de surprenant puisque – nous l'avons dit – la synonymie terminologique est référentielle; le danger est alors patent, puisque le traitement des questions de phraséologie nous entraîne sur le terrain du discours, où la synonymie est d'une autre nature.

Enfin, l'indication des collocations en lexicographie des langues de spécialité pose le problème du degré de similitude entre les collocatifs de l'ensemble d'un terme pluriélémentaire et ceux de la tête de ce terme. Par exemple, si **fonction** est un collocatif attesté de *calculer*, en est-il ainsi, nécessairement, de **fonction brownienne linéaire** ? Il existe des cas nombreux où les propriétés du générique sont héritées par les spécifiques, mais il y a aussi des exceptions : s'agissant de l'exemple ci-dessus il est vrai que certaines fonctions ne sont pas calculables en toutes circonstances.

L'utilisation des outils informatiques dans la fabrication des dictionnaires bilingues de toutes sortes est, en fait, en filigrane, dans la plupart des pages de ce livre. Mais trois chapitres abordent ces questions plus en détail : celui de Frank Knowles, celui de Valerie Grundy et celui d'Elliott Macklovitch. **Frank Knowles** dresse un vaste portrait de ce que l'informatique a modifié dans le travail humain en général, et en particulier en lexicographie bilingue. Il montre comment l'ordinateur sert à accélérer et à systématiser les tâches traditionnelles du lexicographe. Il montre aussi comment l'ordinateur peut faire des choses qu'aucun

lexicographe n'aurait jamais pu faire, comme explorer un corpus, contrôler la symétrie (d'ailleurs en partie utopique) entre les deux parties d'un même dictionnaire, ou gérer des caractères typographiques. Enfin, l'informatique peut être utilisée comme support de l'ouvrage terminé.

**Valerie Grundy** nous fait part de l'expérience qu'elle a acquise dans la compilation du dictionnaire bilingue Hachette-Oxford, réalisé à l'aide de deux corpus, soit un dans chaque langue. Le corpus permet de collecter les formes et de vérifier des traductions, ce qui autorise les auteurs à dire que leur dictionnaire représente la langue telle qu'elle est réellement utilisée dans les discours.<sup>7</sup> L'approche lexicographique des phraséologies, en particulier, a beaucoup bénéficié des apports de l'informatique, qui permet une meilleure appréhension des usages réels (souvent différents, dans ce domaine, des usages reconstruits intuitivement, même par des lexicographes chevronnés) par balayage des corpus et segmentation des fragments répétés. Elle permet surtout, en consultation, d'entrer dans les phraséologies par n'importe lequel de leurs constituants, réglant ainsi le problème de l'adressage que connaissent les dictionnaires «papier». Elle évitera aussi, il est devenu banal de le dire maintenant, les inconvénients liés aux prix de revient élevés des ouvrages sur papier.

Les dictionnaires «en-ligne» (**Elliott Macklovitch**) trouvent leur utilité dans la tâche principale – voire unique – pour laquelle les dictionnaires bilingues ont toujours été conçus, c'est-à-dire la traduction. C'est aussi pour ce même public de traducteurs que l'on peut désormais imaginer des dictionnaires «évolutifs», monolingues ou bilingues d'ailleurs (évoqués dans Thoiron et Béjoint : 1989) et particulièrement utiles lorsque la dynamique de certains domaines rend toute tentative de terminographie «papier» illusoire. Dans le prolongement des dictionnaires actuels, ces recueils, pour lesquels on pourrait parler de «terminographie grise», sont conçus comme des «mémoires de traduction» qui permettent au traducteur de constituer son propre dictionnaire au fur et à mesure qu'il avance dans le travail. Les suites syntagmatiques de tous ordres qui sont retenues sont, grâce à des logiciels spécialement conçus pour cet usage, alignées automatiquement et mises à la disposition du traducteur. Tout ceci conforte la nécessité de former les traducteurs à la lexicographie bilingue, notamment par le recours à l'emploi des systèmes de gestion de bases de données (Thoiron et Maniez : 1989).

Le dernier chapitre, contribution de **Teresa Cabré** et **Cristina Gelpi** fait écho au premier, nous l'avons fait observer ci-dessus : dans le cadre d'une approche «sociolexicographique», les auteurs examinent la place du bilingue dans l'affirmation identitaire d'une culture et d'une langue, en l'occurrence le catalan; la question a souvent été abordée pour les monolingues, mais rarement pour les bilingues. On y voit que la production lexicographique, si on la juge dans son

7 Voir le Collins COBUILD, dictionnaire monolingue de l'anglais dont la première édition est parue en 1987, et qui se présente comme le dépositaire du «real English», au grand dam de ses concurrents.

ensemble pour y discerner les grandes tendances, suit d'une certaine manière l'évolution d'une société, à la fois dans son fonctionnement interne et dans ses rapports avec les sociétés qui l'entourent.

Ce livre ne prétend pas couvrir tous les aspects de la conception et de la réalisation de dictionnaires bilingues. Certains aspects ne sont qu'effleurés, et d'autres sont totalement absents. Il aurait été utile, sans doute, de se pencher par exemple sur de nouveaux types de dictionnaires bilingues, comme les dictionnaires de compréhension, qui supposent, on le conçoit aisément, une réflexion sur les programmes macrostructurels et microstructurels, et sur les modes de présentation, sans parler des supports matériels au moyen desquels ces ouvrages d'un nouveau type pourront dans un avenir proche être mis à la disposition des utilisateurs.<sup>8</sup> Mais le temps et l'espace nous étaient comptés, et il fallait bien commencer. Nous espérons tout de même avoir apporté notre pierre à la construction d'un édifice dont l'importance n'échappe à personne, à un moment crucial dans l'évolution de la discipline.

Ajoutons enfin qu'un tel ouvrage est d'une certaine manière un témoignage de la vigueur de la francophonie. Nous avons voulu y regrouper des auteurs venus de tous les horizons, que rassemblaient le goût pour la lexicographie et le désir de publier le résultat de leurs travaux dans une langue qui n'était pas toujours leur langue principale, à l'intention de lecteurs également divers, et également francophones.

Nos remerciements vont à André Clas, qui est à l'origine de ce livre, et qui a su nous soutenir dans les moments difficiles, à Pascaline Dury et à Christelle Raymond, qui ont donné leur temps sans compter pour que le texte soit aussi irréprochable que possible, et à tous les auteurs, qui ont accepté sans rechigner nos délais exorbitants, les aléas de notre organisation, et nos propositions de modifications – parfois importantes – à apporter à leur texte. A tous, notre reconnaissance.

## **Note sur la terminologie utilisée par les auteurs**

Un effort a été fait pour que la terminologie utilisée par les différents auteurs soit harmonisée. Mais il est clair que cette harmonisation a des limites, si l'on veut respecter les nuances que les auteurs ont voulu exprimer. On trouvera donc, par exemple, tantôt «langue source» et «langue cible», tantôt «langue de départ» et «langue d'arrivée», et même «langue 1» et «langue 2» : ces mots ne sont pas synonymes dans tous les cas, mais le contexte devrait permettre, à chaque fois, de comprendre clairement ce qui est désigné.

<sup>8</sup> Plusieurs équipes travaillent actuellement sur ce concept, parmi lesquelles celle du projet COMPASS financé par l'Union européenne (projet LRE 62.080).



## Chapitre 1

# L'ANTIQUITÉ ET LA VARIÉTÉ DES DICTIONNAIRES BILINGUES

Claude BOISSON

L'ancienneté de la réflexion sur le langage est dramatiquement obscurcie par des images répandues mais simplistes qui ont cours sur l'histoire de la linguistique, caricatures à la fois ethnocentristes et «chronistes», si on me passe ce néologisme, le «chronisme» étant cette fatuité moderne qui consiste à enfler sans pudeur les mérites intellectuels de sa génération et des générations immédiatement antérieures, et de se gausser des lointains prédécesseurs, censés être «naïfs», «pré-scientifiques», quand ce n'est pas «obscurantistes» ou «barbares». Or parmi les témoignages les plus anciens de l'interrogation humaine sur le langage, que l'on pourrait appeler «pré-linguistiques» si l'on tient absolument à ces coupures somme toute assez factices, figurent les dictionnaires, dans le sens le plus général du terme, que ce soient des monolingues ou des bilingues (j'emploie ce terme pour faire simple, mais dans ce qui suit il pourra s'agir de plurilingues). Je m'intéresserai essentiellement aux bilingues (dans un certain nombre de cultures représentatives), dont je rappellerai la vénérable ancienneté, même si, contrairement à ce que l'on avance le plus souvent, dans la plupart des cultures, les dictionnaires monolingues sont apparus avant les dictionnaires bilingues. Nous verrons le pourquoi de ces faits, montrant au passage quels indices précieux les dictionnaires constituent pour la compréhension des attitudes culturelles fondamentales.

Je restreindrai ma présentation aux listes lexicales proprement dites, et je ne tiendrai pas compte des gloses insérées dans des textes; d'autre part, je sélectionnerai les ouvrages les plus anciens, et, préférentiellement ceux qui émanent directement des diverses cultures, plutôt que ceux qui sont dus à des Européens; quant au domaine de la lexicographie européenne, il sera à peine survolé, ayant fait l'objet de nombreux travaux. Pour un traitement plus détaillé et une bibliographie étendue, on se reportera à la monographie de Boisson, Kirtchuk et Béjoint (1991)<sup>1</sup>.

---

<sup>1</sup> On trouvera aussi des informations utiles dans l'*International Encyclopedia of Lexicography* (F. J., Hausmann et al., Ed.), mais la présente contribution permet par sa concision de faire ressortir plus aisément les éléments saillants de l'histoire ancienne des bilingues. Voir aussi la bibliographie méticuleuse de Cop (1990), malgré ses lacunes. Pour les définitions, les typologies de dictionnaires, et différents concepts de la métalxicographie, voir les mises au point de Béjoint (1994).

## 1 Le Proche-Orient ancien : Mésopotamie, Syrie, Égypte

Ce sont les Sumériens qui nous ont donné la première écriture connue, et ce sont également eux les concepteurs des premiers dictionnaires, ou «listes lexicales», dont la conception a été imitée dans tout le Proche-Orient ancien<sup>2</sup>. Ces listes lexicales comptent pour une proportion énorme parmi les tablettes cunéiformes mises à jour, et nous sommes devant la plus étendue des traditions lexicographiques connues, puisque, depuis la fin du IV<sup>e</sup> millénaire, elle s'étend sur trois millénaires. Une part considérable de l'activité intellectuelle de la civilisation mésopotamienne s'absorbait dans ces tâches encyclopédiques et lexicographiques, et dans les écoles de scribes, ou «maisons des tablettes», on enseignait l'apprentissage par cœur et la copie de listes de mots. A l'époque où le sumérien était encore une langue vivante, et même après que, devant la montée de l'akkadien, il fut devenu langue morte classique, on ne semble pas avoir enseigné systématiquement d'autres langues, akkadien mis à part (l'akkadien est une langue sémitique, et donc sans rapport avec le sumérien, langue d'affiliation inconnue). L'école s'intéressait surtout au dialecte principal, l'eme-gir, mais on y a probablement enseigné le dialecte eme-sal (un autre dialecte sumérien). On connaît toutefois des vocabulaires qui comportent un assez grand nombre de mots étrangers, élamites, kassites, subaréens, etc. En sens inverse, à date très tardive (I<sup>er</sup> ou II<sup>e</sup> siècle de l'ère chrétienne), on a des fragments de textes scolaires écrits probablement par des Grecs désireux d'apprendre le sumérien et l'akkadien, et qui transcrivent ces langues dans l'alphabet grec (Sollberger, 1962).

Les listes sont soit des listes de signes (par exemple des syllabaires), soit des listes de mots, soit même les deux. Certaines listes lexicales préfiguraient nos encyclopédies et à vrai dire, le regroupement des mots, qui était thématique, préfigure non seulement nos encyclopédies, mais aussi des ouvrages comme le *Thesaurus* de Roget, dictionnaire conceptuel où les mots anglais sont regroupés par aires sémantiques. Parmi les différents types de listes lexicales, ce sont les listes monolingues qui sont apparues les premières, vers 3300 avant J.-C., sous la forme de signes sumériens. On dispose ainsi de listes de professions, d'animaux, d'arbres et d'objets en bois, de végétaux, de pierres et de minéraux, de noms de pays, de termes mathématiques et économiques, etc.

Mais en Mésopotamie et dans la zone adjacente de la Syrie, les listes monolingues ne sont pas les seules. Il faut noter en effet deux autres catégories. Il y a des listes que l'on pourrait qualifier d'«inter-dialectales», et qui comportent des mots notés en syllabogrammes dans le dialecte sumérien eme-sal, avec leur équivalent dans le dialecte eme-gir. Mais il y a aussi, et surtout, des listes bilingues où chaque terme sumérien est présenté avec sa traduction, listes lexicales bilingues qui constituent bel et bien des dictionnaires de traduction. Du

<sup>2</sup> Voir Cavigneaux (1980-1983), Cavigneaux (1976), et plus brièvement Civil (1975).

moins dans l'état actuel des connaissances, les plus anciens de ces documents n'auraient pas été élaborés en Mésopotamie proprement dite, puisqu'il s'agit de listes sumérien > éblaïte, qui ont été découvertes par une expédition italienne dans les Archives Royales d'Ebla, en Syrie, en 1975<sup>3</sup>. Ces listes remontent au XXIV<sup>e</sup> siècle, soit près d'un millénaire après les premières brèves listes monolingues. L'éblaïte est ainsi la langue sémitique la plus anciennement attestée dans un dictionnaire bilingue et les listes d'Ebla sont les plus vénérables dictionnaires bilingues actuellement attestés (on dispose aussi à Ebla de listes monolingues en sumérien et en éblaïte). On connaît même des listes à trois colonnes : le logogramme sumérien, sa prononciation notée en syllabogrammes, et enfin l'équivalent éblaïte, sinon systématiquement, du moins sporadiquement.

C'est seulement ensuite que sont attestées des listes sumérien > akkadien. Ces dictionnaires bilingues vont jouer un rôle de plus en plus important, pour deux raisons. D'abord, très vite, la civilisation mésopotamienne constitue une symbiose suméro-akkadienne, et ensuite, le sumérien devenant progressivement une langue morte, les scribes akkadophones se trouvaient dans la nécessité d'apprendre avec de plus en plus de difficulté un idiome prestigieux qui finissait par jouer pour eux le rôle que le latin, le grec ancien, le vieux slave, le guèze, l'arabe ou le chinois ont pu jouer dans d'autres cultures.

Je mentionnerai aussi d'autres types de dictionnaires, comme des tablettes bilingues à trois colonnes disposées par groupes synonymiques, qui représentent en même temps les prototypes de nos dictionnaires de prononciation. Le lecteur aura une idée du type de répertoire dont il s'agit d'après le bref fragment suivant, qui translittère dans la colonne de gauche la prononciation du terme sumérien en syllabogrammes (signes à valeur grossièrement «phonétique»), dans la colonne du milieu le mot sumérien en question (écrit avec un logogramme cunéiforme sur la tablette, mais noté dans notre alphabet par des lettres majuscules), et dans la colonne de droite les traductions akkadiennes en syllabogrammes, auxquelles on a ajouté leur équivalent français (le tout d'après Cavigneaux, 1980-1983) :

zu-ú	ZU	la-ma-a-du	«apprendre»
		šu-du-ú-um	«faire savoir»
		e-du-u	«savoir»
		wu-ud-du-ú	«indiquer»
		a- <u>ha</u> -zu	«apprendre»

Mais, là encore, pour ce type de listes, on a maintenant des exemples plus primitifs découverts à Ebla, avec traduction en éblaïte, et non en akkadien. Cependant la «mise en page» est confuse pour nous, puisque la disposition éblaïte comporte une colonne unique, le lecteur étant censé discriminer de lui-même

3 Voir Pettinato (1981), Pettinato et alii (1982), Picchioni (1984), Lambert (1983), Lambert (1984), Von Soden (1984).

dans cet aide-mémoire entre ce qui est du sumérien et ce qui est de l'éblaïte, tandis que l'indication de la prononciation du sumérien, loin d'être systématique, est rare. En tout cas, nous avons ici, à date très ancienne, le prototype des dictionnaires bilingues standards d'aujourd'hui, faisant figurer à la fois la prononciation du mot étranger et sa traduction. La civilisation mésopotamienne ayant rayonné sur presque tout le Proche-Orient ancien, on ne s'étonnera pas de trouver en Syrie (Ougarit, Emar) et en Anatolie (Boghazköy) des exemples de répertoires dérivés en d'autres langues<sup>4</sup>, glossaires bilingues, trilingues, voire quadrilingues, tels des vocabulaires sumérien-hourrite, sumérien-akkadien-hourrite, sumérien-akkadien-hittite, ou sumérien-akkadien-hourrite-ougaritique. L'existence de certains de ces dictionnaires s'explique par la présence d'étrangers, comme c'est le cas avec la colonie hourrite d'Ougarit. Les scribes élamites copièrent eux aussi des vocabulaires.

Pour revenir aux listes suméro-akkadiennes, on trouve même des listes lexicales à quatre colonnes (exemple dans Goetze, 1945) où entre le signe logographique sumérien et la traduction akkadienne, on a inséré le nom du signe : il faut savoir en effet qu'un logogramme sumérien est désigné par son nom, exactement comme la lettre Y de notre alphabet est désignée par le nom «i grec». On aura donc une structure de ce type :

kur	PAP	pa-ap-pu	nak-ru
-----	-----	----------	--------

Elle se lit ainsi : le signe cunéiforme (composés de 2 clous) figurant dans la 2<sup>e</sup> colonne (ici transcrit conventionnellement PAP), signe dont le nom se prononce «pappu» (rendu à la colonne 3 par pa-ap-pu dans la notation du syllabaire), note un mot sumérien qui se prononce «kur» dans cette langue (colonne 1), et qui se traduit en akkadien (colonne 4) par le mot «nakru» (français «ennemi»).

L'Égypte ancienne, pour sa part, fournit une matière lexicographique considérablement plus pauvre que la zone culturelle sumérienne (peut-être à cause des hasards archéologiques, et de la moindre conservation des papyrus par rapport aux tablettes). Mais on y connaît aussi des listes lexicales monolingues, traditionnellement appelées *onomastica* par les égyptologues. Le type le plus ancien, exactement comme en Mésopotamie, est constitué par des collections thématiques de mots. On aurait pu s'attendre à date ancienne à des bilingues égyptien > araméen ou l'inverse, vu l'existence d'une correspondance diplomatique en araméen (devenu langue internationale), mais nous n'en trouvons nulle trace. En revanche, on a trouvé une liste de noms égyptiens d'objets de bois transcrits en cunéiforme, avec quelques traductions égyptien > akkadien (Smith et Gadd, 1925). Ce document a pu être composé par un scribe mésopotamien désireux d'apprendre l'égyptien. En fait, ce n'est qu'aux époques récentes que les

4 Cavigneaux (1980-1983 :617, 633); Civil et al. (1969); Friedrich (1969 :3-4); Güterbock et Carter (1978); Krecher (1969); Landsberger et Hallock (1955); Laroche (1966); Laroche (1968); Lebrun (c. p.); Nougayrol et al. (1968 :230-249); Otten et Von Soden (1968); Thureau-Dangin (1931).

dictionnaires bilingues proprement dits ont fait leur apparition, et l'Égypte est très en retard sur la Mésopotamie dans ce domaine.

Le copte a suscité de nombreux glossaires à organisation thématique : grec > copte d'abord (le plus ancien datant de la seconde moitié du III<sup>e</sup> siècle de notre ère), latin > copte, copte > arabe, voire latin > grec > copte (Kasser, 1972). On a aussi des glossaires qui citent les termes dans deux dialectes, comme le bohaïrique et le saïdique.

## 2 Le Proche-Orient, le Moyen-Orient, l'Éthiopie, l'Inde

Dans la lexicographie arabe ancienne (Haywood, 1965), les bilingues apparaissent aussi après les monolingues, la plupart étant d'ailleurs rédigés par les membres d'autres sociétés en contact avec les arabophones : Persans, Juifs, Turcs, Syriques, etc. Ainsi, dès le IX<sup>e</sup> siècle, un médecin juif de langue maternelle syriaque, Joshua bar Ali, composa un dictionnaire syriaque > arabe (Marello, 1989 : 8). Pour l'hébreu, le premier dictionnaire pourrait être le '*agron*' («Recueil»), répertoire sélectif qui est dû à l'éminent Rabbi Sa'adia ben Joseph, de la ville de Fayyum (Égypte), et dont la deuxième édition (vers 925) est un bilingue hébreu > arabe.

L'Éthiopie, pour sa part, a produit uniquement des bilingues guèze > amharique (entre les X<sup>e</sup> et XVII<sup>e</sup> siècles), puis le but était d'aider à la compréhension de textes religieux rédigés en guèze, devenu une langue morte : la lexicographie éthiopienne, contrairement à ce qui se passe pour l'arabe et pour l'hébreu, comprend uniquement des bilingues. Quant au premier dictionnaire persan, il est monolingue, et après seulement apparaissent les bilingues, comme le dictionnaire arabe > persan *Muqaddimat 'al-'adab* («Introduction à la littérature») de 'al Zamakhsharī (XI<sup>e</sup>-XII<sup>e</sup> siècles).

Pour le sanskrit, les premières productions étaient des glossaires monolingues de mots difficiles dans les Vedas<sup>5</sup>. Le premier bilingue serait un répertoire sanskrit > tibétain renvoyant à la littérature bouddhiste sacrée ou profane, et d'origine exogène, le *Mahāvvyutpatti* («La Grande Connaissance»). Il a été commencé en 814 au Tibet, et terminé par un comité mixte d'experts indiens et tibétains. Puis vient au XII<sup>e</sup> siècle un dictionnaire versifié sanskrit > kannada (le kannada est une langue dravidienne). On trouve ensuite un autre dictionnaire sanskrit > kannada au XIV<sup>e</sup> siècle, un dictionnaire sanskrit > kawi (le kawi est le vieux javanais littéraire, VIII<sup>e</sup>-XIV<sup>e</sup> siècles), pour permettre la lecture de la littérature en vieux javanais, et des ouvrages de l'époque moghole, dictionnaires

5 Sur la lexicographie indienne, on dispose d'un ouvrage très détaillé par Vogel (1979), mais il ne traite que des ouvrages en sanskrit. Pour le domaine dravidien, on peut consulter Dhamotharan (1978), James (1989) et Katre (1965).

marathe > télougou > persan > sanskrit, persan > sanskrit, persan > arabe > sanskrit.

Dans la zone dravidiennne, la lexicographie du tamoul commence avec le *Tolkâppiyam*, de datation problématique (les estimations s'échelonnant du XI<sup>e</sup> siècle avant notre ère jusqu'au V<sup>e</sup> siècle après notre ère), œuvre versifiée, à la fois traité de grammaire, de rhétorique, de lexicographie, et de théorie littéraire. L'ouvrage comporte un petit glossaire de 120 termes littéraires sanskrits avec leur traduction en tamoul. Le premier dictionnaire à être imprimé en tamoul est un bilingue tamoul > portugais, le *Vocabulario Tamulico com a significaçam Portugueza* du Jésuite Antão de Proença (1679).

### 3 L'Europe

Le monde gréco-romain est un domaine incomparablement plus familier. Dans le domaine grec, la lexicographie tire ses origines d'une glose des mots et expressions «difficiles», vieilles, dialectales, spécialisées. Les Grecs étant peu intéressés par les langues des Barbares, on ne pouvait guère s'attendre à ce qu'ils se soient astreints à compiler des dictionnaires de traduction. Chez les Romains, après des glossaires sur les auteurs anciens, grecs d'abord, puis latins, apparurent des glossaires bilingues. Les premiers seraient un glossaire grec > latin, faussement attribué à Cyrille d'Alexandrie, et qui daterait peut-être du VI<sup>e</sup> siècle, et le glossaire latin > grec, dit de Philoxène, censément aussi du VI<sup>e</sup> siècle, mais qui serait en fait du VIII<sup>e</sup> ou du IX<sup>e</sup> siècle (Marello, 1989:7). Citons par curiosité le très bref *Glossaire d'Endlicher* gaulois > latin, avec ses 17 gloses, postérieur au V<sup>e</sup> siècle (Lambert, 1994:203-204).

Pour l'Europe occidentale médiévale et moderne on trouve à la fin du premier millénaire des listes manuscrites (généralement thématiques) de mots latins avec leur traduction en langue vernaculaire. Le premier bilingue ne faisant pas appel au latin date de 1477 : c'est un dictionnaire italien > allemand classé, là encore, thématiquement, le *Introito e porta*, d'Adam von Rottweil (Stein, 1990:31). On observera incidemment qu'en Angleterre, au XVII<sup>e</sup> siècle, des lexiques latin > anglais étaient inclus dans des grammaires, plutôt que publiés séparément (Mitchell, 1994).

Pour l'arménien, c'est aussi par un bilingue que commence la lexicographie, le *Glossaire d'Autun*, dictionnaire latin > arménien datant des IX<sup>e</sup>-X<sup>e</sup> siècles, rédigé par un clerc qui avait conversé avec un Arménien en lui faisant désigner du doigt des objets, ou en l'interrogeant sur des notions courantes, notamment religieuses.

## 4 L'Asie Orientale et l'Asie Centrale

Pour le japonais, le premier dictionnaire est un bilingue chinois > japonais de 3000 entrées, le *Shinsen jikyô* (*Nouveau miroir choisi de caractères*), compilé entre 898 et 901. Le premier dictionnaire bilingue vers une langue européenne serait un dictionnaire japonais > portugais rédigé dès 1563. Mais c'est en 1603 que paraît sur les presses de la mission portugaise de Nagasaki le premier ouvrage de ce type parvenu jusqu'à nous, le *Vocabulario da lingua de Iapan com adeclaração em Portugues*.

En Chine les plus anciens dictionnaires sont monolingues, le premier qui nous soit parvenu remontant à la période des deux premiers siècles avant J.-C.<sup>6</sup> Quant aux dictionnaires bilingues ou plurilingues, ils sont très clairement plus récents, et dus à l'impact de l'étranger. Les premières attestations datent des Ming, mais à cette période ont été édités des manuels compilés plus anciennement. Poppe (1965:22) appelle *Hua-i-i-yü*, qu'il date de 1389, un dictionnaire spécifique, qu'il cite comme étant le premier bilingue chinois > mongol. Apparaissent ensuite des dictionnaires mandchou > chinois sous la dynastie des Ch'ing (Qing). On trouve aussi des vocabulaires de termes du bouddhisme, dont un lexique chinois > sanskrit composé à l'instigation de l'empereur K'ang-hsi (1662-1723), ou bien le «lexique pentaglotte» du *Man han hsi-fan chi-yao*, ou *Wu i ho-pi chi-yao*, registre fournissant les équivalents en sanskrit, tibétain, mandchou, mongol, et chinois. Ce dernier ouvrage, publié à l'instigation de l'empereur Ch'ien-long (1736-1795), n'est en fait qu'un dérivé de la *Mahavyutpatti* indo-tibétaine mentionnée plus haut. Enfin il faut noter évidemment les dictionnaires de chinois compilés par les Occidentaux, spécialement les missionnaires jésuites.

Parmi les langues altaïques, pour ce qui est de la branche mongole, les premières listes de mots se rencontrent au XIII<sup>e</sup> siècle dans des ouvrages d'un Arménien, de Persans et d'Arabes. Aux XIV<sup>e</sup> et XV<sup>e</sup> siècles, des savants musulmans mettent en œuvre une série de glossaires bilingues arabe > mongol et persan > mongol, la langue concernée étant le moyen mongol occidental. Le plus important est un dictionnaire quadrilingue arabe > persan > chaghataï > mongol (le chaghataï est une langue littéraire du stade moyen turke), dérivé du *Muqaddimat* de 'al Zamakhsharî, déjà mentionné. On trouve aussi quelques glossaires arménien > mongol ou géorgien > mongol.

Pour les langues de la branche turke, on notera au stade des langues du moyen turke (XI<sup>e</sup>-XV<sup>e</sup> siècles) les deux ouvrages suivants : pour la langue littéraire du royaume des Qarakhanides, le *Divân Luyât at-Turk* arabe > turc de Mahmûd al-Kâsyrî (1073) et pour le kouman (variété ancienne du kiptchak, le «polovtsien» des Russes), un dictionnaire latin > persan > kouman de la fin du XIV<sup>e</sup> siècle, et

6 On trouvera des informations sur les dictionnaires chinois dans Norman (1988 : 168-172, 28-34) et dans Xue Shiqi (1982). On trouve citées dans Marelllo (1989) les bibliographies de Chien (1986) et de Fu-mien Yang (1985), que je n'ai pas pu consulter.

des glossaires arabe > kouman. Pour le turc osmanli, noter un bilingue italien > turc dès 1641, le turc y étant transcrit en caractères latins (Kurtbötke, 1994).

## 5 Les dictionnaires des langues sans écriture

### 5.1 *Dictionnaires écrits*

Les vocabulaires «exotiques» suivent comme il se doit l'avancée européenne, des listes de mots apparaissant au XVI<sup>e</sup> siècle pour l'Afrique, en 1770 pour l'Australie, en 1828 pour la Nouvelle-Guinée. Le domaine amérindien est spécialement remarquable, et mérite qu'on s'y arrête plus à loisir, à cause de son ancienneté, et de la qualité des travaux (ils font mentir les légendes condescendantes sur l'ethnocentrisme et l'incompétence supposées des missionnaires). Certains des dictionnaires ont été établis très rapidement après le contact avec les Européens, puisque, dès 1600, on disposait déjà de bilingues (vers l'espagnol) pour pas moins de cinq langues mexicaines. D'abord pour le nahuatl (aztèque) parut en 1555 à Mexico le plus ancien dictionnaire, intitulé (*Aquí comienza vn vocabulario en la lengua Castellana y Mexicana*) par Fray Alonso de Molina, huit ans après la première grammaire. En 1571 le même Molina publia un dictionnaire plus étendu, le *Vocabulario de la lengua mexicana*. Peu de temps après sortirent des dictionnaires pour le tarasque (1559), le tzeltal (1571), le zapotèque (1578), et le mixtec (1593) (McQuown, 1967:4). Nous verrons dans la conclusion ce qui a pu favoriser une éclosion aussi rapide.

Pour les langues situées au nord du Mexique, qui ne servaient pas de véhicules à une culture aussi prestigieuse que celle des Aztèques, l'activité lexicographique fut plus tardive. Pour les langues des Caraïbes, on a des mots isolés, puis des listes modestes, dans des récits de voyage espagnols et italiens des XV<sup>e</sup> et XVI<sup>e</sup> siècles, d'abord pour le taïno, maintenant éteint, la première langue rencontrée en Amérique par les Européens, et pour l'arawak (Taylor, 1977:13-18; Warner, 1899:78-79). Enfin, pour l'Amérique du Nord, c'est seulement en 1643 que le pasteur Roger Williams, fondateur de la colonie de Rhode Island, publie à Londres un «phrase book» d'une langue algonquienne maintenant éteinte, le narraganset; un autre missionnaire protestant, Paul Egede, a publié en 1750 un dictionnaire esquimau (groënlandais) > danois > latin (Hoijer, 1976:3-4).

Tous ces ouvrages ont été conçus par des Occidentaux étudiant les langues «exotiques», mais un extraordinaire cas inverse récemment étudié mérite une mention spéciale (Clayton, 1989). La Newberry Library de Chicago renferme sous forme manuscrite un dictionnaire trilingue espagnol > latin > nahuatl, qui est modelé sur le *Vocabulario de romance en latin* d'Antonio de Nebrija, mais avec ajout de traductions en nahuatl. Ce manuscrit, en fait la copie d'un manuscrit antérieur disparu, est datable entre 1530 environ et le premier quart du XVII<sup>e</sup>

siècle. Son examen minutieux indique que l'auteur confond des sons espagnols et commet des erreurs de traduction, le tout démontrant que sa langue maternelle n'est pas l'espagnol, mais le nahuatl. La conclusion est ainsi spectaculaire, puisque nous nous trouvons devant un dictionnaire d'espagnol compilé par un lexicographe nahuatl, et cela à une date remarquablement ancienne.

## **5.2 Dictionnaires oraux**

On a vu que les premiers dictionnaires sumériens sont liés à des exercices scolaires, et que ces exercices présupposent une tradition orale, et une mémorisation des listes, laquelle favorisait l'apprentissage oral (Civil, 1975:130; Landsberger, 1956:126). On a même osé conjecturer qu'il a pu exister des traditions orales de mémorisation de listes monolingues antérieures à l'apparition de l'écriture (Westenholz, 1985:295).

Cela nous donne matière à réflexion. Bien des cultures traditionnelles pratiquent l'intermariage entre conjoints de langues différentes, et se caractérisent donc par une pratique massive d'au moins deux langues (la langue du père et celle de la mère). On peut se demander si certaines de ces cultures connaîtraient des listes lexicales bilingues – ou monolingues – à transmission purement orale, des «dictionnaires oraux», éventualité à première vue douteuse, mais qui appellerait une incursion systématique dans la littérature ethnologique, et constituerait une justification supplémentaire au terme «variété» dans le titre du présent chapitre. Or une publication toute récente sur les Peuls (Guinée, Mali), due à l'ethnolinguiste Christiane Seydou, montre qu'une telle conjecture était parfaitement justifiée (Seydou, 1989a; Dumestre, 1989). Seydou m'indique qu'elle n'a pas entendu parler de faits analogues en dehors des Peuls, mais cela ne prouve nullement que le cas peul est unique, puisque la découverte de dictionnaires oraux est toute fraîche, et l'existence de productions analogues dans d'autres cultures a très bien pu passer inaperçue.

Les Peuls cultivent un genre poétique appelé «mergol», où, dans de subtils entrelacs de jeux stylistiques portant sur les sonorités et le rythme, s'exaltent la virtuosité et la science du récitant. Ces déclamations touchent aux sujets les plus variés sous forme de récits, louanges, souvenirs, etc. Fréquemment, certaines de ces productions peuvent contenir des listes monolingues de mots (recouvrant typiquement un champ sémantique précis), ou même être constituées dans leur entier par une liste lexicale. Or, devenus virtuoses dans l'utilisation de tels dictionnaires thématiques oraux, les Peuls en sont venus occasionnellement à établir de véritables dictionnaires oraux bilingues (versifiés, qui plus est, comme en Inde, ce qui favorise la mémorisation), dont deux exemples sont bien attestés. Il s'agit en premier lieu d'un lexique arabe > peul, à objectif didactique, visant à inscrire dans la mémoire d'étudiants peuls un vocabulaire de base de 365 termes

arabes de la vie courante, sous la forme d'un poème de 110 vers composé sur un modèle métrique arabe. Mis à part l'observation des règles métriques, le regroupement des mots est fondamentalement sémantique. Sont donc associés des termes appartenant à un même champ sémantique (parties du corps, animaux, vêtements, etc.), ou bien des antonymes (soleil / ténèbres, doux / amer, etc.), ou des termes liés par la métonymie ou la métaphore (abeille / miel / cire, feu / fumée / village / maison, etc.). Un exemple de vers en donnera une petite idée (l'arabe est mis conventionnellement entre crochets droits pour le distinguer du peul, et la traduction française figure en-dessous) :

[r-ríhu] henndu, [wa-l-hawâ] ko weeyo  
 [le vent] vent, [et l'air] c'est air

Le deuxième dictionnaire oral, un lexique bambara > peul, constitue une portion d'un mergol de 2200 vers dû à un fameux aède peul enregistré au Mali en 1973. Christiane Seydou indique que cet aède y «a inséré un lexique bambara > peul de cent soixante-douze mots (s'étalant sur deux centaines de vers) [...]». Choisir pour sujet de mergol un lexique bilingue, c'est, à l'instar des héros chevaleresques des épopées peules, créer délibérément la situation la plus périlleuse pour avoir le double honneur de l'affronter – sans y avoir été contraint par autre que soi-même – et d'en sortir vainqueur devant tous» (Seydou, 1989a). Ce dictionnaire bilingue étant un tour de force esthétique et ludique dénué de tout objectif pédagogique, le principe de regroupement des termes n'est plus sémantique, mais strictement formel : il s'agit, pour l'essentiel, de la combinaison d'une structure prosodique précise et de récurrences des sonorités initiales des mots, le tout selon des agencements extrêmement subtils. Dépouillés de la forme poétique qui en fait le véritable intérêt, les vers de ce dictionnaire se présentent sous des équations telles que : «X [le mot bambara] c'est Y [le mot peul]», ou bien «X ils disent Y», ou bien «X est nom de Y», voire simplement «X Y».

## 6 Commentaires et conclusion

Nous avons pu constater que dans la majorité des cas recensés, le dictionnaire monolingue précède le bilingue. C'est au demeurant ce que l'on devait attendre, car les grandes civilisations anciennes à écriture étaient centrées sur elles-mêmes. A chaque fois qu'on se trouve dans le cas contraire, une explication est facile à trouver. Ainsi l'antériorité des dictionnaires bilingues sur les monolingues dans le cas des langues modernes d'Europe peut s'expliquer par l'absence de textes fondateurs internes à la culture européenne, contrairement à ce que les Indiens ont eu avec les Vedas, les Juifs avec la Bible, les Grecs avec Homère, les Chinois avec le canon confucéen, les Arabes avec le Coran. En effet parmi les grands textes qui structurent la civilisation européenne, les plus anciens, et aussi les seuls qui soient vraiment communs à toutes les sous-cultures de ladite civilisation, lui sont en quelque sorte *externes* : il s'agit des classiques grecs et latins, qui établissent la

composante humaniste, et de la Bible, qui établit la composante chrétienne. Ainsi, ce n'est pas l'effet du hasard si les premiers dictionnaires européens sont des bilingues. Dans le cas des langues altaïques, les premiers dictionnaires caractérisés semblent aussi avoir été des bilingues, ce qui s'explique, là encore, par les conditions culturelles. En effet, si considérable qu'ait pu être dans l'histoire de l'Eurasie l'impact des peuples de langues turkes et mongoles, il reste qu'au temps de leur splendeur, une bonne part de leur culture non-matérielle finit par prendre une couleur exogène. Ils perdirent par exemple leurs religions anciennes pour se convertir au manichéisme, à l'Islam et au bouddhisme. Le plus illustre des peuples de langue turke fonda l'Empire Ottoman, dont les liens avec l'ancienne culture turke sont bien distendus. Chez les Japonais également, c'est le dictionnaire bilingue qui apparaît en premier. Les conditions culturelles l'expliquent : c'était le chinois qui dominait la composition des textes culturels et officiels. Le japonais avait ainsi le statut des langues vulgaires de l'Europe face au latin, mais en bien plus marqué encore.

Si maintenant nous revenons au cas «normal», celui des cultures où le dictionnaire monolingue est antérieur au bilingue, nous observons qu'en réalité la distinction entre monolingue et bilingue est moins tranchée qu'il n'y paraît. Pour les premières productions lexicographiques, nous avons vu se dessiner deux grands types, qui paraissent bien correspondre à deux stades historiques. Dans les deux plus anciennes civilisations affirmées, Sumer et l'Egypte, on a élaboré d'abord des listes monolingues «profanes» destinées à la formation des futurs «cadres» de la bureaucratie étatique. On a vu qu'en Mésopotamie proprement dite, les dictionnaires bilingues n'apparaissent pas dans une société qui en aurait besoin dans ses rapports avec une autre société (comme c'est par exemple le cas des dictionnaires des nations modernes de l'Europe), mais dans une société à culture syncrétiste pour laquelle ils constituent un truchement interne. L'ancienneté des dictionnaires bilingues à Ebla (il y a 44 siècles) et en Mésopotamie est directement liée au caractère constitutionnellement bilingue de ces cultures. A ce propos on en profitera pour relever l'ancienneté des concepts qui informent pratiquement tous nos grands types de dictionnaires. En effet, c'est dans la zone proche-orientale Mésopotamie-Syrie que, sous l'impulsion sumérienne, sont apparus les prototypes de quasiment toutes les sous-catégories principales du genre «dictionnaire» : les premiers dictionnaires monolingues conceptuels (thématiques), les premiers dictionnaires bilingues inter-langues et inter-dialectes, ainsi que les premiers dictionnaires de prononciation et les premiers dictionnaires d'homophones; même notre classement alphabétique est très lointainement et obscurément préfiguré par le classement acrographique. Parmi les dictionnaires envisageables dans une culture aussi ancienne, le seul grand type à faire défaut est le dictionnaire monolingue avec définitions systématiques, dont les premiers exemples se trouveront bien plus tard en Grèce, en Inde et en Chine, et qui caractérise donc un stade postérieur dans le développement général des outils lexicographiques. Finalement, une des très rares innovations majeures apportées

par la culture européenne de ces derniers siècles est le dictionnaire plurilingue comparatif : on en a maintenant des exemples impressionnants pour les amples familles linguistiques indo-européenne, dravidienne, ouralienne, bantoue, austronésienne, et, depuis peu, afroasiatique. Mais il ne s'agit que de l'application lexicographique d'une méthode «récente», celle du comparatisme linguistique, inimaginable par les Anciens. En dehors de la dimension historique, l'apport de la culture occidentale moderne semble consister essentiellement en ce classement commode qu'autorise l'écriture alphabétique, en une ampleur, une rigueur et une systémativité sans cesse accrues, puis en une mutation du support, devenu depuis peu électronique, d'où un changement dans le mode de consultation par les utilisateurs, aussi bien que dans la possibilité pour les rédacteurs de rendre plus cohérente la rédaction des entrées. Ainsi donc, si l'Europe moderne a élaboré des chefs-d'œuvre lexicographiques sans équivalents, il reste qu'elle n'a inventé pratiquement aucune des grandes classes de dictionnaires.

Pour revenir au cas spécifique des dictionnaires bilingues, observons qu'une attitude ouverte envers les langues indigènes a pu favoriser le bilinguisme, et cela même dans une situation coloniale à domination européenne. Ainsi au Mexique, les conditions d'apparition des dictionnaires bilingues sont directement déterminées par des décisions ecclésiastiques (Landar, 1977; Suárez, 1983). Après le concile de Trente, l'Eglise avait en effet encouragé l'évangélisation dans les langues vernaculaires, alors même que la messe demeurait en latin, allant jusqu'à interdire de confesser les indigènes dans une langue autre que la leur. Une lettre envoyée en 1575 par l'archevêque de Mexico à Philippe II montre l'importance attachée par la hiérarchie catholique à la connaissance des langues indigènes, et la proportion très élevée du clergé séculier qui les maîtrisaient peu ou prou : 33 % des prêtres résidant à Mexico connaissaient au moins une langue indienne (généralement le nahuatl), et 85 % de ceux qui se trouvaient en dehors de Mexico (Edmonson, 1984). Plus tard, l'Etat espagnol se montrera moins réceptif, et au XVIII<sup>e</sup> siècle, sous l'empire des idées uniformisantes des Lumières, dont en France l'abbé Grégoire fut un illustre représentant, l'objectif avéré devint carrément la disparition des langues indiennes.

Dans les autres cas (mondes indien, chinois, grec, arabe, juif, persan), l'impulsion des premiers dictionnaires écrits vient de la nécessité d'éclaircir des termes obscurs, rares, ou anciens, spécialement ceux qui se rencontraient dans les œuvres classiques fondatrices de chaque culture, notamment les écrits religieux ou assimilés, ou dans des écrits poétiques prestigieux<sup>7</sup> On pourrait aller jusqu'à soutenir qu'il s'agit presque là, au moins dans un certain sens, de dictionnaires inter-dialectaux, le mot «dialecte» étant mis pour toute variété de langue, tant en synchronie (pensons aux dialectes du sumérien, du copte) qu'en diachronie. Dans

7 Pour une synthèse rapide sur les motivations de l'activité lexicographique, voir Hausmann (1989).

le dernier cas, on glose les termes d'un «dialecte» gagné peu à peu par l'obscurité mais demeurant prestigieux, par ceux d'un «dialecte» postérieur, et l'on se trouve dans une situation de diglossie nécessitant une aide de traduction. De sorte que la distinction tranchée entre «monolingue» et «bilingue» devient parfois problématique : il est malaisé de décider jusqu'à quel moment deux codes linguistiques peuvent être considérés comme simples variétés d'une même langue, et à partir de quand ils acquièrent le statut des langues différentes. On peut par exemple poser le problème en ces termes dans le cas de la longue évolution égypto-copte sur plus de quatre millénaires et demi.

Enfin on insistera sur ce qu'ont de capital pour la métalexigraphie générale les données toute fraîches sur les dictionnaires oraux peules, car elles suggèrent la possibilité de dictionnaires purement oraux chez d'autres peuples sans écriture. Elles constituent un appui à la spéculation de Westenholz citée plus haut sur une possible tradition lexicale orale antérieure à l'écriture en Mésopotamie, et laissent du même coup entrevoir la possible existence de dictionnaires avant l'apparition de l'écriture. Au reste, même dans des cultures à écriture, à date ancienne la mémorisation purement orale continue souvent à jouer le rôle principal : ainsi en Mésopotamie, ainsi en Inde pour les «trésors», ouvrages destinés aux poètes, aides à la composition qui devaient être appris par cœur, et qui d'ailleurs, comme une bonne partie des écrits «scientifiques» en sanskrit, étaient versifiés (et les linguistes ne doivent jamais oublier que les admirables grammaires du sanskrit étaient apprises par cœur).

Il est d'autant moins invraisemblable d'envisager de tels dictionnaires purement oraux que l'utilisation explicite de listes de mots est connue dans d'autres cultures : «Les enfants tukano, alors qu'ils apprennent la langue paternelle – qui deviendra leur langue –, se voient inculquer, de façon très explicite, des listes de mots par la mère dans la langue de cette dernière» (Queixalos, 1989:59, citant des données recueillies par Sorensen). Je ne peux donc pas souscrire sans réticences à l'assertion suivante, qui synthétise une position souvent admise implicitement, et qui accorde trop à l'écriture : «It seems reasonable to assume that illiterate speakers of a language do not perceive the flow of speech as consisting of words, but rather of larger units, such as phrases» (Staal, 1979:10). Ce prétendu gouffre entre illettrés et «lettrés» est largement illusoire.

Tout ceci nous permet d'entrevoir l'existence de dictionnaires oraux monolingues et même bilingues chez les peuples sans écriture, et par suite, de toute une réflexion lexicographique qui n'a certes rien de nécessairement «sauvage», l'apparition de l'écriture ne constituant pas un point de rupture aussi radical qu'on aurait pu le penser. Je ne peux qu'approuver la pénétrante conception de Christiane Seydou, contre les thèses fameuses de Goody, lorsqu'elle observe que les données peules «remettent [...] notamment en question la liaison trop souvent établie entre la pratique de l'écriture et les divers processus de conceptualisation et d'organisation applicables au phénomène linguistique, tels que : la capacité d'isoler les lexèmes, et, en ceux-ci, les phonèmes; l'instauration de classifications

et l'établissement de listes; «la formulation explicite de systèmes catégoriels ou de champs sémantiques» et enfin, la rhétorique elle-même»<sup>8</sup>. Il est ainsi licite de poser comme hypothèse de travail qu'avant les listes lexicales écrites les plus antiques, celles des Sumériens, des Eblaïtes et des Égyptiens, avaient pu exister des listes (monolingues, voire bilingues) transmises oralement pendant de nombreuses générations, et que lorsque l'écriture fit son apparition, les hommes connaissaient déjà sous une forme élémentaire l'objet linguistique que nous sommes convenus d'appeler «dictionnaire».

---

8 Seydou (1989a :322); voir aussi Seydou (1989b). Il est fait allusion à Goody (1979); voir aussi Goody (1986:65-66).

# LES DIFFÉRENTS TYPES DE DICTIONNAIRES BILINGUES

Carla MARELLO

### 1 Qu'est-ce qu'un dictionnaire bilingue ?

Le dictionnaire bilingue est un dictionnaire dans lequel des expressions dans une langue (dite langue source ou de départ) sont traduites dans une autre (dite langue cible ou langue d'arrivée). Mais ce n'est pas seulement la présence de deux langues qui fait d'un dictionnaire un bilingue, c'est la raison pour laquelle les deux langues sont mises en contact, c'est-à-dire la communication, par la traduction, entre deux communautés qui ne partagent pas la même langue. Il existe des dictionnaires dans lesquels les adresses (ou entrées ou vedettes) et les gloses appartiennent à deux langues différentes et qui pourtant ne sont pas des bilingues. Par exemple, dans le *Französisches etymologisches Wörterbuch* (1922-1928) de von Wartburg les adresses sont en latin et les gloses en allemand, mais ce n'est pas un bilingue latin > allemand, c'est un dictionnaire étymologique qui parle en allemand de l'origine des mots français dérivant de mots latins choisis comme entrées.

Il existe des dictionnaires compilés par des anthropologues qui expliquent en français ou en anglais, russe, allemand, espagnol (ou toute autre langue accessible à la communauté internationale des savants), le signifié des mots de langues polynésiennes ou africaines, ou amérindiennes ou sibériennes. Ces dictionnaires, qui sont conçus surtout pour décrire des langues inconnues ou mal connues, ont une fonction ambiguë : seuls dictionnaires disponibles pour quiconque veut apprendre à communiquer avec les locuteurs natifs de ces langues, ils sont donc utilisés aussi comme des bilingues. C'est à un processus presque inverse que l'on assiste dans l'histoire des dictionnaires bilingues français : les premiers étaient destinés à aider leurs utilisateurs à traduire du français en latin, mais ils furent utilisés comme des monolingues par une grande partie du public. Le *Dictionnaire françois-latin* de Robert Estienne (1539), comme le dit l'auteur dans l'introduction à l'édition de 1564, était fort utile «à tous desirants entendre la propriété de la langue françoise», parce qu'à cette époque il n'y avait pas de dictionnaire monolingue du français.

Wooldridge (1977), étudiant les origines de la lexicographie française, a montré que le premier dictionnaire monolingue du français, le *Thresor de la langue françoise* (1606) de Jean Nicot est étroitement lié à l'édition de 1573 du *Dictionnaire françois-latin* de Estienne, révisée et augmentée par Dupuys et par Nicot lui-même. Cette filiation était pour ainsi dire inévitable à partir du moment où, jusqu'à la parution du monolingue, le public avait utilisé les gloses latines du dictionnaire français > latin non pas uniquement pour traduire en latin, mais aussi en tant que définitions des signifiés des entrées en français.

Pour les premiers dictionnaires français > latin et en général pour tous les dictionnaires bilingues langue moderne > latin parus au XVI<sup>e</sup> siècle, Quémada (1967:52) suggère la dénomination de semi-bilingues, à cause non seulement de leur utilisation comme monolingues, mais aussi de la forme des entrées en français. Il ne s'agit pas toujours de mots-vedettes, ou de sous-vedettes, mais souvent de syntagmes qui sont des interprétations, des paraphrases contextuelles, du mot latin donné comme équivalent.

La fin du XVI<sup>e</sup> siècle voit la parution de dictionnaires monolingues bilingualisés à la suite d'un procédé nouveau : pour aider les étudiants d'une langue étrangère à employer plus aisément le monolingue pour apprenants, on «bilingualise» un peu les gloses, insérant des explications dans la langue maternelle des étudiants. Ces explications ne sont pas utilisées comme des équivalents, mais aident les étudiants à comprendre les définitions en langue étrangère du monolingue (voir 4 ci-dessous).

Tant dans le cas des semi-bilingues que dans le cas des monolingues bilingualisés les deux langues ne sont pas sur le même plan : l'une est un outil servant à donner accès à l'autre<sup>1</sup>.

Dans la vaste étude qu'elle consacre aux débuts de la lexicographie française, Margarete Lindemann fait remarquer, entre autres, que «Presque tous les dictionnaires français-langues vivantes du XVI<sup>e</sup> siècle ont été publiés à l'extérieur de la France et étaient destinés à l'apprentissage de la langue française» (Lindemann 1994:786). Le fait que ce soient des auteurs résidant à l'étranger qui se chargent de la compilation des dictionnaires bilingues d'une langue a toujours été le signe du prestige international de cette langue.

1 On verra plus loin que dans les articles de moyenne ou grande complexité d'un dictionnaire bilingue une des deux langues est souvent employée de façon métalinguistique, c'est-à-dire comme un outil pour parler de l'autre langue, mais cela se produit à l'intérieur d'un cadre où l'on trouve aussi de véritables items bilingues. Dans les semi-bilingues ou dans les semi-monolingues, au contraire, les items bilingues sont déséquilibrés parce que l'une des deux composantes est métalinguistique; elle n'est employée ni en tant que langue-objet, ni en tant qu'autonyme.

## **1.1 La bipartition des dictionnaires bilingues**

Les dictionnaires bilingues sont en général constitués de deux parties : une partie où la langue source est par exemple le français et la langue cible une langue étrangère, et une partie où la langue source est la langue étrangère et la langue cible le français. Quelle est celle des deux parties qui doit précéder l'autre dans les dictionnaires en un volume ? Il n'y a pas, sur ce point, de politique communément acceptée, ni d'ailleurs de déclarations explicites dans les pages introductives. On peut seulement observer que dans la plupart des cas la partie langue étrangère > langue maternelle précède l'autre, peut-être parce qu'on ouvre et feuillette plus aisément les premières pages d'un très gros volume. On voudrait ainsi faciliter l'emploi le plus fréquent des bilingues, c'est-à-dire la version, la traduction de L2 en L1.

Aujourd'hui, les dictionnaires généraux, qui couvrent tous les secteurs du lexique, sont toujours constitués de deux parties au moment de leur publication. Au début de la lexicographie bilingue on publiait souvent des volumes qui contenaient seulement une partie, et l'autre pouvait paraître beaucoup plus tard. Les lexicographes travaillaient alors presque seuls et s'attachaient à compiler d'abord la partie qui était la plus demandée sur le marché, ou la partie pour laquelle ils avaient le plus de citations, ou encore ils commençaient par la langue source qu'ils maîtrisaient le mieux. De nos jours, les dictionnaires bilingues sont compilés par des équipes rédactionnelles qui travaillent en général en même temps sur les deux parties et qui comptent toujours, en principe, des rédacteurs locuteurs natifs de chaque langue.

Quand une des deux parties du dictionnaire bilingue est compilée avant l'autre, et que la deuxième est basée sur un renversement de la première, le résultat n'est pas de bonne qualité. Tout d'abord, on aboutit à une grande quantité d'items bilingues «déséquilibrés» dans le sens évoqué en 1 ci-dessus. Mais surtout, on court le risque de ne pas respecter la forme que chaque langue donne au contenu. Le découpage du domaine sémantique couvert par un mot, et donc le nombre d'acceptions de l'article lexicographique que doit avoir l'entrée qui lui est consacrée, ne peuvent pas être les mêmes selon qu'on est dans le cadre d'un monolingue, ou d'un bilingue de thème ou d'un bilingue de version (voir ci-dessous : 1.2, 3.1, 3.2). La pratique du renversement est toutefois vivement recommandée en tant que procédé de contrôle en cours de travail et surtout quand les deux parties ont été terminées en parallèle, parce qu'elle permet de vérifier que tous les équivalents donnés dans une partie du dictionnaire figurent également comme adresses dans l'autre partie (sauf décision contraire). L'utilisation de l'ordinateur rend ces contrôles croisés beaucoup plus rapides (voir Knowles, dans ce volume).

Dans les dictionnaires bilingues généraux publiés sous forme de volumes de deux milles pages ou plus, les deux parties ne sont pas toujours de la même longueur. Il est courant que la partie version soit sinon plus longue du moins

constituée d'un plus grand nombre d'adresses, alors que la partie thème a des articles en moyenne un peu plus longs<sup>2</sup>. Les dictionnaires bilingues de langues de spécialité ont souvent deux parties très différentes, l'une plus développée et l'autre réduite à un glossaire bilingue, c'est-à-dire à une liste alphabétique d'items avec leurs équivalents dans l'autre langue, mais sans véritables articles lexicographiques. La partie développée est souvent la partie qui a comme langue source la langue dominante dans le domaine de spécialité : par exemple les dictionnaires bilingues ou plurilingues d'informatique ont souvent une partie anglais > langue X développée et une partie langue X > anglais plus réduite.

Les dictionnaires homoglosses, c'est-à-dire les dictionnaires qui placent côte-à-côte un dialecte ou un patois et une langue nationale (dictionnaires inter-dialectes pour Boisson, voir dans ce volume), sont souvent constitués de la seule partie dialecte > langue nationale. S'ils ont deux parties, la partie langue nationale > dialecte est plus mince, parce qu'on considère qu'ils seront peu ou pas du tout employés comme dictionnaires de thème, pour la traduction en dialecte. De même, dans les dictionnaires bilingues de langue «mortes» la partie langue morte > langue vivante est souvent la seule publiée, ou bien elle est plus développée que l'autre. Inévitablement, au fur et à mesure que l'emploi du latin en tant que langue internationale a diminué on a assisté à une réduction progressive de la partie langue vivante > latin. Lorsque le dictionnaire est consacré à un emploi spécial (voir ci-dessous 2.3), ou qu'il doit mettre en contact des langues très différentes sur le plan du prestige ou même de la morphologie, la bipartition est compromise.

## 1.2 La bidirectionnalité : une utopie ?

Par «bidirectionnalité» d'un dictionnaire bilingue, on entend son aptitude à servir les deux communautés linguistiques soit en tant que dictionnaire de version, soit en tant que dictionnaire de thème. La partie A > B sert soit aux A-phones en tant que dictionnaire de thème, soit aux B-phones en tant que dictionnaire de version. La partie B > A sert soit aux B-phones en tant que dictionnaire de thème, soit aux A-phones en tant que dictionnaire de version.

Un dictionnaire qui prétendrait atteindre ces deux objectifs aurait des articles comme les suivants, tirés de dictionnaires existants (modifiés par Iannucci, 1962:210, 212) :

**ouverture** (d'un objectif) (n.f.), aperture (of a lens);  
 (d'une caverne) (n.f.), mouth (of a cave); (d'un  
 pont, d'une voûte) (n.f.), span (of a bridge,  
 of an arch); (d'un palmer) (n.f.), capacity (of a micrometer)

2 Ces tendances sont difficiles à discerner quand le dictionnaire se veut bidirectionnel et qu'elles sont confrontées aux particularités morphologiques de chaque langue. La nomenclature d'un dictionnaire bilingue français et allemand, par exemple, présentera toujours une partie allemand-français beaucoup plus riche en nombre d'adresses que l'autre, à cause des mots composés allemands

**nice** adj., (delightful - charmant) joli; (delicate - délicat) délicat, fin; (exact - correct) exact, juste; (scrupulous - scrupuleux) (fastidious - difficile) difficile, sévère, rigide, prude; (refined - raffiné) recherché, subtil...

On voit que le souci de donner les indicateurs (informations sémantiques entre parenthèses qui permettent de distinguer entre les différentes acceptions) soit en français soit en anglais alourdit nettement l'article, et le rend difficile à lire.

En fait, dans les dictionnaires bilingues imprimés, la solution adoptée est un compromis.

- Une monodirectionnalité claire, quoique inavouée ou même niée dans l'introduction : le dictionnaire vise à servir avant tout la communauté dans laquelle la maison d'édition veut le vendre;
- Une bidirectionnalité partielle. Par exemple on utilise, si possible, des abréviations grammaticales et des étiquettes pour les champs sémantiques ou les marques d'usage qui sont transparentes pour les deux communautés, (par exemple *Math.* sera transparent à la fois pour les anglophones et pour les francophones, alors que *Chirg.* ne le sera que pour les francophones) et on donne les indicateurs dans la langue de l'adresse. On choisit donc la direction «dictionnaire de thème» dans chaque partie, en supposant que l'utilisateur qui veut traduire dans la langue étrangère a plus besoin d'être aidé que celui qui fait une version (voir 3.2).

Paradoxalement, les dictionnaires bilingues de poche, ceux qui ont les articles lexicographiques les plus simples, presque réduits à des couples de mots, sans indicateurs, marques d'usages etc., sont ceux qui réussissent le mieux à être bidirectionnels. Mais cela se fait au détriment de l'aide apportée aux utilisateurs.

Les dictionnaires bilingues sur CD consultables par ordinateur sont, pour le moment, basés sur les bilingues imprimés (voir Knowles, dans ce volume), et souffrent donc des mêmes imperfections. Mais plusieurs bilingues conçus spécialement pour l'ordinateur sont en cours d'élaboration. Dans ces projets, la bidirectionnalité est à la base de la conception des articles. Chacun a quatre formes différentes selon l'utilisation projetée et la langue maternelle de l'utilisateur : thème et version pour A-phones, thème et version pour B-phones. On peut aussi concevoir un seul texte très complet (du type de celui qui avait été envisagé par Iannucci), mais qui n'apparaît pas tel quel sur l'écran : l'utilisateur pourra choisir de voir seulement ce qui lui sert pour l'opération qu'il, ou elle, est en train d'effectuer.

Les chercheurs qui travaillent dans le cadre du projet européen «COMPASS» préparent des dictionnaires «sensibles au contexte» et en ligne, qui sont conçus pour aider les francophones à comprendre l'anglais, et les anglophones à comprendre l'allemand. Ils constatent que les dictionnaires imprimés qui se veulent bidirectionnels donnent trop d'informations à qui veut décoder seulement, et ils proposent donc de réduire les articles bidirectionnels de plus de la

moitié. B.T.S. Atkins a par exemple réduit l'article *market* dans le *Oxford-Hachette French Dictionary* (1994).

Extrait 1 (Atkins 1994)

Version 1 : article de l'*Oxford-Hachette French Dictionary*

**market** /'ma:kit/ **I n 1** Econ (trading structure) marché *m*; **the art/job/property** ~ le marché de l'art/du travail/de l'immobilier; **the ~ in tea/sugar, the tea/sugar** ~ le marché du thé/du sucre; **at ~ (price)** au prix du marché; **cars at the upper ou top end of the** ~ les voitures haut de gamme; **to be at the upper end of the** ~ [*company*] être au premier rang du marché; **to put sth on the** ~ mettre qch sur le marché; **to be in the ~ for sth** chercher (à acquérir) qch; **to come onto the** ~ [*goods, product*] arriver sur le marché; **2** Comm (potential customers) marché *m* (**for** pour); **domestic/foreign** ~ marché intérieur/extérieur; **the Japanese/French** ~ le marché japonais/français; **a good/poor/steady ~ for** une demande forte/faible/stable de; **it sells well to the teenage** ~ ça se vend bien aux adolescents; **a gap in the** ~ un créneau, un besoin du marché; **3** (place where goods are sold) marché *m*; **flower/fish** ~ marché aux fleurs/halle faux poissons; **covered/open air** ~ marché couvert/en plein air; **to go to** ~ aller au marché; **4** Fin (stock market) Bourse *f*; **to play the** ~ spéculer.

**II modif** Comm. Econ [*share*] de marché; [*conditions, rates, trend*] du marché.

**III vtr 1** (sell) commercialiser, vendre [*product*]; **2** (promote) lancer or mettre [qch] sur le marché.

**IV vi** US **to go ~ing** faire des courses.

**V v refl to ~ oneself** se vendre.

Version 2 : article réduit

**market** /'ma:kit/ **1** marché **cars at the upper ou top end of the** ~ les voitures haut de gamme; **at the upper end of the** ~ au premier rang du marché; **to be in the ~ for sth** chercher (à acquérir) qch; **a good/poor/steady ~ for** une demande forte/faible/stable de; **it sells well to the teenage** ~ ça se vend bien aux adolescents; **a gap in the** ~ un créneau, un besoin du marché; Fin Bourse, **to play the** ~ spéculer. **1** commercialiser, vendre **2** lancer or mettre sur le marché. US **to go ~ing** faire des courses. **to ~ oneself** se vendre.

Atkins a éliminé les usages qui constituent un problème pour l'encodage, mais qui sont tout à fait simples à décoder, ne laissant que les exemples dont le sens ne peut pas être déduit par les utilisateurs francophones, à cause de leur idiomaticité. Elle a aussi enlevé les détails morphologiques (masculin, féminin), les indicateurs (sell, promote etc.), les informations sur les domaines d'emploi (Fin pour finance), qui sont généralement inutiles pour le décodage. Comme le dit très clairement Hausmann (1994:13) «Le dictionnaire de version est obligé de traduire tout ce qui, dans la langue étrangère, n'est pas transparent (...) En revanche, il peut faire l'économie de tout ce qui relève de la compétence en L1, puisque l'utilisateur est compétent dans cette langue qui est la langue d'arrivée du dictionnaire».

## **2 Typologie des dictionnaires bilingues : la macrostructure**

Une typologie des dictionnaires bilingues basée sur les caractéristiques de la macrostructure (nombre des adresses, nature, disposition) permet de distinguer des dictionnaires généraux plus ou moins encyclopédiques, des dictionnaires abrégés, de poche, des dictionnaires spécialisés, ainsi que des ouvrages sémasiologiques et des ouvrages onomasiologiques, en fait les mêmes catégories que pour la typologie des monolingues. Ces catégories étant aisément repérables par les utilisateurs, j'en parlerai brièvement.

### **2.1 Dictionnaires généraux et dictionnaires abrégés**

Les dictionnaires bilingues généraux sont les dictionnaires scolaires par excellence, d'ordinaire avec deux parties reliées dans un seul volume comptant entre 1600 et 2200 pages. Ces dictionnaires sont constitués dans la partie français > langue étrangère (par exemple) par une macrostructure (ou ensemble des adresses, en général classées par ordre alphabétique) formée par le lexique central de la langue française et par un certain nombre de technicisms tirés des différents domaines spécialisés. En outre, si l'ouvrage n'est pas destiné aux francophones mais aux étrangers, on trouvera dans cette partie les archaïsmes utilisés dans les classiques de la littérature française et un certain nombre d'adresses qui renvoient d'une forme irrégulière ou rare à l'article où cette forme est expliquée.

Par rapport à la macrostructure des dictionnaires monolingues généraux, les bilingues généraux devraient contenir aussi un certain nombre de noms propres de personnes et de lieux, surtout quand ces noms propres diffèrent d'une langue à l'autre. Beaucoup de bilingues n'englobent pas les noms propres dans la macrostructure principale du dictionnaire, mais les présentent dans une partie séparée, sous forme de listes, qui en donnent la graphie, mais négligent souvent la prononciation.

Enfin, dans la partie qui sert de dictionnaire de version, on trouve toujours des mots qui désignent des aspects typiques de la civilisation : par exemple, dans un dictionnaire français > langue X conçu pour des locuteurs X-phones, on trouvera des termes de la viticulture et de la vinification. Dans la partie langue X > français les mêmes termes peuvent constituer des adresses si la communauté parlant la langue X produit aussi du vin ou au moins connaît les procédés de vinification et les a nommés.

On appellera «dictionnaires bilingues abrégés» les ouvrages comprenant 1000 pages environ. Le public de ces dictionnaires est avant tout celui de personnes ayant des besoins «pratiques» de communication. Leur macrostructure ne contient pas d'archaïsmes, retient un nombre plus réduit de technicisms et surtout les gloses sont allégées. Il y a moins d'équivalents, et moins d'exemples.

L'extrait 2 montre deux versions de l'article *réseau* : tout d'abord l'article tiré d'un dictionnaire de 2160 pages et 127 000 articles, puis l'article abrégé provenant d'un dictionnaire de 1380 pages et 75 000 articles. On notera que tout ce qui correspondait à *reticolo*, et à *ant.(ico)*, (forme ancienne de *reticella*) a été supprimé. Les indications grammaticales ont été maintenues, parce que le dictionnaire abrégé ne sert pas seulement pour décoder (contrairement à l'exemple de l'extrait 1, version 2).

Extrait 2 (a : Garzanti 1992, et b : Garzanti 1994)

a)

**réseau** [Rezo] (pl. *-aux*) *s.m.* **1** *rete* (f.) : *aérien, hydrographique, routier, urbain, rete aerea, idrografica, stradale, urbana*; – *de chemins de fer* (o *ferroviaire*), *rete ferroviaria* // (tel.) : – *téléphonique, rete telefonica*; – *automatique, teleselezione* // (tv) : – *câblé, rete (televisiva) via cavo*; – *sauvage, emittente (televisiva) pirata* // (cine., comm.) – *de distribution, rete di distribuzione* // (elettr.) – *de portes logiques, rete di porte logiche* // (inform.) : – *en anneau, rete ad anello*; – *en bus, rete bus*; – *local à jeton, rete Token Ring; ordinateur à – neuronal, computer, elaboratore a rete neuronale* (o *neurale*) // (fig.) : – *d'espionnage, rete di spionaggio*; – *d'entreprise, rete aziendale*; *un inextricable – de petites ruelles, un inestricabile dedalo di viuzze* **2** *reticolo* // (fis., min.) – *cristallin, reticolo cristallino* // (ottica) – *de diffraction, reticolo di diffrazione* // (med.) – *de veines, reticolo venoso* **3** (*arti grafiche*) *retino* **4** (*ant.*) *reticella* (f.), *retina* (f.).

b)

**réseau** [Rezo] (pl. *-eaux*) *s.m.* **1** *reticolo* **2** (*estens.*) *rete* (f.) : – *de chemin de fer, rete ferroviaria*; – *de distribution, rete di distribuzione*; – *d'espionnage, rete di spionaggio* // – *téléphonique, rete telefonica*; (tv) – *câblé, rete via cavo*; – *sauvage, emittente (televisiva) pirata* // *inextricable – de petites ruelles, dedalo di viuzze*.

## 2.2 Dictionnaires bilingues de poche

Ces dictionnaires de 400-600 pages sont ceux qui se vendent le plus. Ils sont avant tout conçus pour le voyage. Ils étaient autrefois d'une qualité très médiocre, mais on en trouve aujourd'hui qui sont fort bien faits, attentifs aux exigences de la clarté typographique et à la nécessité d'aider les lecteurs à distinguer les différentes traductions des mots polysémiques.

Les bilingues de poche électroniques sont pour le moment assez rudimentaires. On cherche à les améliorer, mais le fait qu'on ne puisse lire qu'une ligne à la fois interdit les articles complexes. Dès que les progrès technologiques permettront d'avoir des écrans de visualisation plus larges, bon marché quoiqu'insérés dans des appareils qui se tiennent dans la paume de la main, et un système fiable et économique de recharge des batteries, ces bilingues électroniques sont destinés

à connaître une expansion dans le monde entier. Il y aura alors plus de concurrence entre les éditeurs, et on peut espérer qu'il en résultera une amélioration de la qualité des produits.

### 2.3 Dictionnaires bilingues spécialisés

Par rapport à la macrostructure d'un dictionnaire général, la macrostructure d'un dictionnaire spécialisé est beaucoup plus réduite, mais elle contient des termes que les dictionnaires généraux n'ont pas. Les adresses sont plus souvent des syntagmes, et sont la plupart du temps des noms. Il n'y a en général pas d'entrées homonymiques.

La microstructure de ces bilingues spécialisés ne contient en général pas d'informations grammaticales ou d'exemples traduits, mais seulement des collocations et les contextes les plus courants du mot vedette. On a vu apparaître depuis peu des dictionnaires phraséologiques spécialisés bilingues. L'article ci-dessous est tiré d'un ouvrage réalisé à partir d'un dépouillement manuel, mais grâce aux corpus électroniques bilingues ou plurilingues ce type de dictionnaires pourrait devenir courant.

Parmi les dictionnaires spécialisés plus typiquement linguistiques il y a les dictionnaires de faux-amis. Leur macrostructure et la forme de leurs articles sont différentes, parce que les faux-amis forment des «grappes» des mots. Il faut décrire le vrai signifié d'un mot pour montrer qu'il n'est pas le faux suggéré par une ressemblance superficielle avec un mot de la langue maternelle de l'utilisateur. Il y a donc une partie de la microstructure qui rappelle celle des monolingues, mais il faut aussi donner l'équivalent des mots. Parfois on donne des mots presque synonymes ou corrélés pour en délimiter le signifié; on devrait alors donner l'équivalent pour chacun de ces presque synonymes ou corrélés.

Extrait 3 (article *saluto* tiré de Boch, 1988)

**saluto** s.m.

Negli scambi epistolari ai saluti corrispondono *les salutations*, e non *les saluts*: *la prego di gradire i miei più distinti saluti*, **veuillez agréer, je vous prie, mes salutations les plus distinguées.**

Ma sulle cartoline che si spediscono agli amici dal mare o dai monti (approfittando magari di una giornata piovosa quando le ferie sono ormai agli sgoccioli) si scriverà: **un bonjour de Capri** (oppure, più affettuosamente, **amitiés de Capri**).

*Le salut* è il **saluto** solo quando se ne definisce il tipo: **le salut militaire**, *il saluto militare*. E come interiezione, *salut!* corrisponde a **salve!** o, se vogliamo, a **ciao!**: **salut, les gars!**, *salve, ragazzi!*

□ Si noti la locuzione: **à bon entendeur, salut**, *a buon intenditor poche parole*.

## 2.4 Dictionnaires sémasiologiques et onomasiologiques

Les ouvrages sémasiologiques partent de la forme du signe pour atteindre le contenu, le signifié. Les dictionnaires bilingues qui ordonnent leurs adresses en suivant l'ordre alphabétique – c'est-à-dire la plupart – sont sémasiologiques. Il y a toutefois une tradition d'ouvrages bilingues, et souvent plurilingues, qui ont toujours organisé leurs adresses selon le sens, donc de façon onomasiologique : il s'agit des vade-mecums pour marchands et voyageurs qui à partir du Moyen-Âge ont regroupé les mots autour de centres d'intérêt (les repas, les vêtements, l'auberge, les transports, la famille, les jeux, etc.), proposant soit des items bilingues, soit des morceaux de dialogues.

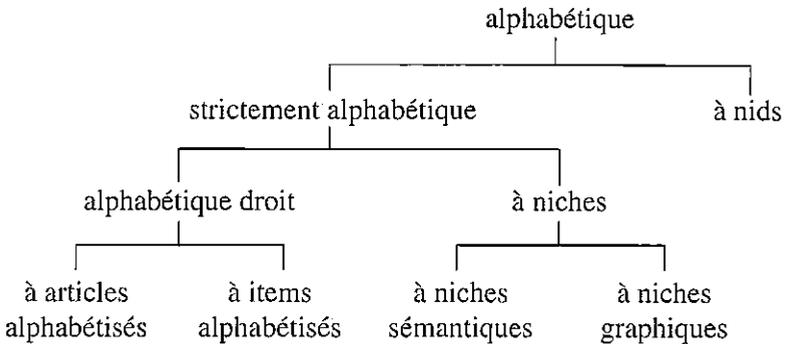
Mais, exception faite pour ce genre d'ouvrages bilingues qui ne sont pas de véritables dictionnaires, l'ordre alphabétique a toujours dominé la lexicographie bilingue des langues qui ont des écritures alphabétiques. C'est donc avec curiosité qu'on accueille le *Cambridge Word Routes Anglais-Français. Lexique thématique de l'anglais courant* (1994). Conçu pour les apprenants francophones, ce dictionnaire onomasiologique bilingue aura aussi des éditions pour apprenants espagnols, italiens, grecs et catalans. Comme on peut le constater dans l'article *tension* (extrait 4), il s'agit d'un dictionnaire synonymique avec définitions, exemples, collocations et illustrations. Il donne des informations sur les restrictions d'emploi (par exemple «informel» sous *het up*) et sur les comportements grammaticaux (par exemple «(non) dénombrable» sous *tension*, «toujours après v(erbe)» sous *on edge*). Naturellement le *Cambridge Word Routes* a aussi des listes alphabétiques des mots français et anglais traités.

Extrait 4 (*Cambridge Word Routes*, 1994)

**tension** *n* **1** *ni* [inquiétude nerveuse] *tension nervous tension* tension nerveuse  
*Tension is mounting as the time for the announcement draws near.* La tension monte au fur et à mesure qu'on se rapproche de la date de l'annonce. **2** *nid* [atmosphère désagréable] *tension international tension* tension internationale  
*racial tensions in inner-city areas* tensions raciales dans les centres urbains **3** *ni* [rigidité] *tension a cable under tension* un câble sous tension

## 2.5 Différents types de macrostructures alphabétisées

On empruntera à Hausmann (1994:25) le schéma suivant.



Le «nid» est un bloc lexicographique qui regroupe des mots composés ou dérivés à la suite du mot de base. À l'intérieur du nid les composés ou dérivés sont disposés par ordre alphabétique. La structure à nids oblige parfois à s'écarter de l'ordre alphabétique strict de la macrostructure : on peut avoir des composés ou dérivés qui devraient être précédés, du point de vue de l'ordre alphabétique, par des mots qui n'appartiennent pas à la famille lexicale du mot base.

Le classement macrostructurel à nids, en soulignant les liens dérivationnels, aide à l'apprentissage du vocabulaire, mais dans le cas de la version il est, selon les termes de Hausmann, «un péché mortel, puisqu'il suppose de la part de l'utilisateur une compétence d'analyse sémantique qui, s'il l'avait, le dispenserait de se servir du dictionnaire. (...) Demander à l'utilisateur de chercher le mot *Bauzeit* avant le mot *Bauch*, c'est se moquer de lui» (Hausmann, 1994:22). La macrostructure à nids était répandue il y a une quinzaine d'années, surtout dans les bilingues avec l'allemand. Elle a désormais été remplacée par sa variante plus raisonnable, c'est-à-dire la macrostructure à «niches» sémantiques, ou plus radicalement par une macrostructure strictement alphabétique à articles alphabétisés.

La «niche» est un nid qui laisse place à l'insertion d'un article complet, qui n'a rien à faire avec la base, là où l'ordre alphabétique strict l'impose. Les niches sémantiques ne regroupent seulement les mots dérivés du mot-base ou composés avec lui. On en trouvera un exemple dans le *Langenscheidts Schulwörterbuch Deutsch-Französisch*. Certains dictionnaires, pour gagner de la place, avaient adopté des niches purement graphiques (par exemple, *égoïsme*, *-ïste*, *-rger* (égorger), *-ut* (égoût), dans la même niche), mais cette macrostructure donne une fausse idée des familles de mots, et on ne la voit presque plus, même dans les dictionnaires de poche.

La macrostructure à articles alphabétisés est la plus répandue aujourd'hui dans les dictionnaires généraux bilingues de langues qui ont une graphie alphabétique. Elle est «droite» parce que l'oeil a du mal à suivre le parcours sinueux imposé par les nids ou les niches.

La macrostructure alphabétisée, qui insère comme des adresses des syntagmes qui d'ordinaire font partie de la phraséologie, est utilisée surtout dans les dictionnaires spéciaux. Son avantage consiste dans la facilité de repérage, mais elle est sans objet pour les dictionnaires en ligne ou sur CD, puisque ces outils permettent une recherche rapide sur tout le dictionnaire, dans le corps du texte des articles et non pas sur les vedettes seulement.

### 3 Typologie des dictionnaires bilingues : la microstructure

La microstructure d'un dictionnaire bilingue comprend en général :

- 1 la tête de l'article. Dans un dictionnaire de thème on aura l'adresse suivie de sa catégorie grammaticale et d'autres informations grammaticales (comme la flexion); dans un dictionnaire de version l'adresse est souvent accompagnée de ses variantes graphiques et de sa prononciation;
- 2 les traductions ou équivalents de l'adresse. Si l'adresse correspond à un élément polysémique, les équivalents sont numérotés ou divisés par un moyen typographique et l'entrée contiendra des indicateurs, des abréviations qui indiquent le niveau de langue ou le champ sémantique, etc., pour aider l'utilisateur à choisir la bonne traduction;
- 3 les exemples, les collocations, les combinaisons phraséologiques, c'est-à-dire la partie où l'on montre le mot-vedette dans des contextes d'usage.

Les dictionnaires diffèrent surtout dans la manière dont ils disposent les contextes dans l'article. Cela peut provenir soit du respect de traditions lexicographiques nationales soit de l'obéissance à des impératifs nés de la rencontre entre ces traditions et des caractéristiques morphologiques des langues décrites. On a vu (2.4) que l'allemand, où la composition est le moyen principal d'enrichissement lexical, s'est orienté vers des macrostructures à nids ou à niches. Dans des langues comme le français, l'italien, l'espagnol, qui ont une morphologie dérivationnelle très vivante et qui présentent des syntagmes lexicalisés avec le modificateur à droite de la tête là où l'allemand a des mots composés et l'anglais des composés ou des syntagmes lexicalisés avec modificateur à gauche, la macrostructure à nids ou à niches a été moins répandue<sup>3</sup>. Par contre dans la

---

3 La macrostructure à nids ou à niches sert à réunir, à l'intérieur d'une disposition sémasiologique, les mots formés par la même base. Si la langue construit ses syntagmes avec le modificateur à gauche, une macrostructure à nids ou à niches pourra regrouper seulement une partie des membres de la famille lexicale, mais elle ne pourra jamais avoir les syntagmes avec la base à droite de têtes différentes.

lexicographie des langues romanes les mots composés et les syntagmes lexicalisés, jusqu'à une période récente, ont fait partie de la section phraséologique des articles. Au cours des dernières décennies la lexicographie française bilingue a commencé à adopter un plus grand nombre d'adresses formées de deux ou trois mots. Les lexicographes de l'italien et de l'espagnol sont encore fidèles à l'idée que l'adresse doit être un seul mot.

Plus les articles qui ont comme adresse un syntagme lexicalisé sont nombreux, plus maigres deviennent les articles concernant les mots très productifs de syntagmes. La lexicographie bilingue a toujours eu plus d'intérêt encore que la lexicographie monolingue à réduire les articles longs, pour les rendre moins complexes et plus lisibles. Une première solution a été trouvée avec l'augmentation des adresses homonymiques et des adresses formées de plus d'un mot; une deuxième consiste à faire «remonter» tous les équivalents le plus près possible de l'adresse, tout en cherchant des modalités pratiques d'organisation des unités phraséologiques.

On a démontré (Marello, 1989:77-98) que les microstructures des dictionnaires oscillent entre trois solutions. La première consiste à regrouper un maximum de phraséologie sous chaque équivalent : c'est à ce type qu'appartiennent les articles *market* dans le *Oxford-Hachette* (extrait 1) et *réseau* dans le *Garzanti* 1992 (extrait 2). La deuxième consiste à distribuer seulement une partie des contextes sous les sections numérotées et à présenter à la fin de l'article une section phraséologique pour les expressions idiomatiques, les proverbes et parfois pour les composés allemands et anglais. Dans la troisième, enfin, le dictionnaire fournit un minimum de contextes à proximité des équivalents, qui sont alors suivis d'une section syntagmatique détaillée, plus ou moins clairement organisée<sup>4</sup>.

Les deux premiers types de microstructures sont les plus répandus et ceux qui ressemblent le plus aux microstructures des monolingues. Cela ne veut pas dire qu'ils sont ceux qui répondent le mieux aux exigences des utilisateurs de bilingues. Les études expérimentales sur l'emploi des dictionnaires bilingues ont montré que l'utilisateur moyen arrête de lire l'article dès qu'il/elle trouve un équivalent à peu près convenable et ne poursuit pas sa recherche pour trouver un meilleur équivalent. Plus l'article est long, plus les équivalents sont éloignés de

---

(suite note 3)

Si l'on compare la suite des mots allemands formés par *der Brief* *lettre*, à partir de «Briefabholung» *levée des lettres*, en passant par «Briefannahme», «Brieffaufgabe», «Brieffaufschritte», «Briefbogen», «Briefbote», «Brieffeinwurf», jusqu'à «Briefwechsel» *correspondance* ou les adresses anglaises composées ou formées par deux mots dont le premier est *mail* *poste* («mailbag», «mailboat», «mailbomb», «mailbox», etc.), on verra que les équivalents français (ou italiens ou espagnol) ne seront jamais proches dans l'ordre alphabétique.

4 Dans ces dictionnaires, on assiste à l'adoption de microstructures différentes pour les articles très longs, tels que «faire», «être», «avoir», «tenir», «devoir» etc. Les dictionnaires abrégés et de poche ont moins de contextes et en général ils n'ont pas de sections phraséologiques : ils distribuent tout le matériau dont ils disposent à proximité de chaque équivalent.

l'adresse par des contextes, plus le risque d'une telle pratique augmente. Dans une microstructure du premier ou du deuxième type, «L'inconvénient majeur (...) réside en ceci que l'utilisateur ne peut détecter le syntagme recherché avec suffisamment de rapidité que si sa compétence linguistique conduit au même découpage sémantique de l'entrée lexicale en question que celle du lexicographe qui a rédigé l'article. C'est donc un mode d'organisation qui peut à la rigueur convenir au dictionnaire de thème, mais qui est carrément inadéquat dans le cas du dictionnaire de version» (Al, 1991:2831).

En effet pour qui veut traduire dans sa langue maternelle un texte écrit en langue étrangère une microstructure du troisième type est beaucoup plus pratique (voir 3.2.2 l'article tiré du *Van Dale* 1983 [extrait 6], qui donne très rapidement un cadre pour tous les équivalents). C'est seulement si l'on n'est pas satisfait de ces équivalents, ou si l'on veut être rassuré par un contexte, que l'on aborde la section syntagmatique. Les macrostructures du premier et du deuxième type obligent parfois à lire tout l'article avant de trouver ce que l'on recherche<sup>5</sup>.

L'utilisateur étant toujours pressé, il est vital d'établir avec soin l'ordre des significations, quel que soit le type de microstructure. Pour cela on ne peut pas se contenter de l'analyse des signifiés des mots polysémiques proposée par les dictionnaires monolingues. Même les bons dictionnaires, qui privilégient les critères logiques (du général au particulier, du propre au figuré) et la fréquence (du signifié le plus fréquent au plus rare), ne peuvent servir que de point de départ, et le résultat doit être adapté en fonction de l'analyse contrastive de la paire de langues. L'anisomorphisme des langues naturelles correspond à des découpages différents du «monde extralinguistique», qui se réfléchissent dans le découpage des sens d'un article de dictionnaire monolingue ou bilingue. Par exemple, l'italien ne fait pas de distinction lexicalisée entre *fleuve* et *rivière*, et les dictionnaires monolingues italiens dans leurs définitions de *fiume* ne font donc pas mention explicite du fait qu'un cours d'eau atteint ou non la mer; seul un dictionnaire bilingue de thème italien > français présente une microstructure avec un indicateur qui sépare les deux équivalents.

### 3.1 Dictionnaires bidirectionnels

Nous avons déjà noté (1.2) que la bidirectionnalité de la plupart des dictionnaires en deux parties est partielle. Ils offrent deux moitiés qui visent surtout à aider pour le thème, partant de l'hypothèse que traduire en langue étrangère est plus difficile que traduire vers sa propre langue.

5 Un des grands avantages des dictionnaires consultables par ordinateur réside justement dans le fait que l'utilisateur peut demander la traduction d'un syntagme et l'obtenir sans se soucier de l'endroit où le lexicographe l'a placé.

Un dictionnaire comme *Le Robert & Collins Français - Anglais Anglais - Français* (1993) est un bon exemple de dictionnaire bidirectionnel. L'article *pleasure* (extrait 5) aide les anglophones en leur disant que le mot dans le sens de «satisfaction» se traduit par *plaisir*, qui est masculin, et en donnant la traduction d'expressions comme *they took great pleasure in his success*, qui est parfaitement compréhensible pour un français, mais qu'on ne peut pas traduire mot à mot en anglais (la traduction suggérée est : «ils se sont réjouis de son succès»). La traduction de locutions figées comme *we await your pleasure* («nous sommes à votre entière disposition») dans des lettres commerciales est utile à la fois aux francophones et aux anglophones. Par contre l'étiquette «frm» (formal) placée avant toute une série d'expressions anglaises sert seulement aux francophones qui traduisent vers le français; pour savoir si les traductions françaises proposées sont du même niveau de langue, l'anglophone doit consulter l'article *plaisir*. L'indication «Non C» (non dénombrable) qui accompagne l'équivalent b est superflue pour les anglophones et presque inutile ici aux francophones qui par contre en auraient besoin, pour traduire vers l'anglais, à l'article *plaisir*. Cette indication est là parce qu'elle se trouve dans tous les dictionnaires monolingues anglais et que les dictionnaires bilingues ont du mal à renoncer complètement à décrire une langue, même si ça n'est pas leur objectif premier.

Extrait 5 (*Robert & Collins*, 1993)

**pleasure** ['pleɪzə] **1 n a** (*satisfaction*) plaisir *m.* **with ~** (*with enjoyment*) *listen etc* avec plaisir; (*willingly*) *do, agree, help* avec plaisir, volontiers; **one of my greatest ~s** un de mes plus grands plaisirs, une de mes plus grandes joies; **it's a ~ !, my ~ !, the ~ is mine !** je vous en prie ! **it's a ~ to see you** quel plaisir de vous voir !; **it has been a ~ to talk to you** j'ai eu beaucoup de plaisir à parler avec vous; **it gave me much ~ to hear that...** cela m'a fait grand plaisir d'apprendre que...; **if it gives you any ~** si ça peut vous faire plaisir; (*frm : at dance*) **may I have the ~ ?** voulez-vous m'accorder cette danse ?; (*frm*) **may we have the ~ of your company at dinner ?** voulez-vous nous faire le plaisir de dîner avec nous ?; (*frm*) **I have ~ in accepting...** j'ai l'honneur d'accepter...; (*frm*) **Mrs A requests the ~ of Mr B's company at dinner** Mme A prie M. B de lui faire l'honneur de venir dîner; (*frm*) **Mr. and Mrs. A. request the ~ of the company of X at the marriage of their daughter...** M. et Mme A. seraient très honorés de la présence de X au mariage de leur fille...; **he finds or takes great ~ in chess** il trouve *or* prend beaucoup de plaisir aux échecs; **what ~ can you find in doing that?** quel plaisir pouvez-vous trouver à faire cela ?; **to take great ~ in doing** éprouver *or* avoir *or* prendre *or* trouver beaucoup de plaisir à faire; (*pej*) se complaire à faire; **they took great ~ in his success** ils se sont réjouis de son succès; **it takes all the ~ out of it** ça vous gâche le plaisir; **has he gone to Paris on business or for ~ ?** est-il allé à Paris pour affaires ou pour son plaisir ?; **a life of ~** une vie de plaisirs; *see* **business**.

**b** (*NonC : will, desire*) bon plaisir, volonté *f.* **at ~** à volonté; **at your ~** à votre gré; (*Jur*) **during the Queen's ~** aussi longtemps qu'il plaira à Sa Majesté, pendant le bon plaisir de la reine; (*Comm*) **we await your ~** nous sommes à votre entière disposition.

**2 comp** ► **pleasure boat** bateau *m* de plaisance ► **pleasure craft** (*collective*) bateaux *mpl* de plaisance ► **pleasure cruise** croisière *f*, (*short*) promenade *f* en mer *or* en bateau ► **pleasure-loving** qui aime le(s) plaisir(s) ► **the pleasure principle** (*Psych*) le principe de la recherche du plaisir ► **pleasure-seeker** hédoniste *mf* ► **pleasure-seeking** *adj* hédoniste ► **pleasure steamer** vapeur *m* de plaisance ► **pleasure trip** excursion *f*.

### 3.2 Dictionnaires monodirectionnels

Au cours des dix ou quinze dernières années le nombre des dictionnaires qui se disent ouvertement monodirectionnels a augmenté. On en trouve même parmi les dictionnaires de poche. *Le Dictionnaire de l'anglais d'aujourd'hui* de Girard, Barrie, Chaptal et Yvinec (1982) est pour les francophones qui veulent traduire de l'anglais vers le français. Les deux tiers de ses 633 pages sont occupés par la partie anglais > français, la partie français > anglais n'étant qu'un complément. Comme l'expliquent les auteurs dans l'introduction «La partie français > anglais ne comporte pas d'exemples. Ils seraient inutiles pour un francophone. Mais cette économie considérable a permis d'étoffer la partie anglais > français à laquelle il est vivement conseillé de se reporter pour choisir entre les mots proposés celui qui correspond le mieux à ce que l'on cherche». Le parti-pris des auteurs selon lequel dans un dictionnaire pour francophones les exemples sont inutiles dans la partie français > anglais, se justifie uniquement si l'on considère ce dictionnaire comme une sorte de monolingue bilinguisé, un dictionnaire qui sert pour lire l'anglais ou – marginalement – pour produire des textes en anglais. A vrai dire, ceci est plus une déclaration d'intentions qu'une description de la réalité, mais il est tout de même intéressant de voir comment les théorisations metalexigraphiques sur la direction des bilingues ont envahi, sinon le contenu véritable des articles, au moins les introductions de dictionnaires, même de dimensions modestes.

On peut être certain d'avoir entre les mains un dictionnaire monodirectionnel pour Italiens quand on trouve des articles comme les suivants :

**eutichianésimo** *sm* (*stor. relig.*) eutychieisme [-kia-]

**scultùra** *sf* sculpture [*skylyr*] : **scultura in legno**, sculpture sur bois (Ferrante-Cassiani 1991)

On le comprend par la langue des indicateurs de champ sémantique, mais surtout par l'indication de la prononciation qui est là où les Italiens en ont besoin, après l'équivalent, et non pas près de l'adresse dans la partie français > italien, comme d'ordinaire dans les autres dictionnaires.

## La monodirectionnalité extrême

Le dictionnaire *Van Dale Frans-Nederlands* paru en 1983 a montré aux lexicographes comment on pouvait poursuivre jusqu'au bout une vision distributionnelle de la microstructure bilingue. Partant de la constatation que l'étranger qui lit un texte doit s'appuyer sur la forme des signes qu'il rencontre, B. Al et la rédaction Van Dale ont imaginé une microstructure du troisième type, qui présente une organisation efficace de la partie syntagmatique.

Extrait 6 (article *bouger* du *Van Dale*, 1983)

**bouger** [buʒe] ⟨f⟩ ⟨→ t 12⟩

**I** ⟨onov. ww.⟩ **0.1** *bewegen* ⇒ *zich verroeren* **0.2** ⟨inf.⟩ *veranderen* ⟨vooral ontkennend⟩ **0.3** ⟨pol.⟩ *in beweging zijn, komen* ⇒ *in opstand komen* ◆ **3.1** avoir besoin de ~ *behoefte aan beweging hebben* **5.1** ⟨inf.⟩ bouge pas ! (*blijf zitten waar je zit en verroer je niet !*) **6.1** je ne bouge pas de chez moi aujourd'hui *ik doe vandaag geen stap buiten de deur*; **sans** ~ de sa chaise *zonder van zijn stoel op te staan* **6.2** ne pas ~ **au** lavage *kleurecht zijn* ⟨v.stof in de was⟩

¶.2 les prix n'ont pas bougé *de prijzen zijn niet veranderd*

¶.3 le peuple commence à ~ *het volk begint in opstand te komen*;

**II** ⟨ov.ww.⟩ ⟨inf.⟩ **0.1** *verplaatsen* ⇒ *bewegen* **1.1** ~ un meuble *een meubel verplaatsen*; **sans** ~ le petit doigt *zonder een vin te verroeren*;

**III** se ~ ⟨wk.ww.⟩ ⟨inf.⟩ **0.1** *zich verroeren* ⇒ *bewegen* **0.2** *iets doen, ondernemen*.

Les chiffres I, II, III divisent l'article en trois parties : verbe intransitif, transitif et pronominal; 0.1, 0.2, 0.3 montrent les trois signifiés pour lesquels on donne des équivalents. Après le losange noir on trouve les contextes. Tout au long du dictionnaire le premier chiffre arabe signifie «combinaison avec 1. nom, 2. adjectif, 3. verbe, 4. pronom, 5. adverbe, 6. préposition, 7. article, 8. conjonction, 9. interjection du mot vedette avec le signifié.1., 2., etc.». Ainsi le verbe *bouger* dans le signifié.1 se combine avec un verbe (3) dans *avoir besoin de bouger* (3.1), avec un adverbe dans *bouge pas !* (5.1), avec une préposition dans *je ne bouge pas de chez moi aujourd'hui* (6.1). Dans le signifié.2 il se combine avec une préposition dans *ne pas bouger au lavage* (6.2) et fait partie des expressions *les prix n'ont pas bougé* et *le peuple commence à bouger*, respectivement avec le signifié.2 et .3. La petite bannière signifie que le mot vedette dans la combinaison n'a aucun des sens donnés au commencement de l'article.

L'attribution d'un contexte à telle ou telle partie du discours n'est pas toujours automatique, mais la microstructure est très ordonnée et peut vraiment être appliquée à tous les articles<sup>6</sup>. C'est une microstructure pour Néerlandophones qui traduisent à partir du français. Al (1991:2835) explique comment on peut la

6 Pour une discussion détaillée des avantages et des inconvénients de la microstructure Van Dale on peut lire Hausmann (1988:146-148) et Marcello (1989:89-98).

transformer pour en faire un dictionnaire de thème. Dans le thème on part du signifié du mot; la microstructure Van Dale pourrait donner les contextes regroupés sur la base du deuxième chiffre des combinaisons ci-dessus.

Le *Robert & Van Dale* (1988) sert de contrepois parce qu'il se veut monodirectionnel pour francophones. On peut comparer l'article *cadre* dans les deux dictionnaires.

Extrait 7 (a : *Le Robert & Van Dale* 1988, et b : *Van Dale* 1983)

a)

**cadre** <n.m.> **0.1** [encadrement] *lijst* <de> ⇒ *omlijsting* <de (f.)> **0.2** [limites imposées] *kader* <het> ~ *raam* <het> **0.3** [décor] *omgeving* <de (f.)> ◇ *decor* <het> **0.4** [salarie] *leidinggevend functionaris* <de (m.)> ~ *staflid* <het (m.)> het (f.), <plur.> *leidinggevend personeel* <het>. <milit.> *kader* <het> **0.5** [châssis] *raam(werk)* <het> ~ *frame* <het> **0.6** [du béton] *bekisting* <de (f.)> ◆ **1.2** <milit.> ~ de réserve *reservekader* **1.3** ~ de verdure *groene omgeving*; ~ de vie *leefklimaat* **1.4** femme ~ *vrouw met leidinggevende functie* **1.5** ~ d'une porte *deurkozijn* **2.1** ~ amovible *wissellijst* **2.2** <Belg.> *Cadre Spécial Temporaire, CST Bijzonder Tijdelijk Kader, BTK* <Belg.> **2.4** ~ sédentaire *grondpersoneel* **3.1** mettre une photo dans un ~ *een foto inlijsten* **3.2** entrer dans le ~ de ses attributions *tot zijn bevoegdheden behoren*; sortir du ~ de ses fonctions *buiten het kader van zijn bevoegdheden vallen* **3.4** être passé ~ *een leidinggevend functionaris (geworden) zijn* **3.¶** figurer sur les ~s *op de loonlijst staan* **6.2** dans le ~ de in *het kader van*.

b)

**cadre** [kadr] <f> <m.> **0.1** *lijst* ⇒ *omlijsting* **0.2** <fig.> *kader* ⇒ *raam* **0.3** *omgeving* ⇒ *decor, inrichting, milieu* **0.4** *leidinggevend functionaris* ⇒ *staflid, (mv.) leidinggevend personeel, (mil.) kader* **0.5** <techn.> *raam(werk)* ⇒ *omraming, frame, beugel, raamantenne, (biljart, vis.) kader* **0.6** *laadkist* **0.7** <bouw.k.> *bekisting* ⇒ <mijnw.> *stutwerk* **0.8** <scheep.> *hangmat* ⇒ *hangende kooi* ◆ **1.3** ~ de verdure *groene omgeving*; ~ de vie *leefklimaat* **1.4** femme ~ *vrouw met leidinggevende functie*; <mil.> ~ d'activité *kader in actieve dienst*; ~ de réserve *reservekader* **1.5** ~ de bicyclette *fietsframe*; ~ d'une porte *deurkozijn* **1.6** ~ de déménagement *gecapitonneerde verhuiskist* **2.4** ~ moyen *iem. die tot het middenkader behoort*; ~ supérieur *iem. die tot het hoger leidinggevend personeel behoort*; <mil.> ~ noir <cavaleriekader van Militaire School in Saumur>; <mil.> ~ sédentaire *grondpersoneel* **3.1** mettre une photo dans un ~ *een foto inlijsten* **3.2** entrer dans le ~ de ses attributions *tot zijn bevoegdheden behoren*; sortir du ~ de ses fonctions *buiten het kader van zijn bevoegdheden vallen*; cela sortirait du ~ de mon exposé *dat zou buiten het kader van mijn exposé vallen* **3.4** être passé ~ *een leidinggevend functionaris (geworden) zijn* **3.¶** figurer sur les cadres *op de loonlijst staan*; être rayé des cadres *van de loonlijst afgevoerd worden, van zijn functie ontheven worden, ontslagen worden* **6.2** dans le ~ de in *het kader van, passend in het geheel van* **6.4** officier *hors* ~ *gedetacheerd officier*.

La différence n'est pas très remarquable : les indicateurs du *Robert & Van Dale* sont en français; il y a sous 2.2 un item bilingue belge (les francophones de Belgique étant le public visé par le dictionnaire), tandis que dans le *Van Dale* de 1983 il y a davantage de contextes.

Plus passionnante est la comparaison avec l'article *cadre* tiré du *Robert-Signorelli* (voir Extrait 8). Elle donne en effet une idée de ce qu'est une microstructure difficilement lisible.<sup>7</sup> C'est en fait la présentation, dans une macrostructure du deuxième type, des contextes présents dans le *Robert & Van Dale* et dans *Van Dale* 1983.

## 4 Dictionnaires hybrides et bilingués

Le désir de conquérir de nouvelles tranches de marché et aussi l'effort pour mieux servir les utilisateurs a conduit les éditeurs à publier des dictionnaires hybrides, à mi-chemin entre les bilingues et les monolingues.

### 4.1 Le *Robert-Signorelli*

Ce dictionnaire, projeté par Paul Robert lui-même, ne veut pas seulement offrir des traductions, mais aussi décrire les deux langues mises en contact. Il donne donc des définitions complètes et, comme le dit l'introduction, «Au lieu d'analyser chaque mot d'une langue selon les différences de traduction dans l'autre, nous nous sommes basés sur l'analyse interne pour chaque langue, comme dans un dictionnaire unilingue. Cette analyse peut correspondre soit à une même traduction pour plusieurs sens, soit à des équivalences différentes». Les définitions françaises ont été tirées du *Micro Robert* ou du *Petit Robert* et un *Micro Signorelli* avec définitions des mots italiens a été préparé spécialement. Selon la rédaction du *Robert & Signorelli* les définitions servent à mieux choisir le bon équivalent et deviennent elles-mêmes des exemples de langue correcte.

Extrait 8 (article *cadre* du *Robert-Signorelli*, 1981)

**cadre** [kadʁ(ə)] *s.m.* [*bordure entourant une glace, un tableau*] cornice (*f.*): *cadre rectangulaire, rond, ovale* cornice rettangolare, rotonda, ovale; *cadre en bois, peint, doré* cornice di legno, dipinta, dorata; *mettre une peinture dans un cadre* mettere in cornice, incorniciare un dipinto 2 [*châssis fixe*] telaio\* : *le cadre d'une porte, d'une fenêtre* il telaio di una porta, di una finestra ► **chambranle**; *cadre de bicyclette* telaio di bicicletta # *cadre de déménagement* gabbia da imballaggio 3 (*mar.*) branda *f.* 4 (*radio*) antenna *f.* a telaio II (*fig.*) 1

7 Certains trouvent que la présentation des articles du *Robert-Signorelli* est gâchée par la subdivision sur trois colonnes qui rend les articles très longs et étroits.

[ce qui circonscrit, entoure un espace, une scène] ambiente : j'aimerais vivre dans un cadre aussi accueillant mi piacerebbe vivere in un ambiente così accogliente ► **décor, entourage, 2 milieu 2** [arrangement des parties d'un ouvrage] struttura.f., schema ► **2 plan 3 loc.** DANS LE CADRE DE [dans l'ensemble organisé] nel quadro, nell'ambito di : les deux pays ont signé cet accord dans le cadre d'une politique de détente i due paesi hanno firmato questo accordo nel quadro di una politica di distensione; il a agi dans le cadre de ses fonctions ha agito nell'ambito delle sue funzioni **4 (milit.) quadri (pl.)** : les cadres d'un bataillon, d'un régiment i quadri di un battaglione, di un reggimento; cadre de réserve quadri di riserva # un officier hors cadre un ufficiale fuori quadro ◊ (aviat.) cadre sédentaire quadri sedentari, di terra **5 surtout au plur.** [tableau des emplois et du personnel qui les remplit] organico\* (sing.) : figurer sur les cadres figurare in organico; être rayé des cadres essere radiato dall'organico **6 au plur.** LES CADRES [le personnel d'encadrement des entreprises] i quadri : les cadres moyens i quadri, i capi intermedi; gli intermedi; les cadres supérieurs i (quadri) dirigenti ◊ (au sing.) jeune cadre giovane dirigente; c'est un cadre moyen, supérieur è un intermedio, un dirigente (appos.) ♦ une femme cadre una dirigente.

#### 4.2 Dictionnaires monolingues pour apprenants «bilingualisés»

L'idée de bilingualiser un dictionnaire monolingue n'est pas si moderne qu'on le prétend. Un exemple parmi d'autres : le premier grand dictionnaire monolingue le *Vocabolario degli Accademici della Crusca* (1612) a maintenu jusqu'à sa quatrième édition (1729-1738) un grand nombre de mot latins et grecs employés comme indicateurs sémantiques ou définitions synonymiques. Le *Vocabolario* avait un index des mots latins et un index des mots grecs qui permettaient de retrouver l'article italien où ils étaient mentionnés. Les auteurs déclaraient explicitement dans l'introduction que ces renvois étaient conçus pour faciliter l'emploi du dictionnaire aux étrangers cultivés qui connaissaient les langues classiques, et même aux Italiens qui connaissaient le toscan moins que le latin – situation semblable à celle des Français non francophones qui avaient employé le *Dictionnaire françois-latin* (1573) de Estienne pour apprendre le français.

Une caractéristique moderne de la bilingualisation des monolingues pour apprenants est la présence d'un seul mot en langue maternelle pour toutes les significations des articles du monolingue de langue étrangère. Le dictionnaire qui en résulte est utile aux étudiants qui doivent apprendre la langue étrangère et pas seulement traduire. L'objectif de cette bilingualisation est de rendre caduque l'opposition entre l'utilisation du monolingue de langue étrangère (pédagogiquement utile et recommandé par les enseignants) et l'utilisation du bilingue (plus rapide et plus facile, donc préféré par les étudiants). La maison d'édition japonaise Shogakukan a bilingualisé le *Petit Robert* pour les étudiants japonais. Pour les étudiants français qui apprennent l'anglais la maison d'édition

Modulo de Montréal a publié en 1989 un dictionnaire sous le titre de *Password* qui se base sur le *Oxford Student's Dictionary*<sup>8</sup>. Ce dictionnaire est distribué aussi en France (par Belin), tandis que Modulo a préparé des éditions spéciales pour le Cameroun et la Côte d'Ivoire.

Extrait 9 (*Password* 1989)

- scene** [si:n] *noun* **1** the place where something real or imaginary happens : *A murderer sometimes revisits the scene of his crime; The scene of this opera is laid/set in Switzerland.* □ **scène**
- 2** an incident *etc* which is seen or remembered : *She recalled scenes from her childhood.* □ **incident**
- 3** a show of anger : *I was very angry but I didn't want to make a scene.* □ **scène**
- 4** a view of a landscape *etc* : *The sheep grazing on the hillside made a peaceful scene.* □ **tableau**
- 5** one part or division of a play *etc* : *The hero died in the first scene of the third act of the play.* □ **scène**
- 6** the setting or background for a play *etc* : *Scene-changing must be done quickly.* □ **décor(s)**
- 7** a particular area of activity : *the academic/business scene.* □ **scène**

C'est surtout en Israël, où des dictionnaires bilingués sont disponibles sur le marché depuis 1986 que l'on a conduit des expériences sur l'emploi de ces dictionnaires et sur leur efficacité comparée à celle des monolingues pour apprenants et des bilingues. Laufer et Melamed (1994:565-576) concluent que les utilisateurs de niveau élémentaire, moyen ou avancé qui emploient le dictionnaire pour comprendre des mots obtiennent de meilleurs résultats avec le dictionnaire bilingués. Pour ce qui concerne la production de phrases le dictionnaire bilingue est le plus utile pour les utilisateurs de niveau élémentaire; les utilisateurs moyens et avancés obtiennent de meilleurs résultats avec le dictionnaire bilingués.

## 5 Conclusion

Le type de dictionnaire bilingue le plus courant jusqu'à maintenant, c'est-à-dire le dictionnaire général en deux parties, bidirectionnel et imprimé restera encore longtemps le plus répandu, parce qu'il répond aux besoins d'un large public. Les traducteurs professionnels tireront davantage de profit des outils informatiques, c'est-à-dire des banques des données bilingues ou plurilingues. Les étudiants

8 Le premier monolingue d'anglais bilingués a été publié en 1986 par Kernermann Publishing : c'était une version pour Israël de l'*Oxford Student's Dictionary*. Il existe des dictionnaires bilingués pour l'arabe (1987), l'italien (1989), le grec (1989), le polonais (1990), l'espagnol (1991), le portugais (1991), le chinois (1991), le swahili (1992) et d'autres langues encore.

auront à leur disposition un plus grand nombre de dictionnaires pédagogiques : un nombre croissant de langues est en train de se doter de monolingues pour apprenants et de monolingues bilingualisés. Nous n'en sommes qu'aux premiers pas de l'hybridation en matière de dictionnaires et on a l'impression que les nouveautés éditoriales ont avancé beaucoup plus rapidement que la pratique pédagogique.

## Chapitre 3

# LE ROLE DE LA SYNTAXE DANS LA LEXICOGRAPHIE BILINGUE

Brigitte LÉPINETTE

### 1 Introduction

Notre but, dans ce chapitre, est de présenter un type de dictionnaire bilingue qui ne s'appuie pas, comme dans la lexicographie conventionnelle, sur des bases sémantiques (voir, entre autres, Rey-Debove 1971: 28-29), mais dont le principe d'organisation est de nature syntaxique. Dans ce nouveau type de dictionnaire, ce sont les possibilités de combinaison de chaque item lexical avec les autres unités de la phrase qui déterminent l'analyse dans la langue source et, secondairement, l'identification des équivalents de la langue cible. Le cadre de la phrase dans laquelle peut s'insérer l'item est obligatoirement pris en compte lors de la phase de préparation, à la suite d'une analyse distributionnelle inspirée de travaux réalisés dans le cadre du Laboratoire d'Analyse Documentaire et Linguistique (L.A.D.L.).

Les premières bases de ce dictionnaire ont été jetées dans une publication récente (Lépinette, 1994), mais nous insisterons ici sur ce que nous pensons être la nouveauté principale, et nous présenterons une description plus complète de la phase d'analyse, qui doit faire intervenir des distinctions plus précises que celles que nous avons initialement retenues.

Nous ferons en préambule un rapide tour d'horizon historique qui aura pour but de mettre en évidence l'absence relative, dans les dictionnaires bilingues traditionnels, de la syntaxe en tant que principe organisateur de l'analyse lexicographique. Nous examinerons le système contrastif des dictionnaires bilingues, qui n'ont pas profité des recherches récentes dans le domaine de la lexicographie monolingue. Dans un second temps, nous présenterons les grandes lignes de l'article de dictionnaire bilingue tel que nous le concevons, dans la perspective syntaxique dont nous voulons ici montrer l'intérêt.

## 2 Analyse sémantique et analyse syntaxique dans les dictionnaires bilingues traditionnels

### *Les présupposés linguistiques de la lexicographie bilingue conventionnelle :*

Le rôle de la syntaxe : le dictionnaire bilingue constitue un système de mise en relation d'un item lexical appartenant à une langue source (LS) avec un ou plusieurs items lexicaux d'une langue cible (LC), en un rapprochement qui, implicitement, pose comme principe l'identité entre le sémantisme de l'item de LS et celui de l'item (ou des items) de LC donné(s) comme équivalent(s) traductif(s). Le dictionnaire bilingue suppose aussi, d'une manière tout aussi implicite, l'existence d'une organisation lexico-sémantique en LS fondamentalement parallèle à celle de LC. Cette mise en regard de deux unités lexicales appartenant à deux langues différentes concerne uniquement l'axe paradigmatique, excluant la prise en considération des possibilités de combinaisons de ces items de LS et de LC dans des phrases ou dans des syntagmes, sauf dans le cas particulier des séries figées. Il est ainsi possible de constater que, dans la lexicographie bilingue, les environnements et les constructions sont rarement présentés systématiquement, et que, dans les cas où ils le sont, ils apparaissent toujours en tant que simples adjuvants de la description sémantique. Ainsi l'organisation des différents équivalents traductifs en LC d'une unité polysémique de LS a pour base principale la sémantique, et la syntaxe n'y joue qu'un rôle secondaire.

Certes, les informations syntaxico-grammaticales ne sont pas totalement exclues de l'article de dictionnaire bilingue. Dans ce qu'on peut considérer comme un des premiers «véritables» dictionnaires latin-langue vernaculaire, le *Vocabulario de romance en latín* (1495) de A.-E. de Nebrija (recueil espagnol-latin), les entrées ne sont pas seulement traduites; elles sont également accompagnées, dans certains cas, de la mention de la catégorie grammaticale à laquelle elles appartiennent. C'est spécialement le cas pour les adverbes :

**Manojo a manojo** adverbio *manipulativum*

Cependant, ces indications sont évidemment en relation avec l'optique d'encodage de ce dictionnaire qui, pour cette opération, se doit de signaler la classe morphologique et le type de déclinaison (vs. absence de déclinaison dans le cas de l'adverbe que nous venons de citer) dont relèvent les items traduits. Ces informations ne signifient donc pas que l'axe syntagmatique soit pris en compte en tant que tel, mais seulement dans les changements morphologiques qu'il suppose pour chaque unité considérée isolément.

La lexicographie bilingue de deux langues vernaculaires, français-espagnol par exemple, ne suivra pas la direction indiquée par l'humaniste Nebrija. Les dictionnaires bilingues omettront, durant le XVII<sup>e</sup> et XVIII<sup>e</sup> siècle, la mention de la catégorie grammaticale des entrées, en contraste avec la lexicographie mono-

lingue française, qui, se posant ainsi en héritière de la tradition monolingue latine (du *Catholicon* par exemple) depuis Nicot (1606), a utilisé ce point de vue pour classer les unités qui constituent la macrostructure. Tout spécialement, dans le cas des verbes, Nicot mentionnait la sous-classe à laquelle ils appartenaient (actifs, neutres ou pronominaux). Cependant cette spécification – de nature purement syntaxique, cette fois – qui supposait une base pour la différenciation sémantique, ne pouvait être exploitée systématiquement car, dans l’optique de description des sens qui était celle de ces ouvrages, cette division tripartite traditionnelle est loin d’être toujours pertinente.

Dans l’histoire de la lexicographie bilingue, la spécification des possibilités de complémentation des verbes apparaît dans les dictionnaires du XIX<sup>e</sup> siècle (Capmany 1804, dans le cas du domaine français-espagnol), puis elle est intégrée systématiquement à l’ensemble des informations de nature diverse (diachronique, diastatique, pragmatique, etc.), généralement codifiées, que présenteront les dictionnaires bilingues, et ce, sans véritable transformation jusqu’à nos jours. Actuellement, seule la terminologie a changé. Les verbes, abandonnant les étiquettes traditionnelles, apparaissent comme appartenant à la classe des intransitifs, des transitifs ou des pronominaux. Dans certains dictionnaires bilingues français-espagnol récents figurent également les spécifications *impersonnel* et *auxiliaire* (c’est le cas du Larousse 1987, français-espagnol). Cependant, outre les difficultés intrinsèques que soulèvent les analyses auxquelles les lexicographes soumettent chaque entrée en particulier, nous avons conclu à propos du principe même, après un examen de cette question, que :

les distinctions syntaxiques consignées après les entrées ont un niveau de généralités qui les rend inopérantes pour l’analyse du sémantisme de ces termes. Sauf cas exceptionnel, elles ne sont pas directement utiles car elles renvoient à une catégorisation trop large pour rendre compte des significations. (Lépinette, 1992:496)

Ce très rapide survol de quatre siècles de lexicographie bilingue nous permet donc de vérifier que celle-ci n’a changé radicalement ni ses méthodes ni les présupposés qu’elles impliquent depuis les premiers lexiques qui présentaient en contraste le latin et une langue vulgaire ou bien une langue vulgaire en face d’une autre, à l’époque de la Renaissance (Lépinette, 1992:483-7). L’embryon d’analyse syntaxique qui figure dans ces ouvrages ne change pas la nature fondamentalement sémantique de la description de la langue source qui conduira à la mise en regard des unités traductives.

Une remarque s’impose. Le dictionnaire monolingue, dans la presque totalité des cas, ne présente pas, dans sa zone purement sémantique (i.e. la définition), de différence de nature avec le dictionnaire bilingue. Il offre une (des) traduction(s) intralinguale(s), sous forme de paraphrase(s) (la définition), alors que le dictionnaire bilingue présente une (des) traduction(s) interlinguale(s) sous forme d’équivalent(s) en langue cible. Baldinger (1971:138-39) signalait déjà que la différence essentielle entre les dictionnaires monolingues et les bilingues réside

dans le fait que, dans les premiers, l'analyse sémémique est explicite (les définitions) alors que dans les seconds elle est seulement présupposée dans les équivalents.

***Les instruments d'analyse des unités polysémiques en lexicographie conventionnelle :***

Est-il utile de rappeler qu'entre deux langues mises en contraste la relation lexico-sémantique est rarement univoque terme à terme ou «homomorphique» (Gross, 1987:47) ? Le seul cas de *symétrie* totale est celui des unités monosémiques de LS traduites par d'autres unités monosémiques de LC. Il y a donc, dans une lexicographie bilingue qui se veut opératoire, une double nécessité en face des unités dites «polysémiques». Il faut :

- a) mettre au point des instruments d'analyse pour discriminer les différents lexèmes d'un vocable<sup>1</sup>, et
- b) signaler ces instruments de manière explicite, à l'intérieur des articles, parce qu'ils sont à la fois la base des «découpages» effectués et l'explication de ces découpages, constituant ainsi une donnée à fournir à l'utilisateur. En effet, celui-ci doit avoir des critères pour sélectionner de manière adéquate l'item lexical de LC correspondant à l'acceptation de l'item de LS qui est actualisée dans le discours à traduire. Or, les moyens auxquels la lexicographie bilingue conventionnelle recourt en général pour marquer cette délimitation de la zone de recouvrement de chacune des unités (de LC) données comme traductions (par rapport à la totalité de l'aire sémantique du vocable qui constitue l'entrée) sont des ébauches de définitions, des synonymes en LS, des mentions de l'appartenance à un domaine thématique de LS, etc.<sup>2</sup>, instruments majoritairement de nature sémantique, et en outre linguistiquement hétérogènes.

Si la délimitation des différents lexèmes d'un vocable est évidemment un impératif pratique, elle a en même temps une dimension théorique, car elle constitue une analyse sémantique qui est obligatoirement conditionnée par la nature des moyens linguistiques retenus pour la réaliser. Par conséquent, elle ne sera pas unifiée si ces moyens sont hétérogènes. Ajoutons que si l'objectif final de la mise en contraste d'unités de LS et de LC est de constituer un ensemble de données exploitables mécaniquement, il est indispensable que les instruments d'analyse retenus soient homogènes.

---

1 Dans la terminologie que nous emploierons ci-dessous, le lexème est l'unité lexicale considérée dans une seule acceptation, le vocable est l'ensemble des lexèmes, unité lexicale de même forme renvoyant donc à plusieurs acceptations (ou polysémique).

2 Voir aussi Lépinette, 1992 sur cette question des marqueurs et de l'indication de domaines thématiques employés asystématiquement dans la lexicographie bilingue.

Il apparaît donc que, non seulement pour être linguistiquement satisfaisante mais aussi pour prétendre à une utilisation sur machine, la lexicographie bilingue devra utiliser des instruments d'analyse des vocables qui soient unifiés et qui puissent être appliqués de manière systématique, sinon à l'ensemble du lexique, au moins à des classes d'unités composant ce dernier.

Notons encore ici que, lorsque l'entrée est une unité polysémique, le dictionnaire monolingue conventionnel ne procède pas d'une manière radicalement différente de celle du dictionnaire bilingue. La délimitation des sens n'est pas réalisée à l'aide d'un système d'analyse unique. En effet, si traditionnellement cette lexicographie a recouru à la définition (qui ne représente pas en elle-même un système), elle a aussi employé, pour compléter ou remplacer celle-ci, un éventail de moyens encore plus variés que ceux que l'on peut observer dans le dictionnaire bilingue. L'ouvrage monolingue se sert parfois de critères logiques (par exemple en définissant un sens par rapport à un autre : *par extension*), rhétoriques (ex. : *par métaphore*) ou lexicaux (ex. : *archaïsme*), etc. Ainsi la lexicographie monolingue conventionnelle, comme la lexicographie bilingue, emploie des procédures de discrimination entre les lexèmes qui ne sont ni homogènes ni même dans certains cas explicites.

Des études théoriques relativement récentes visant précisément cette délimitation des sens nécessaire dans le dictionnaire proposent des instruments d'analyse à valeur générale, dont l'utilisation permettrait sans doute d'atteindre un certain degré d'homogénéisation dans les procédures de discrimination lexématique. Cependant, dans la lexicographie bilingue et monolingue il y a peu de traces des recherches structurales (analyse sémémique<sup>3</sup>) et encore moins des génératives (Katz et Fodor). Il est vrai que dans les deux cas l'ensemble des traits sémantiques utilisables pour distinguer les lexèmes est mal délimité et ne constitue pas une liste fermée, de sorte que ces modèles ne fournissent pas un cadre d'analyse assez contraignant pour être opératoire en lexicographie. Notons cependant que des analyses influencées par ces théories seraient sans doute plus homogènes, donc plus satisfaisantes théoriquement, que celles qui sont mises en œuvre actuellement, même si elles restent – c'est ce qui concerne directement notre propos ici – de nature exclusivement sémantique<sup>4</sup>, ne tenant aucun compte des comportements syntaxiques pour déterminer l'existence des différents lexèmes constituant un vocable.

Cette nécessité d'une analyse unifiée des items a aussi conduit à la conception et à la réalisation d'instruments spécifiques qui ont pour caractéristique d'être établis sur des bases explicites et homogènes. Ainsi Igor Mel'čuk, dans son *Dictionnaire explicatif et combinatoire du français contemporain* (D.E.C., 1984),

3 Rappelons que dans ce domaine, le *Dictionnaire du français contemporain* de J. Dubois et al. (1966) représente une exception bien connue.

4 C'est l'analyse automatique du sens des mots qui semble avoir repris le flambeau d'une méthode utilisant des marqueurs hiérarchisés et en nombre restreint (voir par exemple et entre autres : Sciarone, 1983:145).

ouvrage monolingue théorique qui fait partie intégrante de la théorie sens-texte, rend compte du contenu de chaque lexème par une définition qui remplit un certain nombre de conditions (Mel'čuk, 1984 et 1986, Mel'čuk *et al.*, 1983) : langage définitionnel univoque, adéquation (i.e. la définition doit être nécessaire et suffisante) et décomposition (i.e. la définition doit seulement comprendre des termes qui sont plus simples que le terme défini). Ces principes d'analyse peuvent, en théorie, être adoptés avec profit par la lexicographie bilingue; c'est ce que nous avons montré (Lépinette, 1989a). Ils peuvent donner lieu également à un système explicite et cohérent qui permet la mise en évidence du degré de différence et d'identité entre les contenus sémantiques de termes qui sont normalement considérés équivalents dans deux langues en contraste (Lépinette, 1989 b).

Mais la méthode de différenciation des lexèmes mise en œuvre par Mel'čuk ne s'appuie pas non plus sur une analyse des possibilités de combinaisons de ces derniers dans la phrase<sup>5</sup>, la finalité du D.E.C. étant évidemment différente. Partie intégrante de la théorie sens-texte, le D.E.C. veut réunir, de manière exhaustive, toute l'information sur un vocable et réaliser l'inventaire des unités lexicales qui sont des cooccurrents fréquents d'un lexème, constituant avec d'autres lexèmes le vocable, entrée de l'article. L'auteur procède à un classement des cooccurrents fréquents selon le type de relations sémantiques – répertoriées et étiquetées sous ce qu'il appelle les «Fonctions Lexicales» (FL) – qui existent entre ce lexème-clé et ses cooccurrents. La perspective n'est pas «typiquement onomasiologique» comme l'affirme R. Vivès (1989 :100), même dans la «zone lexicale» du D.E.C. : il y a rapprochement d'unités du lexique fréquemment associées et, en même temps, analyse du lien sémantique qu'elles entretiennent entre elles<sup>6</sup>.

Parce qu'elle est basée sur le sens elle aussi, la perspective adoptée par Mel'čuk pour séparer ses vocables en lexèmes ne diffère pas radicalement de celle qui est utilisée dans les dictionnaires monolingues, même si l'appareil théorique qui préside à l'élaboration des définitions du D.E.C. ne permet en aucun cas de les assimiler à celles d'un ouvrage lexicographique conventionnel. Dans cette perspective linguistique, l'unité définie est conçue en tant que contenu sémantique mis en valeur par l'ensemble des composantes de la définition, ce qui fait que la considération de ses possibilités d'insertion dans le cadre de la phrase, quand il s'agit de la discrimination des différents lexèmes d'un vocable donné, n'est pas directement pertinente.

5 Il faut souligner que Mel'čuk signale dans l'article de son dictionnaire théorique *D.E.C.* (1984) le régime syntaxique des lexèmes traités. Cependant ce régime n'est pas la base sur laquelle s'effectue l'analyse d'un vocable en différents lexèmes.

6 Il est cependant vrai comme le dit ce même auteur (R. Vivès, 1989:100) que, comme le *Lexique-grammaire*, le D.E.C. «allie de façon originale, une théorie de la description linguistique et la description d'une langue à travers son lexique».

Certains dictionnaires ont pour objectif de fournir à l'utilisateur une description précise d'un trait syntaxique – la possibilité de complémentation – des termes qui constituent leur macrostructure. Ce sont les dictionnaires de valences (pour les verbes) et les dictionnaires combinatoires. Depuis la publication du dictionnaire allemand de valences de Helbig et Schenkel (1968), plusieurs de ces ouvrages ont vu le jour, par exemple et également pour l'allemand, Engel et Schumacher (1976), Sommerfeldt et Schreiber (1974 et 1977), et pour le français Busse et Dubost (1977/1983). Les destinataires des dictionnaires de valences ou combinatoires sont, dans presque tous les cas, des apprenants étrangers de la langue analysée qui ont besoin de cette description à des fins pratiques (encoder correctement des énoncés). Pour l'anglais, il faut citer entre autres *The BBI Combinatory Dictionary of English* (BBI, Benson, Benson et Ilson 1986) qui présente non seulement les collocations grammaticales (constructions des verbes, des noms prédicatifs et des adjectifs) mais aussi les collocations lexicales que l'on peut considérer, sous certains aspects, comme similaires aux cooccurrences de Mel'čuk. Cependant, dans le BBI, les combinaisons apparaissent classées formellement à partir des catégories grammaticales (V + N, Adj + N, etc.) et hors de tout cadre théorique comme dans le D.E.C. Pour l'anglais, il faut signaler aussi la proposition de Herbst (1987). Le domaine de l'espagnol, pour autant que nous le sachions, est absent de ce panorama, du moins pour ce qui concerne les réalisations d'une certaine envergure. En tout état de cause, ces dictionnaires, dont la base est indubitablement de nature syntaxique, décrivent les constructions d'une langue, mais ils le font à la seule fin de présenter ces constructions, et non pour relier syntaxe et sémantique, et encore moins pour traduire de LS en LC.

Contrairement aux méthodes d'analyse sémantique dont nous avons mis en évidence la présence dans les dictionnaires que nous venons d'envisager, le cadre méthodologique général du Lexique-grammaire offre, à notre avis, une base unifiée sur laquelle appuyer la distinction des différents lexèmes d'un vocable de LS. Rappelons que le Lexique-grammaire (dont les bases ont été exposées d'abord dans *Méthodes en syntaxe*, 1975) est né du sentiment de la nécessité pour son auteur, M. Gross, «d'accumuler des données empiriques aussi complètes et détaillées que possible» (1975:45) sur les possibilités du système linguistique. Dans cette optique, la première tâche consistait à décrire les propriétés syntaxiques des items lexicaux d'une langue. Le domaine choisi au départ a été celui des verbes<sup>7</sup> qui admettent une construction complétive et dont on a étudié les propriétés distributionnelles et transformationnelles. On sait que la méthode a été appliquée ensuite à d'autres catégories linguistiques, comme les «verbes composés», les «adverbes composés» qui appartiennent à ce qu'on a appelé en lexicologie les «phrases figées», classe hétérogène dont les caractéristiques syntaxiques n'avaient jamais été décrites. Si ces travaux ont l'ambition d'améliorer les

7 Il y a approximativement 3000 verbes étudiés. Les propriétés distributionnelles et transformationnelles impliquées sont en tout au nombre d'une centaine.

traitements automatisés du langage, ils fournissent également des instruments pour une description lexicographique sur des bases qui ne sont pas sémantiques. Au lieu de prendre en considération les différents signifiés du vocable, l'analyse, dans ce cadre, commencera par distinguer les différents comportements syntaxiques d'un item. Dans un second temps, en partant de l'hypothèse selon laquelle il y a souvent homogénéité entre les comportements syntaxiques et sémantiques (autrement dit, une construction et un environnement lexical donnés correspondent à un lexème et à un seul), il paraît possible d'organiser, en regard de ces unités de LS déterminées par leur comportement syntaxique, la présentation des équivalents traductifs de ces dernières en LC.

Les méthodes mises au point au L.A.D.L. devraient donc permettre de mener à bien les deux tâches principales d'un dictionnaire bilingue linguistiquement adéquat, à savoir l'analyse par des procédures homogènes (dans ce cas de nature grammaticale) des différents lexèmes constituant un vocable en LS, puis la mise en regard de chacun de ces lexèmes d'un équivalent traductif appartenant à LC. Nous précisons ci-dessous les conditions de cette analyse syntaxico-sémantique et les contraintes que l'optique contrastive imposera par rapport aux réalisations effectuées dans la perspective avant tout monolingue qui est celle du L.A.D.L.

### **3 Conditions générales de l'analyse syntaxique contrastive dans une perspective lexicographique**

L'analyse que nous présentons ici renouvelle les bases lexicographiques conventionnelles parce qu'elle considère les unités de LS en tant qu'éléments intégrés dans une phrase susceptibles de se distinguer par leur comportement syntaxique plutôt que par leur contenu sémantique. Les critères de distinction des sous-unités lexicographiques traductives sont les caractéristiques syntaxiques différenciées dans chaque cas. L'unité d'équivalence lexicographique traductive est la phrase entière (libre ou figée) qui inclut cette unité.

Le cadre méthodologique général est celui du Lexique-grammaire des Langues romanes (c'est-à-dire l'ensemble des études syntaxiques réalisées à partir des travaux de M. Gross), dont l'objectif principal, outre les considérations théoriques (observations sur les fonctionnements syntaxiques d'une langue), est l'amélioration des traitements automatisés du langage par la création de banques de données linguistiques exploitables sur machine. Le Lexique-grammaire, nous l'avons dit, offre des instruments grâce auxquels il est possible d'établir la distinction des différents lexèmes d'un vocable de LS. L'analyse, au lieu de prendre en considération les différents contenus sémantiques d'un vocable pour déterminer ses «acceptions», se base sur une distinction des différents comportements syntaxiques du vocable.

Le procédé est justifié par le fait bien connu que différentes constructions supposent, en général, des différenciations sémantiques parallèles. C'était le principe sur lequel s'appuyaient les analyses syntaxiques présentes dans le *Dictionnaire du français contemporain* (D.F.C., Dubois *et al.* 1966), qui visaient à la division de certains vocables (verbes) en lexèmes. Ces observations se sont trouvées confirmées par les analyses de Gross (1975), qui ont permis de constater l'existence d'une homogénéité fréquente entre les comportements syntaxiques et sémantiques. En conséquence, il paraît possible d'organiser, en face des sous-unités lexicales de LS (ou lexèmes) déterminées en fonction de leurs propriétés syntaxiques, la présentation des équivalents traductifs de LC.

Nous avons décrit de manière précise dans l'étude déjà mentionnée (Lépinette, 1994) les conditions de l'analyse syntaxico-sémantique qu'impose l'optique contrastive par comparaison avec la visée monolingue qui est celle des travaux de l'équipe de M. Gross – qui n'a qu'en de très rares occasions considéré deux langues en contraste<sup>8</sup>. Nous nous bornerons ici à en rappeler quelques aspects.

Les entrées de cet ensemble de données bilingues qui a la même finalité traductive que n'importe quel dictionnaire bilingue sont des unités lexicales (*courir, entrer, monter*), traitées selon une méthode qui dépendra de la classe (verbe, substantif, adjectif) à laquelle ils appartiennent (nous ne nous occupons dans ces pages que du verbe). Les entrées ne se présentent pas de la même manière que dans un dictionnaire conventionnel, c'est-à-dire séparées en fonction de leur contenu sémantique en une série de sous-unités discriminées intuitivement et associées, pour chacune d'entre elles, à des équivalents interlinguaux, multiples dans la plupart des cas. Elles sont analysées en fonction des différents types de phrases dans lesquels elles peuvent être intégrées. L'analyse syntaxique appliquée en LS (français, langue source) à un item que nous appellerons A, permet de diviser A en diverses sous-entrées (A1, A2, A3, A4, etc.) et d'associer à chacune d'elles une unité lexicale de LC (l'espagnol, langue cible ici) : a1, a2, a3, etc. Ces unités de LC ont obligatoirement une acception équivalente du point de vue sémantique à A1, A2, etc.

Le dictionnaire est conçu dans l'optique de l'encodage en LC, c'est-à-dire qu'il est destiné à un locuteur de LS qui veut traduire de LS à LC, dans notre cas, dans le sens français > espagnol. La recherche de l'information dans le dictionnaire correspond dans ce cas à un processus onomasiologique. Cette situation présuppose que la LS est analysée de manière à permettre la mise en parallèle d'un équivalent unique en LC. Un utilisateur qui recourt au dictionnaire doit avoir toutes les données pour réaliser, de manière complète, l'opération de traduction. Il ne doit donc pas avoir à choisir entre deux ou trois équivalents. S'il y avait

---

8 Une exception : la thèse de doctorat réalisée à l'Université de Valence (Espagne) sous notre direction, par P. Mogorrón sur les phrases *Ser Jestar + prep et Être + prep* qui envisage le contraste français-espagnol dans les expressions qui répondent à ce schéma syntaxique.

plusieurs équivalents traductifs, le dictionnaire n'aurait facilité qu'une partie du processus de traduction.

La présence d'un équivalent unique en face de chacune des sous-entrées est la condition pour que le dictionnaire puisse être utilisé en traduction automatique. Pas plus que l'utilisateur humain peu compétent, la machine ne peut traduire de manière adéquate si elle n'a pas, pour chaque sous-entrée, un équivalent unique, et une mise en correspondance univoque.

La méthode d'analyse est de type distributionnel et transformationnel et constitue une application au domaine contrastif des principes généraux qui ont été appliqués dans les classifications des réalisations incluses sous le titre générique de Lexique-grammaire des langues romanes.

La perspective contrastive du dictionnaire suppose que l'analyse des sous-entrées d'un vocable de LS ne se fasse pas seulement en fonction des schémas syntaxiques des phrases dans lesquelles il peut être inclus. Il existe un autre critère dont il est nécessaire de tenir compte : c'est l'organisation lexico-sémantique de la langue cible. Par exemple, si pour l'entrée *apprendre* nous ne tenons compte que du schéma N hum + apprendre + N [comme, par exemple, le fait le D.F.C. (1966)<sup>9</sup>], la présence de l'unité lexicale nouvelle dans le syntagme de LS *apprendre la nouvelle* impose qu'on présente parallèlement le syntagme attendu en LC : *enterarase de la noticia*. Cela a pour conséquence, dans l'optique traductive qui est la nôtre, qu'il est nécessaire de séparer *apprendre la nouvelle* (N hum + apprendre + N) de *apprendre la géographie* (N hum + apprendre + N) parce que, dans ce dernier cas, l'association attendue est *aprender geografía*. *Noticia* [nouvelle] et *geografía* [géographie] supposent des traductions différentes pour le verbe *apprendre*. Cette séparation entre deux classes de noms en position de complément dans le syntagme N hum + apprendre + N n'aurait pas été nécessaire dans une perspective d'analyse non contrastive.

#### 4 Description des instruments utilisés pour la détermination des sous-unités de LS

Les instruments utilisés pour l'analyse syntaxique qui permettra de délimiter les sous-unités lexicographiques traductives et de justifier la présence de l'équivalent sont de type distributionnel et également transformationnel, ce qui veut dire

9 Les deux unités *apprendre 1* et *apprendre 2* apparaissent non séparées dans le D.F.C. L'une des constructions distinguées est *apprendre quelque chose + apprendre que* (+ ind.), qui regroupe donc en une seule catégorie, les deux types de phrases que nous avons séparées. Les auteurs de cet ouvrage lexicographique n'opposent à cette unité *apprendre* ainsi discriminée, qu'une seconde unité qui a pour schéma *apprendre quelque chose à quelqu'un* ou *apprendre à + V inf. + à quelqu'un*. (glosé dans ce dernier cas par *enseigner*).

que les caractéristiques syntaxiques peuvent être manifestées au moyen d'opérations d'insertion, de déplacement, de passivation, etc.

Exemples d'insertion discriminant entre deux comportements superficiellement identiques :

- (1) N + entrer + Loc N [le café / l'hôpital]
- (1a) Max entre dans le café (Max entra en el bar)
- (1b) Max entre boire dans le café (insertion de un Vinf.)
- (1c) Max entre à l'hôpital (= est hospitalisé) (Max ingresa en el hospital)
- (1d) \*Max entre se soigner à l'hôpital

Exemples de transformation (passivation) discriminant entre deux comportements superficiellement identiques :

- (2) N + tourner + N (la manivelle / le coin de la rue)
- (2a) Max tourne la manivelle (Max gira la manivela)
- (2b) La manivelle est tournée par Max
- (2c) Max [la voiture] tourne le coin de la rue (Max dobla la esquina de la calle)
- (2d) \*Le coin de la rue est tourné par Max [la voiture]

Exemples de transformation (N = que P) discriminant entre deux comportements superficiellement identiques :

- (3) N + apprendre + N (la géographie de la France / la mort de Léo)
- (3a) Max apprend la géographie de la France (Max estudia geografía de Francia)

Il n'y a aucune transformation possible en (3a), mais en (3b) elle est possible.

- (3b) Max apprend la mort de son père (Max se entera de la muerte de su padre)
- (3c) Max apprend que son père est mort

Il est nécessaire de souligner que, dans le cas de verbes opérateurs en phrases libres que nous examinons, il n'existe pas toujours d'opérations syntaxiques qui, comme c'était le cas pour le verbe *apprendre*, permettent de discriminer entre deux fonctionnements superficiellement identiques et, en conséquence, de présenter en regard de chacun d'eux un équivalent interlingual appartenant à LC. C'est le cas par exemple de :

- (4) Max abat l'arbre / l'avion / le bœuf (Max tala el árbol, derriba el avión, sacrifica el buey)

Dans ce cas, il n'existe pas de procédés de discrimination de type syntaxique entre abattre l'arbre / l'avion / le bœuf, et il est donc nécessaire de recourir aux procédés utilisés dans les dictionnaires électroniques, spécialement dans celui du français (Gross, 1992:94), qui se sert de la notion de classe de noms.<sup>10</sup>

<sup>10</sup> Gross parle de *classes d'objets*. Nous préférons parler de classes (syntaxico-sémantiques) de noms. Il s'agit des noms qui ont une même aptitude à la combinaison avec un opérateur donné.

Considérons les V *boire, supporter, saouler, distiller* : il est possible de constater que si un substantif est susceptible d'être complément de l'ensemble de ces opérateurs, il fait obligatoirement (ou presque) partie de la classe «boissons alcoolisées». Les conditions pour former une classe de noms sont donc de nature syntaxico-sémantique, car elles sont gouvernées par l'acceptabilité d'un ensemble donné de phrases.

(5a) Max boit, supporte, distile du whisky / de l'eau de vie / du calvados, etc.  
(Max bebe, soporta, destila güisqui / aguardiente / orujo, etc.)

(Classe de noms : boissons alcooliques)

(5b) \*Max boit, supporte, distille de la viande / de l'eau minérale / du jus de fruit  
(Max bebe, soporta, destila carne / agua mineral / zumo, etc.)

(Aucune classe de noms homogène)

On peut ainsi observer que la constitution de classes de noms n'est pas uniquement conditionnée par la sémantique, c'est-à-dire par l'existence de signifiés communs isolés de façon intuitive. Cette notion de classe de noms, de caractère instrumental, s'applique à des ensembles dont l'existence est conditionnée par l'acceptabilité des phrases dans lesquelles un opérateur donné est présent. Nous constatons, si nous reprenons l'exemple (4) Max abat l'arbre / l'avion / le bœuf, qu'il est possible de distinguer trois sous-unités :

(6a) abattre (C : arbres / árboles) :  
Max tala el arbol / el haya / el tilo / la encina / etc.  
(Max abat l'arbre / le hêtre / le tilleul / le chêne / etc.)

(6b) abattre (C : avions / aviones) :  
Max derriba el avión / el helicóptero / el F-16 / etc.  
(Max abat l'avion, l'hélicoptère, le F 16, etc.)

(6c) abattre (C : animaux domestiques / animales domésticos) :  
Max sacrifica el buey / el cordero / el conejo / el perro, etc.  
(Max abat le bœuf / l'agneau / le lapin / le chien / etc.)

## 5 Zones figées des phrases qui incluent l'item traité

Les analyses d'un dictionnaire bilingue d'encodage<sup>11</sup> doivent prendre en compte le caractère figé de certaines zones des phrases qui incluent le vocable traité dans l'article. Ces phrases sont classées en fonction de la présence, et de la situation, de la partie figée par rapport aux éléments variables. Il est indispensable dans la perspective de la traduction automatique de répertorier et de traduire les phrases qui incluent des syntagmes figés : en effet, leur interprétation sémantique ne correspond pas à la somme des signifiés des éléments qui constituent la séquence

<sup>11</sup> Il est évident que la même exigence existe pour les dictionnaires de décodage.

et, en conséquence, elle détermine une traduction divergente par rapport à cette «somme».

- (7a) Max ne porte pas Léa dans son cœur  
(Max no le tiene ningún cariño a Lea)

ne peut pas être interprétée de la même manière que

- (7b) Max ne porte pas Léa dans ses bras  
(Max no lleva a Lea en brazos)

Le dictionnaire bilingue que nous décrivons offre parmi le matériel qu'il a inventorié la phrase figée «N0hum ne pas porter N1hum dans son cœur»<sup>12</sup> dans laquelle il y a deux éléments libres (N0hum et N1hum) et il permet la dissimilation des phrases (non figées) dont l'interprétation ne se fait pas de manière spécifique.

Un dictionnaire établi sur le modèle tabulaire des inventaires du L.A.D.L. est informatisable, et fait partie des instruments dont l'élaboration vise la réalisation de la traduction automatique. D'un point de vue plus général, les procédés d'analyse que nous avons présentés ont une fonction de discrimination syntaxique et une optique spécifiquement contrastive. En effet, il ne s'agit pas de la description du fonctionnement de x items en vue de leur classification dans le cadre de la LS, comme le fait le Lexique-grammaire en un classement à valeur générale. Pour nous, la syntaxe est un moyen mais non une fin en soi. Notre objectif est de nature lexicographique, donc finalement pratique. Il s'agit de mettre en face d'une unité de LS un équivalent traductif et un seul, adéquat pour l'opération de traduction manuelle mais aussi de traduction sur machine et de le faire sur des bases d'une autre nature que celle des dictionnaires conventionnels. Néanmoins, du point de vue théorique, il est probable que ces analyses contrastives qui relient sémantique et syntaxe et inventorient les fonctionnements syntaxiques d'items intégrés dans des phrases libres et dans les phrases figées fourniront des matériaux qui pourront contribuer à la construction d'une linguistique contrastive français-espagnol systématique, aussi bien dans le domaine de la syntaxe comparée des deux langues que dans celui du lexique.

<sup>12</sup> Dans ce cas, la phrase se formalise de la manière suivante N0 + C (V porter + dans son cœur) N1, en laquelle C est la constante et N0 et N1 les variables.

**Présentation de l'article de dictionnaire monter :  
légende du tableau**

1/ entrée monter (colonne A)

a : schéma de la construction phrase simple

ex : N0 + V (+ Ni)

ou

N0 + V (+N1 (+ prép N1))

N0 représente le sujet du verbe (V);

N1 représente le premier complément;

N2 représente le second complément.

b : schéma de la construction expression figée

ex. : N0 + c (V + la garde)

c'est la constante et l'autre (ou les autres) élément(s) est (sont) variable(s)

2/ Traduction (colonne B)

Equivalent de l'entrée.

L'univocité étant le principe du dictionnaire, il ne peut figurer dans cette colonne qu'une seule unité lexicale.

3/ Exemple de la construction qui correspond au schéma de la colonne A (colonne C)

4/ Nature syntaxico-sémantique du N0 (colonnes D, E, F) :

N0 = Nhum (nom humain)

N0 = N-hum (nom non humain)

N0 = Nnr (nom non restreint = Nhum ou N-hum ou expression :

Ex. : Ce que tu me dis me surprend

Ce que tu me dis = Nnr)

5/ Classe de N0 : nature sémantique du N0 (colonne G)

6/ Nature syntaxico-sémantique du N1 (colonnes H, I, J)

7/ Classe de N1. Nature sémantique du N1 (colonne K)

8/ Entrées différenciées et opérations ou éléments qui servent à cette différenciation (colonnes L et M);

Cette colonne montre les propriétés (syntaxiques et sémantiques) qui différencient les lexèmes entre eux (en phrases libres).

Dans la zone des expressions figées, l'existence d'une constante rend ces différenciations superflues.

*Présentation de l'article «monter» : texte*

## LES DICTIONNAIRES BILINGUES

	A	B	C	D	E	F
1	Phrases libres/ Frases libres					
2				Nohum	No-hum	No nr
3	1-a No. + V (+ N1loc)	Subir	Max monte (fa côte)	+	+	-
4	1- b No + + V inf (+ N1 Loc)	Subir	Max monte ouvrir (la porte)	+	-	-
5	2- a No + V + N1	Subir	Max monte les caisses	+	-	-
6	2-b No + V + N1	Montar	Max monte l'âne	+	-	-
7	2-c No + V + N1	Montar	Max monte la roue	+	-	-
8	2-d No + V + N1	Armar	Max monte le meccano	+	-	-
9	2-e No + V + N1	Engastar	Max monte le diamant	+	-	-
10	2-f No + V + N1	Montar	Max monte une affaire	+	-	-
11	3- No + V + N1 contre N2	Poner a N en contra de N	Max monte Léa contre Téo	+	-	+
12	4-a No + V + prép N1loc	subir(se) en	Max (la moto)monte sur la voie	+	+	-
13	4-b No + V + (en - dét) (dans + dét.) N1	Subirse a/en	Max monte dans le train	+	-	-
14	4-c No + V + dans/à N1	Trepar	Max monte à/dans l'arbre	+	-	-
15	5- No + se V + en N1	Equiparse en	Max se monte en linge	+	-	-
16	6-a No + V (à N1)	Asciende	L'avion monte (à 2000 m)	-	+	-
17	6-b No + V	Subir	L'eau monte (à 2 m.)	-	+	-
18	7-a No + V	Subir	le niveau (prix) monte	-	+	-
19	7-b No + V + à N1.	Ascender à	La note monte à 200 Fr.	-	+	-
20	8- No + V + N1	Montar	L'étalon monte la jument	-	+	-
21						
22						
23	Phrases figées/ frases fijas					
24	No + C ( V + la garde)	estar de guardia	Max monte la garde	+	-	-
25	No + C ( V + un coup)	Preparar un golpe	Max prépare un coup	+	-	-
26	No + C (V + à pied) + Nloc	Ir andando	Max monte à pied la colline	+	-	-
27	No + C (V + à cheval)	montar a caballo	Max monte à cheval	+	-	-
28	No + C (V + à bicyclette)	montar en bicicleta	Max monte en bicyclette	+	-	-
29	No + C (V+ en grade)	Ascender	Max monte en grade	+	-	-
30	No + C (V+ sur le trône)	Llegar al trono	Max est monté sur le trône	+	-	-
31	N0 + C ( V + en chaire)	Subir al púlpito	Max monte en chaire	+	-	-
32	No + C ( V + à l'assaut)	Lanzarse al asallo	Max monte à l'assaut	+	-	-
33	No + C ( V + au front)	Ir al frente	Max monte au front	+	-	-
34	No + C ( se V + la tête)	Ir iritándose	Max se monte la tête	+	-	-
35	No + C (V+ la tête à ) N1	Soliviantar	Max monte la tête à Léa	+	+	+
36	C ( Le rouge + V + au(x) joues (visage)de ) N1	Ponerse colorado	Le rouge monte au visage de Max	-	+	-
37	No + C ( V + à la tête de ) N1	Subirse a la cabeza	Le vin monte à la tête de Max	-	+	-
38	No + C (V+ en graine)	Granar	La laitue monte en graine	-	+	-
39	La génération montante	los jóvenes		+	-	-

	G	H	I	J	K	L	M	N
1							Entrées différenciées	Opérations/ Eléments
2	Classe de No	N1hum	N1-hum	No nr	Classe de N1	Domaine		de différenciation
3		-	+	-	Loc		1a /1b, 2, 3, 4, ...	N1 = Loc: facultatif
4		+	+	-			1b/ 1a, 2, 3, ...	No hum
5		+	+	-			2/1, 3, 4, ...	No hum, N1 -hum obligatoire
6		-	+	-	Montures			No hum, N1 -hum obligatoire
7		-	+	-	Obj. montables			No hum, N1 -hum obligatoire
8		-	+	-	Pièces à assembler			No hum, N1 -hum obligatoire
9		-	+	-	Pierre précieuse ou perle	Bijouterie		No hum, N1 -hum obligatoire
10		-	+	-	Activités lucratives			No hum, N1 -hum obligatoire
11		+	-	+			3/1, 2, 4, 5, ...	N1 hum contre N2 Hum
12	Nohum ou N-hum =objets mobiles	-	+	-	Loc +horizontalité		4/ 1, 2, 3, 5, 6...	Prép sur/dans N1 Loc horizontal
13		-	+	-	Véhicules			Prép sur/dans (dans + dét.) N1 N1 véhicule
14	Nohum ou N-hum =objets mobiles	-	+	-	Loc + verticalité			Prép sur/dans N1 Loc vertical
15		-	+	-	Objets		5/ 1, 2, 3, 4, 6, 7,..	Prép en (det. 0) N1
16	Obj. mobiles déplacement vertical	-	+	-	Distance		6/ 1, 2, 3, 4, 7,..	No objet + N1 déplacement vertical
17	Fluides déplacement vertical	-	+	-	Distance			No fluide + N1 déplacement vertical
18	abst						7/1, 2, 3, 4, 6, 8	N1 quantité
19		-	+	-	Quantité d'argent			
20	Animal mâle	-	+	-	Animal femelle		8/ 1, 2, 3, 4, 5, ...,	No , N2 animaux
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35		+	-					
36		+	-					
37		+	-					
38		-	+					
39								



# LA GRAMMAIRE DANS LES DICTIONNAIRES BILINGUES

Marcel LEMMENS

### 1 Introduction

Dans les dictionnaires monolingues pour apprenants, comme le *Collins Cobuild Dictionary of the English Language* et le *Longman Dictionary of Contemporary English*, il est courant d'introduire des informations relativement détaillées sur le comportement grammatical des mots vedettes. La plupart des dictionnaires bilingues, qu'ils soient pour débutants ou d'un niveau plus avancé, contiennent également des informations grammaticales, mais, en règle générale, elles sont moins détaillées que celles contenues dans les dictionnaires monolingues.

Les raisons sont évidentes. Tout d'abord, les locuteurs non anglophones qui utilisent les dictionnaires monolingues ont souvent besoin d'une information grammaticale spécifique pour savoir comment employer les mots qu'ils recherchent. Ensuite, les dictionnaires monolingues s'adressent à un public bien plus vaste que les dictionnaires bilingues. Ils sont commercialisés dans le monde entier et sont utilisés par de nombreux locuteurs non natifs dans toutes sortes de pays. Les dictionnaires bilingues, par contre, s'adressent à des usagers, locuteurs natifs d'une langue et souhaitant en utiliser une autre, qui sont donc censés posséder les connaissances de base nécessaires, y compris une certaine intuition à propos du comportement grammatical des mots de leur langue maternelle.

Bien qu'il y ait de nombreuses différences entre les dictionnaires monolingues pour apprenants et les dictionnaires bilingues en ce qui concerne la quantité d'informations grammaticales données, la plupart des dictionnaires bilingues donnent tout de même quelques renseignements sur l'utilisation des mots vedettes, en donnant par exemple la catégorie grammaticale à laquelle ils appartiennent ainsi que leurs formes fléchies. Afin de savoir si les utilisateurs néerlandais des dictionnaires bilingues ressentent le besoin d'y trouver de l'information grammaticale sur les mots qu'ils recherchent, et si oui, en quelle quantité, et afin de déterminer s'ils savent réellement quel genre d'information grammaticale ils peuvent trouver dans leurs dictionnaires favoris, une enquête a

été menée auprès de 111 étudiants de l'Ecole de Traduction et d'Interprétation de Maastricht, et de 18 professionnels de la traduction.

Puisque l'on compare les réponses données par les personnes interrogées et les informations grammaticales effectivement données dans un certain nombre de dictionnaires bilingues, cette enquête offre également un aperçu général des besoins grammaticaux supposés, c'est-à-dire du type et de la quantité d'informations grammaticales que les auteurs et/ou les éditeurs de dictionnaires estiment nécessaires.

## **2 L'enquête**

Les dictionnaires bilingues jouent un rôle important tout au long des différents cursus de traduction offerts par l'Ecole de Traduction et d'Interprétation. Certains départements fournissent même des cours d'initiation à l'utilisation des dictionnaires. C'est ainsi que, au moment de l'enquête, à la fin du premier semestre universitaire (décembre 1993/janvier 1994), les étudiants ayant choisi l'anglais parmi leurs deux langues étrangères (soit 98 des 111 étudiants interrogés) possédaient des connaissances de base sur le Van Dale.

Notre enquête avait pour objectif principal de déterminer ce que les utilisateurs de dictionnaires, lorsqu'ils sont à un niveau moyen supérieur ou avancé, connaissent de l'information grammaticale contenue dans ceux des dictionnaires bilingues qu'ils consultent le plus, ou qu'ils préfèrent, et de quelle genre d'information grammaticale ils ont le sentiment d'avoir besoin.

Le questionnaire (qui peut être obtenu en écrivant à l'auteur à l'adresse suivante : Kalverhaf 57, NL-6241, CZ Bunde, Pays-Bas) remis aux sujets comportait des questions ouvertes et un certain nombre de questions appelant une réponse par oui ou par non. Les sujets ne devaient pas consulter leur dictionnaire pour répondre aux questions, mais seulement indiquer ce qu'ils pensaient que leur dictionnaire contenait. De plus, ils avaient pour consigne de répondre aux questions 1 à 10 dans l'ordre croissant sans revenir sur les questions auxquelles ils avaient déjà répondu, autrement dit sans corriger ou ajouter des informations après avoir vu les autres questions.

## **3 Les résultats pour l'ensemble des sujets**

Le nombre de personnes ayant répondu au questionnaire est de 129 au total : 111 étudiants et 18 professionnels.

Tableau n° 1: Sujets

Etudiants		Professionnels	
1 <sup>re</sup> année	46	traducteurs	14
2 <sup>e</sup> année	30	secrétaires	1
3 <sup>e</sup> année	12	professeurs de traduction	3
4 <sup>e</sup> année	23		

Total	111	Total	18
-------	-----	-------	----

Tableau n° 2 : Première langue étrangère des sujets

Etudiants		Professionnels	
anglais	49	anglais	13
français	14	français	2
russe	12	allemand	1
chinois	11	arabe	1
italien	7	néerlandais	1
arabe	6		
portugais	6		
espagnol	5		
néerlandais	1		

Total	111	Total	18
-------	-----	-------	----

En réponse à la question 4 : («Quel est votre dictionnaire préféré pour votre langue préférée ?»), les sujets ont mentionné 21 dictionnaires différents. La question 5 comportait deux parties. Tout d'abord, les sujets devaient indiquer les différents types d'information qu'ils considéraient comme des «informations grammaticales importantes dans un dictionnaire» (ils devaient répondre à cette question sans consulter les dictionnaires concernés). La seconde partie de la question exigeait une réponse par oui ou par non : «Avez-vous regardé la page 4 du questionnaire ?». Cette page contenait cinq exemples d'entrées provenant de

différents dictionnaires bilingues. On avait souligné et expliqué dans une colonne à part les informations grammaticales contenues dans ces entrées.

Tableau n° 3: Que considérez-vous comme relevant de la grammaire dans les dictionnaires?

Type d'information	Nombre de sujets (sur 112)	%
genre des noms	62	55,4 %
conjugaison des verbes	55	49,1 %
indicateur de la catégorie grammaticale	47	42,0 %
formes des noms au pluriel	38	33,9 %
flexions des autres catégories grammaticales	32	28,6 %
types de verbes	27	24,1 %
combinaison avec d'autres mots	12	10,7 %
position en fonction d'autres mots	7	6,3 %
formation de dérivés	2	1,8 %
restrictions concernant l'usage	2	1,8 %
divers	60	53,6 %

Pour permettre la comparaison entre les réponses données aux différentes questions de l'enquête, on a utilisé pour les réponses aux questions ouvertes 5a, 6b, 7b et 10b les mêmes catégories que celles utilisées à la question 9. Certaines de ces catégories pouvaient se comprendre facilement, mais d'autres furent explicitées pour éviter tout risque de confusion. On a ainsi donné les explications et les exemples suivants : la catégorie grammaticale (par exemple si le mot vedette est un nom, un verbe), le type de verbe (transitif, intransitif, pronominal, réfléchi), les restrictions concernant l'usage («s'emploie seulement au passif», ou «s'emploie souvent au pluriel»), la position en fonction d'autres mots («adjectif placé devant le nom»), la formation de dérivés (adverbes dérivés d'adjectifs, noms dérivés de verbes) et la combinaison avec d'autres mots (par exemple quelle préposition s'utilise le plus fréquemment avec le mot vedette). Bien que certaines catégories puissent se recouper (c'est le cas par exemple pour «restrictions concernant l'usage» et «position en fonction d'autres mots»), cette classification permet d'opérer des distinctions assez subtiles entre les différentes réponses. Une catégorie, «utilisation du pluriel ou du singulier», très proche des «restrictions

concernant l'usage», n'est pas mentionnée dans le tableau n° 3 ci-dessus, mais était incluse dans la liste de la question 9.

Tableau n° 4: Avez vous consulté la page 4 pour savoir ce qu'on entend par grammaire dans les dictionnaires ?

	Nombre de sujets	%
oui	60	46,5 %
non	49	38,0 %
réponses non valables	20	15,5 %

Les questions 5a et 5b avaient pour objet de découvrir si les étudiants et les professionnels tels que des traducteurs avaient une idée précise de ce qu'est la grammaire dans les dictionnaires. En fait, un grand nombre de sujets (46,5 %) ne savaient pas, ou pas bien, ce que «l'information grammaticale» dans un dictionnaire représente, et ils ont dû vérifier les exemples donnés à la page 4 du questionnaire. Sur tous les sujets, 17 n'ont pas répondu à la question (principalement des étudiants de première année), alors que beaucoup d'autres n'ont donné qu'une ou deux caractéristiques, mentionnant même parfois des informations de type non grammatical. Parmi ceux qui ont donné une réponse, plusieurs sujets ont fourni des indications très vagues, telles que «les choses que je ne sais pas».

Bien que de nombreux sujets aient en effet fait référence à des contenus grammaticaux typiques des dictionnaires bilingues (catégorie grammaticale, inflexion, genre), d'autres sujets ont également mentionné des caractéristiques que l'on ne considère pas d'habitude comme relevant de l'information grammaticale. Les nombreux exemples donnés dans la catégorie «divers» comprenaient des allusions aux proverbes, à la prononciation, aux symboles d'accentuation, à l'orthographe, à la fréquence, au style, aux exemples, à la ponctuation et aux synonymes. On peut considérer comme purement grammatical un petit nombre seulement des éléments de cette catégorie, comme le fait qu'un mot soit comptable ou non (cité 3 fois), la référence à un précis grammatical (cité 2 fois), les temps, la signification des radicaux en arabe, les cas, et la référence aux tableaux de déclinaisons (tous cités 1 fois).

#### 4 Les résultats pour les quatre dictionnaires les plus utilisés

Tableau n° 5: Les quatre dictionnaires le plus souvent cités par les étudiants et les professionnels

Dictionnaires	Paires linguistiques	Nombre de sujets	%
1 Van Dale Groot Woordenboek Nederlands-Engels & Engels-Nederlands	néerlandais ↔ anglais	69 <sup>1</sup>	53,5 %
2 Van Dale Groot Woordenboek Nederlands-Frans & Frans-Nederlands	néerlandais ↔ français	13	10,1 %
3 Prisma Woordenboek Nederlands-Italiaans & Italiaans-Nederlands	néerlandais ↔ italien	8	6,2 %
4 A Dictionary of Modern Written Arabic (Wehr)	arabe > anglais	7	5,4 %

Le nombre de sujets ayant indiqué que l'anglais était leur langue préférée, ou celle qu'ils utilisaient le plus fréquemment, est beaucoup plus élevé que pour les autres langues. Ceci s'explique par le fait que la plupart des étudiants de l'École de Traduction et d'Interprétation choisissent l'anglais comme langue étrangère, et que, la plupart des 18 professionnels ayant répondu au questionnaire avaient également participé à un séminaire de travail sur la traduction dans le sens néerlandais > anglais qui a eu lieu en janvier 1994.

Tout comme le tableau n° 3, le tableau n° 6 donne les réponses à la question 5, mais uniquement pour les quatre dictionnaires cités ci-dessus. Les chiffres entre parenthèses en haut des colonnes indiquent combien de réponses valables ont été données pour chaque dictionnaire.

<sup>1</sup> Il faut noter qu'il y a un écart entre le nombre de sujets qui ont répondu que l'anglais était leur langue favorite (49 + 13 = 62) et le nombre de personnes qui ont rempli le questionnaire pour les dictionnaires Van Dale hollandais > anglais (69). On ne sait pas pourquoi ces 7 sujets ont répondu ainsi.

Tableau n° 6: Que considérez-vous comme relevant de la grammaire dans les dictionnaires ?

Dictionnaires	1:A ↔ N (56) <sup>2</sup>	2:F ↔ N (13)	3:I ↔ N (8)	4:Ar > A (7)
---------------	------------------------------	-----------------	----------------	-----------------

types d'information	nombre	%	nombre	%	nombre	%	nombre	%
indicateur de la catégorie grammaticale	23	41,1	7	53,8	4	50	4	57,1
formes des noms au pluriel	18	32,1	3	23,1	4	50	5	71,4
genre des noms	25	44,6	9	69,2	6	75	4	57,1
types de verbes	17	30,4	4	30,8	1	12,5	0	0
conjugaison des verbes	23	41,1	6	46,2	6	75	6	85,7
restrictions concernant l'usage	2	3,6	0	0	0	0	0	0
position en fonction d'autres mots	5	8,9	0	0	1	12,5	0	0
flexions des autres catégories grammaticales	12	21,4	5	38,5	4	50	3	42,9
formation de dérivés	2	3,6	0	0	0	0	0	0
combinaison avec d'autres mots	5	8,9	1	7,7	2	25	1	14,3
divers	32	57,1	4	30,8	4	50	2	28,6

2 Parmi les sujets ayant choisi le dictionnaire anglais ↔ néerlandais (dictionnaire n° 1), 13 n'ont pas répondu à la question 5. Ce sont, en majeure partie, des étudiants de première année. Tous les sujets ayant choisi les trois autres dictionnaires ont donné au moins une caractéristique grammaticale.

La principale différence entre les réponses pour les quatre dictionnaires réside dans le fait que le consensus parmi les sujets ayant choisi les dictionnaires 2, 3 et 4 est plus important que parmi ceux ayant choisi le dictionnaire 1. Les étudiants de français, italien et arabe semblent désigner d'un commun accord les types d'informations suivants comme étant caractéristiques de la composante grammaticale dans les dictionnaires : l'indication de la catégorie grammaticale, la forme des noms au pluriel (en particulier les étudiants d'arabe), le genre des noms (en particulier les étudiants de français et d'italien), la conjugaison des verbes (surtout les étudiants d'italien et d'arabe) et les inflexions des autres catégories grammaticales. En ce qui concerne le dictionnaire 1, il y a moins de réponses pour la catégorie «flexion des autres catégories grammaticales» que pour les autres dictionnaires et beaucoup plus de réponses pour la catégorie «divers» (comme pour le dictionnaire 3). Relativement peu d'utilisateurs du Van Dale néerlandais ↔ anglais semblent savoir qu'il existe des flexions pour les adjectifs par exemple. Etant donné le nombre important d'éléments non grammaticaux contenus dans la catégorie «divers», beaucoup de sujets ayant choisi l'anglais ont donc une vision plus large de ce que représente «l'information grammaticale» que les autres.

Tableau n° 7 : Avez vous consulté la page 4 pour savoir ce qu'on entend par grammaire dans les dictionnaires ?

Dictionnaires	1:A ↔ N		2:F ↔ N		3:I ↔ N		4:Ar > A	
	nombre	%	nombre	%	nombre	%	nombre	%
oui	31	44,9	6	46,2	2	25,0	6	85,7
non	25	36,9	4	30,8	6	75,0	1	14,3
aucune réponse donnée	13	18,8	3	23,1	0	0	0	0

Les pourcentages trouvés pour les dictionnaires 1 et 2 sont proches des résultats généraux (voir tableau n° 4). Les résultats pour les dictionnaires 3 et 4, cependant, montrent des différences majeures. La majorité des étudiants d'italien, (tous étaient étudiants de seconde année) semblent être certains de leurs réponses : seulement 2 d'entre eux ont estimé nécessaire de lire les exemples donnés à la page 4 du questionnaire. Par contre, les 6 étudiants d'arabe étaient bien moins sûrs de leurs réponses et ont consulté la page 4. Seul le professeur d'arabe n'a pas eu besoin de consulter cette page.

Alors que la question n° 5 ne concernait aucun dictionnaire en particulier, les questions 6 à 10 ont pour objet d'en savoir plus sur la «perception» de la composante grammaticale dans des dictionnaires spécifiques. La question 6

demandait aux participants d'indiquer quel genre d'information grammaticale ils pensent trouver dans leur dictionnaire préféré.

## 5 Les résultats pour le Van Dale anglais ↔ néerlandais

Les deux volumes du *Van Dale Groot Woordenboek* représentent le dictionnaire bilingue anglais ↔ néerlandais le plus moderne et le plus exhaustif (approximativement 1500 pages par volume) actuellement sur le marché. Ces deux volumes sont particulièrement populaires à la fois parmi les étudiants en langue et les utilisateurs professionnels de l'anglais. La principale fonction du volume anglais > néerlandais est d'aider l'utilisateur à comprendre l'anglais et celle du volume néerlandais > anglais est d'aider l'utilisateur à produire de l'anglais. La présentation est différente de celle des autres dictionnaires bilingues néerlandais puisque les principaux équivalents sont groupés au début de chaque entrée (voir Marelló, dans ce volume). Les exemples sont également groupés et sont systématiquement associés aux sens appropriés grâce à un système de renvois sous forme de chiffres. Le volume anglais > néerlandais contient un précis grammatical de 75 pages classé par ordre alphabétique.

Sur tous les sujets, 69 ont répondu que le Van Dale anglais ↔ néerlandais étaient leur dictionnaire préféré et/ou celui qu'ils utilisaient le plus souvent. Seulement 52 de ces 69 sujets (75,4 %) ont répondu à la question 6a<sup>3</sup>. Les 52 sujets ont indiqué qu'ils savaient que leur dictionnaire contenait des informations grammaticales. Ils devaient ensuite préciser (question 6b) quel type d'information ils pensaient que leur dictionnaire contenait, mais seulement 39 sujets ont répondu à cette question. Tous ceux qui ont répondu à la question 6a ont également répondu à la question 6c («Sous quelle forme la grammaire se présente-t-elle dans votre dictionnaire ?») et à la question 9 («Veuillez cocher les 3 types d'information grammaticale que vous utilisez le plus souvent»). Les sujets ayant choisi de répondre pour les trois autres dictionnaires principaux se sont montrés plus cohérents ou plus consciencieux et ont répondu à toutes les questions auxquelles ils devaient répondre. Le nombre important d'étudiants de première année (31 sur les 46 participant au questionnaire) ayant choisi le Van Dale anglais ↔ néerlandais peut en être la cause. Il est peut-être plus difficile pour les étudiants de première année de décrire ce qu'il y a dans leur dictionnaire que pour les autres étudiants et pour les professionnels, même s'ils ont déjà suivi un cours d'introduction sur les dictionnaires monolingues et bilingues.

3 Les résultats de ce questionnaire montrent bien combien il est difficile d'extraire des informations d'un groupe de personnes au moyen d'une enquête, ou de tirer des conclusions non équivoques de ces résultats (voir le tableau n° 21 ci-dessous).

Le tableau n° 8 compare trois choses. La seconde colonne donne les résultats pour la question 6b : elle indique combien de sujets ont mentionné les différents types d'information grammaticale énumérées dans la première colonne du tableau (les chiffres entre parenthèses représentent le nombre de réponses valables). La troisième colonne montre combien de sujets ont coché les caractéristiques grammaticales telles qu'elles étaient énumérées à la question 9. Enfin, les symboles + et - de la quatrième colonne indiquent si les caractéristiques mentionnées se trouvent dans les deux volumes du dictionnaire ou non.

Tableau n° 8

Type d'information grammaticale	Question 6b (39)	Question 9 (52)	Présentation systématique ?
---------------------------------	------------------	-----------------	-----------------------------

	nombre	%	nombre	%	N > A	A > N
indicateur de la catégorie grammaticale	17	43,6	14	26,9	+	+
formes des noms au pluriel	17	43,6	8	15,4	-	+
genre des noms	19	48,7	19	36,5	+	-
types de verbes	13	33,3	14	26,9	+	+
conjugaison des verbes	14	35,3	16	30,8	-	+
restrictions concernant l'usage	1	2,6	15	28,8	-	+
position en fonction des autres mots	1	2,6	14	26,9	-	+
flexion des autres catégories grammaticales	5	12,8	0	0	-	+
formation de dérivés	1	2,6	18	34,6	+	+
combinaison avec d'autres mots	3	7,7	31	59,6	+	+
divers	23	59,0	0	0		

Une comparaison entre les réponses aux questions 6a et 9 montre une fréquence relative plus importante pour la question 9 en ce qui concerne les catégories suivantes : «restrictions concernant l'usage», «position en fonction des autres mots», «formation de dérivés», et «combinaison avec d'autres mots». Au contraire, les pourcentages apparaissant dans les réponses à la question 9 sont moins élevés pour les catégories suivantes : «indications de la catégorie grammaticale», «forme des noms au pluriel», «utilisation du pluriel ou du singulier», «inflexion des autres parties du discours» et la catégorie «divers». Il faut noter que pour la question 9, les sujets devaient faire leur choix parmi une liste de caractéristiques et en cocher 3 seulement (ce que tous ont fait) : la valeur relative des résultats est donc plus importante que les nombres absolus. La question 6b était ouverte; les sujets donnaient eux-mêmes la liste des caractéristiques. Dans ce cas, ce sont donc les nombres absolus les plus importants; en d'autres termes, il est intéressant de voir quelles catégories les sujets ont mentionnées d'eux-mêmes. D'ailleurs, les nombres absolus apparaissant dans les résultats de la question 6b sont faibles : la plupart des sujets n'ont pu citer plus d'une ou deux caractéristiques.

Lorsqu'ils se sont trouvés confrontés à une liste de types possibles d'information grammaticale, beaucoup de sujets semblent s'être rendu compte que leur dictionnaire contenait des informations qu'ils n'étaient pas capables de décrire. Ceci concerne en particulier les catégories «restrictions concernant l'usage», «position en fonction d'autres mots», «formation de dérivés» et «combinaison avec d'autres mots». Bien que certains ne considèrent pas la catégorie des prépositions et des adverbes à utiliser dans les collocations comme une catégorie «grammaticale», elle a tout de même été la catégorie la plus souvent citée. Les utilisateurs semblent considérer qu'il s'agit d'une information essentielle dans un dictionnaire bilingue.

Une des conclusions que l'on peut tirer des réponses aux questions 6b et 9 est que les types d'information grammaticale que l'on trouve le plus fréquemment dans les dictionnaires bilingues, (soit l'indication de la catégorie grammaticale, la forme pluriel, le genre des noms, et le type de verbes et/ou la conjugaison des verbes) sont aussi ceux cités le plus souvent par les sujets.

Mais jusqu'à quel point les utilisateurs connaissent-ils leurs dictionnaires ? Pour le savoir, on peut comparer les informations données dans la quatrième colonne avec les réponses à la question 6d («Y a-t-il des différences entre les deux volumes du dictionnaire ?»). Sur les 42 sujets qui ont répondu à la question, 11 ont affirmé qu'il y avait en effet des différences. Un seul sujet a cité une réelle différence («un volume donne le genre, l'autre ne le fait pas»). Un autre sujet a ajouté un commentaire intéressant : «peu ou pas du tout d'information grammaticale sur le néerlandais dans le volume néerlandais > anglais et sur l'anglais dans le volume anglais > néerlandais». Cette réponse n'est pas complètement exacte, mais elle révèle un des aspects typiques de beaucoup de dictionnaires bilingues comme le Van Dale néerlandais > anglais.

Comme la quatrième colonne du tableau n° 8 le montre, il y a quelques différences importantes entre les deux volumes : le volume néerlandais > anglais contient bien moins de références systématiques à des aspects grammaticaux que le volume anglais > néerlandais. Ceci s'explique par un problème majeur dans la présentation de la grammaire dans le Van Dale néerlandais > anglais (et dans beaucoup d'autres dictionnaires bilingues) : les marqueurs grammaticaux dans le volume néerlandais > anglais renseignent moins sur le comportement des traductions anglaises que sur les mots vedettes néerlandais. Pour le verbe néerlandais *irriteren*, par exemple, les abréviations <*onov.ov.www.*> (signifiant à la fois intransitif et transitif), renseignent seulement sur le comportement grammatical du verbe néerlandais. On ne trouve pas ces informations sur les traductions données dans l'entrée. Il n'y a, par exemple, aucune abréviation indiquant que les traductions anglaises *irritate* et *annoy* ne peuvent être utilisées que comme des verbes transitifs. Autre exemple :

**materie** <de (v.)> **0.1** [stof] **matter** **0.2** [zaak, kwestie] (**subject**) **matter** **2.1**  
*ruwe - raw material, vormloze - mass.*

Le marqueur grammatical placé après le mot vedette <de (v.)> indique que le nom est féminin et qu'il s'emploie avec l'article défini *de* (au lieu de *het*). L'entrée ne précise pas que le nom anglais *matter* est non comptable dans son premier sens, alors qu'il l'est dans le deuxième. Les utilisateurs du dictionnaire qui souhaiteraient obtenir ce genre de renseignements doivent consulter le volume anglais > néerlandais ou un dictionnaire anglais unilingue. Les phrases données en exemple dans le volume néerlandais > anglais illustrent souvent l'emploi des équivalents, mais elles ne couvrent pas toutes les possibilités. D'ailleurs, l'information donnée est bien plus souvent implicite qu'explicite, si bien que les utilisateurs ne peuvent avoir aucune certitude. Puisqu'un seul participant semble être plus ou moins conscient de cette différence importante entre les deux volumes, on peut en déduire que lorsque les utilisateurs traduisent du néerlandais vers l'anglais et consultent le Van Dale néerlandais > anglais, il existe quatre cas de figure :

- a) ils vérifient le comportement grammatical des équivalents proposés dans un autre dictionnaire,
- b) ils s'inspirent de leurs intuitions grammaticales (qui ne sont pas celles d'un anglophone),
- c) ils extrapolent à partir des exemples donnés s'il y en a,
- d) ils ne se rendent tout simplement pas compte que les mots en anglais peuvent ne pas se comporter comme leurs équivalents en néerlandais.

Beaucoup d'étudiants plus âgés et la plupart des traducteurs professionnels (qui de toute façon utilisent généralement des dictionnaires unilingues plutôt que bilingues) vérifient dans d'autres sources lorsqu'ils utilisent des dictionnaires bilingues. Les utilisateurs moins expérimentés, cependant, ne le font pas et se contentent généralement de l'intuition qu'ils ont de leur propre langue en l'appliquant à la langue étrangère.

Le fait que le dictionnaire de traduction Van Dale néerlandais ↔ anglais ne donne pas d'information grammaticale explicite pour les équivalents proposés nous amène à envisager les résultats du tableau n° 8 dans une perspective différente. Les marqueurs de la catégorie grammaticale sont très utiles à la fois dans le sens néerlandais > anglais et anglais > néerlandais, parce qu'ils contribuent à réduire le temps de recherche dans le dictionnaire et permettent de repérer la bonne entrée ou la bonne acception immédiatement. Les utilisateurs peuvent avoir des doutes sur la forme plurielle des noms et pour cela être prêts à consulter le volume anglais > néerlandais. Le score élevé pour la catégorie «genre des noms» est cependant tout à fait étrange. On ne trouve les marqueurs des genres généralement que dans le sens néerlandais > anglais et ces derniers ne s'appliquent qu'aux mots vedettes de la langue maternelle des utilisateurs.

En général, il ne faut aucune connaissance relative au genre pour traduire du néerlandais vers l'anglais. Et si les locuteurs ayant le néerlandais comme langue maternelle traduisent de l'anglais vers le néerlandais, ils peuvent dans la plupart des cas se servir de leur intuition à propos de leur langue maternelle. En règle générale, ils ne consultent pas le volume néerlandais > anglais du Van Dale pour vérifier le genre des mots. L'utilisation des noms au pluriel ou au singulier est une catégorie choisie relativement fréquemment à la question 9 (9 sujets l'ont choisie comme l'un des trois types d'information grammaticale qu'ils utilisaient le plus souvent). Cela peut refléter de façon disproportionnée l'information grammaticale donnée dans les dictionnaires, puisque dans le sens néerlandais > anglais cette information, si elle est donnée, ne l'est qu'implicitement au travers d'exemples. Tout comme les indicateurs de la catégorie grammaticale, les marqueurs verbaux («types de verbes») accélèrent et simplifient le processus de recherche, puisqu'on les utilise généralement pour distinguer les différents sens du mot vedette.

Cependant, aucun équivalent contenu dans les entrées des deux volumes ne comporte de marqueur grammatical. La même remarque s'applique aux catégories «restrictions concernant l'usage» et «position en fonction des autres mots», aucune des deux, n'étant, en fait, fréquemment donnée dans les deux volumes du dictionnaire Van Dale. Le score obtenu pour la catégorie «formation de dérivés» à la question 9 est élevé, ce qui est surprenant. Alors qu'un seul sujet se rappelait qu'il, ou elle, utilisait les deux volumes du Van Dale pour vérifier si, par exemple, tel adjectif possédait un adverbe dérivé en *-ly* et un nom dérivé en *-ness*, presque 35 % des sujets ont indiqué que la formation de dérivés est un type important d'information grammaticale. On a déjà évoqué plus haut la catégorie «combinaison avec d'autres mots».

Aucun des sujets n'a mentionné de caractéristique non grammaticale dans la catégorie «divers» de la question 9. Il n'est donc pas nécessaire de revenir sur cette catégorie.

On a demandé aux sujets qui savaient que les deux volumes du Van Dale contenaient des informations grammaticales (53 d'entre eux), d'indiquer com-

ment l'information grammaticale est présentée dans les dictionnaires. Ils pouvaient cocher le nombre de réponses qu'ils voulaient dans la liste reproduite dans le tableau n° 9 :

Tableau n° 9: Comment l'information grammaticale se présente t-elle ?

Forme de l'information grammaticale	Nombre de sujets (53)	
	nombre	%
flexions complètes (par exemple forme des pluriels)	25	47,2
abréviations (par exemple «adv.» pour «adverbe»)	45	84,9
nombres (par exemple «122» pour «tableau 122»)	10	18,9
codes (par exemple «mv14» pour «pluriel, paragraphe 14»)	10	18,9
exemples	29	54,7

Le Van Dale contient tous les types d'information contenus dans la liste, ce qui signifie que toutes les réponses auraient dû totaliser 53 (autrement dit 100 %), mais ce sont les abréviations et les exemples que l'on retrouve le plus fréquemment dans les deux volumes. Bien que le système des exemples soit utilisé très souvent, il n'est ni cohérent, ni complet. Dans le volume anglais > néerlandais, les flexions ne sont données en entier que pour les verbes irréguliers. Dans le volume néerlandais > anglais, elles apparaissent dans une annexe. Dans le volume anglais > néerlandais, tous les autres types de flexions sont codés et expliqués dans un précis grammatical. Cela signifie donc que seul le volume anglais > néerlandais contient des codes tels que : «mv14».

Puisque si peu de sujets savaient qu'il y avait des différences entre les deux volumes du Van Dale (voir plus haut), on peut considérer que le nombre des sujets qui ont coché toutes les cases est peu important. Un tout petit nombre d'entre eux a coché 4 ou 5 cases. Cela signifie que peu d'utilisateurs du Van Dale savent que l'information grammaticale se présente sous différentes formes dans le dictionnaire. Par contre, le score élevé pour les abréviations et le score peu important pour les chiffres et les codes correspond aux fréquences d'apparition relatives de ces informations. Le score relativement élevé (47,2 %) obtenu pour les flexions complètes est un peu surprenant. Tout d'abord, les flexions ne sont données en entier que dans un nombre restreint d'entrées dans l'un des deux volumes. Ensuite, les étudiants ne semblent pas vérifier la forme des verbes irréguliers dans leur dictionnaire. Surtout lors de la première année, les étudiants continuent de faire des fautes lorsqu'ils utilisent le passé et le participe passé des verbes irréguliers comme *teach*. Le score totalisé pour les exemples (54,7 %) est lui aussi

très intéressant. Puisque le Van Dale ne donne aucune information grammaticale quelle qu'elle soit pour les équivalences proposées, les utilisateurs vont donc devoir se fier aux phrases-exemples données, s'il y en a. Il semble, en fait, qu'environ la moitié seulement des sujets se rendent compte que les exemples peuvent fournir des informations grammaticales.

Sur les 53 sujets qui ont répondu à la question 7 : «Le dictionnaire que vous avez mentionné à la question 4 contient-il suffisamment d'informations grammaticales ?», 42 ont coché OUI et 11 ont coché NON. On a demandé à ceux qui avaient coché NON de préciser ce qui leur semblait manquer. Les 10 sujets qui ont répondu à cette question ont donné les réponses suivantes :

Tableau n° 10

étudiants de 1 <sup>re</sup> année	<ul style="list-style-type: none"> <li>- plus d'expressions figées et de combinaisons de mots</li> <li>- beaucoup de termes ne sont pas expliqués correctement</li> <li>- plus d'exemples</li> </ul>
étudiants de 2 <sup>e</sup> année	<ul style="list-style-type: none"> <li>- information à propos du néerlandais</li> </ul>
étudiants de 3 <sup>e</sup> année	<ul style="list-style-type: none"> <li>- il n'y en a jamais assez</li> </ul>
étudiants de 4 <sup>e</sup> année	<ul style="list-style-type: none"> <li>- information sur la position des adverbes insuffisante (par exemple position de <i>however</i>)</li> <li>- formes du pluriel</li> <li>- trop restreint, trop compact</li> </ul>
professionnels	<ul style="list-style-type: none"> <li>- collocations, exemples avec des prépositions</li> <li>- connotations, changements, évolution du sens des mots</li> </ul>

Seul un nombre réduit de sujets appartenant au petit groupe de ceux qui pensaient que les volumes néerlandais > anglais et anglais > néerlandais ne contenaient pas assez d'information grammaticale se sont montrés précis quant aux améliorations à apporter («position des locutions adverbiales», «formes plurielles», «collocations»). Les autres sont restés très vagues ou ont mentionné des éléments à caractère non grammatical.

La plupart des sujets sont satisfaits du contenu grammatical de leur volume néerlandais > anglais et ne ressentent pas le besoin d'y trouver plus d'information grammaticale. Ceux qui pensent que les volumes devraient contenir plus d'information grammaticale ont eu des difficultés à exprimer exactement quel genre d'amélioration ils souhaitaient.

Tableau n° 11: Selon quelle fréquence utilisez-vous l'information grammaticale contenue dans votre dictionnaire ?

Fréquence	Nombre de sujets (52)	
	nombre	%
une fois par jour	14	26,9
une fois par semaine	29	55,8
une fois par mois	3	5,8
moins d'une fois par mois	4	7,7
jamais	2	3,8

Cette question ne demandait aucun type précis d'information, mais les résultats sont néanmoins très encourageants. Plus de 80 % des sujets utilisent leur dictionnaire pour résoudre des problèmes grammaticaux au moins une fois par semaine.

Les questions 10a à 10c ne concernaient que ceux qui pensaient que le Van Dale néerlandais ↔ anglais ne contenait pas d'information grammaticale (question 6a). Aucun des sujets n'ayant répondu NON à la question 6a, personne n'aurait dû répondre aux questions 10a à 10c. A la question 10a : «Le dictionnaire que vous avez mentionné à la question 4 doit-il contenir de l'information grammaticale ?», 39 sujets ont tout de même répondu OUI et 5 NON. Sur ces 5 sujets; un seul a ajouté que la grammaire était de trop dans un dictionnaire. Les 4 autres ont répondu qu'ils utilisaient d'autres sources d'information pour la grammaire.

Les utilisateurs du Van Dale anglais savent que leur dictionnaire contient des informations grammaticales, mais seul un petit nombre d'entre eux savent exactement de quel type il s'agit. Peu ont conscience des différences entre les deux volumes et la plupart des sujets ont du mal à définir exactement ce qu'est l'information grammaticale. La majorité des sujets affirment qu'ils ont recours à l'information grammaticale fréquemment, mais il est difficile de savoir s'ils l'utilisent vraiment. Après tout, la présentation de l'information grammaticale dans le Van Dale, et tout particulièrement dans le volume néerlandais > anglais, n'est pas vraiment destinée à faciliter la production de l'anglais.

## **6 Les résultats pour le Van Dale français ↔ néerlandais**

Tout comme sa version néerlandais ↔ anglais, le Van Dale néerlandais ↔ français est ce que l'on trouve de mieux sur le marché. Les principes qui régissent ces deux dictionnaires sont les mêmes et ils font partie de la même série. Le volume français > néerlandais est destiné aux utilisateurs dont la langue maternelle est le néerlandais et qui désirent comprendre le français écrit et parlé. Le volume néerlandais > français s'adresse à ceux qui veulent produire du français écrit ou parlé. L'introduction au volume néerlandais > français précise que ceci a des conséquences importantes sur la quantité d'informations requises pour chaque équivalent donné dans les entrées. Nous y reviendrons plus tard. Le volume français contient un précis grammatical de 92 pages classé selon l'ordre alphabétique. On peut trouver un résumé de 37 pages de ce précis à la fin du volume néerlandais > anglais.

Treize sujets ont répondu au questionnaire sur le Van Dale néerlandais ↔ français : 5 étudiants de première année, 5 étudiants de seconde année, 2 étudiants de quatrième année et 1 traducteur. Bien que le nombre total de sujets soit bien moins élevé que pour le Van Dale néerlandais ↔ anglais, ils sont également répartis sur les différents «niveaux».

A la question 6a, 11 sujets ont répondu qu'ils savaient que le Van Dale néerlandais ↔ français contenait des informations grammaticales, un seul a affirmé que les deux volumes n'en contenaient pas et un a répondu NE SAIT PAS (ces derniers étaient des étudiants de première année).

Le tableau n° 12 est du même type que le tableau n° 8. La troisième colonne indique quels types d'informations sont donnés par les deux volumes. Seuls les sujets ayant coché OUI à la question 6a auraient dû répondre aux questions 6b à 9, mais l'étudiant ayant choisi NON a également répondu. Cela explique pourquoi le nombre total de réponses données aux questions 6b et 9 s'élève à 12.

Tableau n° 12.

Type d'information grammaticale	Question 6b (12)		Question 9 (12)		Présentation systématique ?	
	nombre	%	nombre	%	N > F	F > N
indicateur de la catégorie grammaticale	6	50,0	7	58,3	+	+
forme des noms au pluriel	3	25,0	4	33,3	+	+
genre des noms	9	75,0	11	91,7	+	+
utilisation du singulier ou du pluriel	0	0	1	8,3	+	+
types de verbes	3	25,0	4	33,3	+	+
conjugaison des verbes	6	50,0	1	8,3	+	+
restrictions concernant l'usage	0	0	2	16,7	+	+
position par rapport aux autres mots	0	0	1	8,3	-	-
flexion des autres catégories grammaticales	3	3,0	1	8,3	+	+
formation des dérivés	0	0	3	25,0	+	+
combinaison avec d'autres mots	0	0	1	8,3	+	+
divers	4	33,3	0	0		

Comme les symboles + de la dernière colonne le montrent, les deux volumes du Van Dale néerlandais ↔ français contiennent de nombreux types d'information grammaticale. Le seul type qu'on ne trouve pas systématiquement concerne les «positions». Les entrées pour des adjectifs comme *ancien* et *petit* par exemple ne précisent pas explicitement s'ils doivent être placés avant ou après le nom. Ceci ne peut être déduit que grâce aux exemples donnés dans les entrées.

On a réuni ci-dessous divers exemples qui montrent comment l'information grammaticale est présentée dans le volume néerlandais > français.

**bespotting** <de (v.)> **0.1 moquerie** <v., vaak mv.>  
**diesI** <bw.>...<GC : voegwoord>

**hollijst** <de>...0.1...gueule de loup <v.> <getal 10>

**oog**...0.1...**oeil** <m.>, **yeux** <mv.>...

**precies I** <bn., bw.>...0.2...<bn.> **précis** ⇒ consciencieux, -euse,

minutieux, -euse, méticuleux, -euse; <bw.> **précisément** ⇒

consciencieusement, minutieusement, méticuleusement, avec précision...

**uitrukken II** <onov.ww.> 0.1... **se mettre** <t60 wk.> **en marche** ⇒ sortir <t38>

L'information grammaticale se trouve entre les signes <>. Dans l'entrée pour *bespotting*, on trouve à la fois le genre du nom néerlandais et de sa traduction française. L'entrée précise également que *moquerie* s'utilise généralement au pluriel. L'entrée de *dies* contient un code qui renvoie à la partie sur les conjonctions dans le précis grammatical du dictionnaire. A *hollijst*, on renseigne les utilisateurs sur le genre de l'équivalent français *gueule-de-loup* et sur la formation de son pluriel : le code «10» signifie que c'est la première partie du nom composé qui peut être mise au pluriel. Ce code renvoie également à l'entrée pour *Getal* (c'est-à-dire Nombre) dans le précis grammatical. A *oog*, on trouve la forme plurielle du nom *oel* imprimée en entier (l'abréviation «mv» signifie «pluriel»). A l'entrée *precies*, on donne en entier les formes masculin singulier et féminin singulier des équivalents et on fait clairement la distinction entre l'utilisation adverbiale et adjectivale du mot. Le dernier exemple concerne le premier sens du verbe *uitrukken*. S'il est utilisé de façon intransitive, il peut être traduit par trois verbes différents. Les deux références données à *se mettre* et *sortir* renvoient l'utilisateur aux tableaux de conjugaison 60 et 38 du précis grammatical. L'abréviation *wk.* signifie que *se mettre* est un verbe réflexif.

Il est donc clair que les deux volumes du Van Dale néerlandais ↔ français donnent beaucoup de types d'informations grammaticales. Les chiffres et les pourcentages donnés dans le tableau n° 12, pourtant, indiquent que beaucoup de sujets n'ont pas conscience de la richesse de l'information procurée. Une conclusion possible pourrait être que les utilisateurs des deux volumes du Van Dale néerlandais > français n'en mesurent pas la juste valeur. Dans le meilleur des cas, ils n'utilisent qu'un petit nombre des différents types d'information grammaticale donnés. Les seules catégories qui totalisent un score élevé aux questions 6b et 9 concernent «le genre» et «les catégories grammaticales». Sachant que le volume néerlandais > français procure fréquemment des informations sur la traduction des verbes, les résultats obtenus pour les catégories «types de verbes» et «inflexion des verbes», surtout à la question 9, sont donc particulièrement faibles. Comme pour le Van Dale néerlandais ↔ anglais, aucun des étudiants en français n'a mentionné des éléments non grammaticaux sous la rubrique «divers», bien que les quatre réponses données aient été un peu vagues : «voir le dictionnaire» (deux fois), «forme» et «fonction des mots».

Une preuve supplémentaire que les utilisateurs ne mesurent pas à sa juste valeur la grammaire dans les deux volumes du dictionnaire est que la majorité des 11 sujets qui ont répondu à la question 6c ne savaient pas que les cinq types d'information grammaticale donnés dans cette question se trouvent effectivement dans le Van Dale.

Tableau n° 13 : Comment l'information grammaticale se présente-t-elle ?

Forme de l'information grammaticale	Nombre de sujets (11)
-------------------------------------	--------------------------

	nombre	%
flexions complètes (par exemple forme des pluriels)	5	45,5
abréviations (par exemple «adv.» pour «adverbe»)	10	90,9
nombres (par exemple «122» pour «tableau 122»)	3	27,3
codes (par exemple «mv14» pour «pluriel, paragraphe 14»)	4	36,4
exemples	5	45,5

Bien que les abréviations représentent sans doute la forme d'information grammaticale la plus fréquente dans le Van Dale néerlandais ↔ français, les quatre autres types d'information sont suffisamment courants et importants pour qu'on les remarque. Si l'on considère que seulement deux sujets (tous deux en première année) ont coché les cinq caractéristiques et un seulement quatre d'entre elles, on peut dire que la plupart des sujets ignorent que leur dictionnaire contient un certain nombre d'éléments caractéristiques en ce qui concerne la grammaire.

Sur les 12 sujets qui ont répondu à la question 6d («Y a-t-il des différences entre les deux volumes du dictionnaire ?») : six ont répondu NON, deux NE SAIS PAS et quatre ont répondu OUI. Ceux qui ont répondu OUI ont donné les précisions suivantes : «voir le dictionnaire» (étudiant de première année), «conjugaison des verbes» (étudiant de première année), «bien que les types d'information soient les mêmes, un des deux volumes donne plus d'informations que l'autre» (étudiant de deuxième année), et «il y a moins de grammaire dans le volume français > néerlandais» (étudiant de deuxième année). La première remarque n'est pas utile et la seconde est fautive, mais la troisième et la quatrième sont correctes. Le tableau n° 12 montre que les deux volumes donnent les mêmes types d'information. La différence principale réside dans le fait que le volume français > néerlandais ne donne des informations que sur les mots vedettes alors que le volume néerlandais > français fournit une aide grammaticale pour beaucoup d'équivalents. Cela signifie que, dans une certaine mesure, le petit groupe d'étudiants apprenant le français se rend mieux compte des différences entre les deux volumes du Van Dale néerlandais > français que ceux qui utilisent fréquemment le Van Dale néerlandais > anglais.

Environ 20 % des sujets apprenant l'anglais (11 sur 53) avaient affirmé que leur dictionnaire ne contenait pas assez de grammaire. Seulement un sur les 12 sujets apprenant le français qui ont répondu à la question 7a a indiqué qu'il, ou

elle, n'était pas satisfait(e) par les deux volumes du Van Dale néerlandais ↔ français. Cet(te) étudiant(e) de première année a indiqué que les deux volumes devraient contenir plus d'information grammaticale, sans spécifier quel genre d'information.

Le tableau n°14 montre que plus de 80 % des sujets étudiant le français ont indiqué qu'ils utilisaient l'information grammaticale dans leur dictionnaire au moins une fois par semaine. Ceci concorde avec les résultats obtenus pour le Van Dale néerlandais ↔ anglais (cf tableau n° 11). La seule différence réside dans le fait que le nombre de sujets qui étudient le français et qui vérifient des points de grammaire au moins une fois par semaine dans leur dictionnaire est bien plus élevé (50 %), que celui des sujets qui étudient l'anglais et utilisent le Van Dale néerlandais ↔ anglais avec le même objectif (26,9 %).

Tableau n° 14 : Selon quelle fréquence utilisez-vous l'information grammaticale dans votre dictionnaire ?

Fréquence	Nombre de sujets (12)	
	nombre	%
une fois par jour	6	50,0
une fois par semaine	4	33,3
une fois par mois	0	0
moins d'une fois par mois	2	16,7
jamais	0	0

Comme pour le Van Dale néerlandais ↔ anglais, un trop grand nombre de sujets ont répondu à la question n° 10. Seulement un d'entre eux aurait dû y répondre, c'est-à-dire l'étudiant(e) qui avait répondu NON à la question 6a, alors que les 11 participants ont répondu. L'étudiant(e) qui était supposé(e) répondre à cette question a indiqué qu'il, ou elle, n'avait pas besoin que son dictionnaire fournisse d'information grammaticale. Sur tous les sujets, 9 ont répondu que le Van Dale néerlandais ↔ français devrait continuer à fournir des informations grammaticales. Un(e) étudiant(e) de seconde année a répondu que le dictionnaire devait se concentrer sur la question de savoir «quand un mot en particulier peut, ou ne peut pas, être employé» plutôt que sur la grammaire. Il, ou elle, a ajouté que la grammaire est ce que l'on trouve dans les livres de grammaire.

Les deux volumes du dictionnaire Van Dale néerlandais ↔ français contiennent plus d'informations grammaticales que leurs équivalents néerlandais ↔

anglais, particulièrement le volume néerlandais > français. Les types d'informations grammaticales contenus dans ce volume et la manière dont ils sont présentés montrent bien qu'il s'agit clairement d'un dictionnaire qui vise à la production. En règle générale, les sujets sont satisfaits de la partie grammaticale du dictionnaire et ils utilisent l'information qu'il fournit relativement souvent. Pourtant si l'on en juge par les réponses données, beaucoup de sujets ne savent pas apprécier à sa juste valeur la somme d'informations données. Tout d'abord, la plupart d'entre eux n'ont pas conscience que la grammaire se présente sous différentes formes et en des endroits différents. Ensuite, quelques types d'information seulement sont vus comme étant importants : genres et indicateurs des catégories grammaticales en particulier. Plusieurs autres types de marqueurs grammaticaux, comme la référence aux tableaux de conjugaison sont mentionnés par peu de sujets; ils ne semblent donc pas frapper l'utilisateur par leur utilité.

## 7 Les résultats pour le Prisma italien ↔ néerlandais

Les deux volumes du dictionnaire Prisma néerlandais ↔ italien format livre de poche sont bien plus petits que les deux volumes des dictionnaires Van Dale<sup>4</sup>. Bien que l'introduction du dictionnaire ne le précise pas, le Prisma s'adresse à un public plus vaste. Etant donné son prix abordable, on peut présumer qu'il est également acheté par un plus grand nombre de personnes non spécialistes des langues que les deux volumes du Van Dale. On peut d'ailleurs s'attendre à ce que ceci ait des conséquences sur la façon dont l'information grammaticale contenue dans le Prisma est présentée. Chaque volume se termine par un précis grammatical de 11 pages.

Comme le tableau n° 7 le montre, la majorité des huit étudiants de seconde année qui ont répondu au questionnaire pour le Prisma en italien connaissent bien le concept de grammaire dans les dictionnaires : 6 étudiants sur 8 n'ont pas consulté les exemples donnés à la page 4 du questionnaire. De plus, les résultats à la question 6a montrent qu'ils savent quel type d'information grammaticale ils peuvent trouver dans leur dictionnaire : les huit étudiants ont précisé qu'ils savaient que les entrées des deux volumes de leur dictionnaire contenaient des informations grammaticales.

4 Ils sont aussi considérablement moins chers. Un volume du Van Dale coûte généralement 145 NGL, alors que le dictionnaire de poche Prisma ne coûte que 17,50 NGL.

Tableau n° 15

Type d'information grammaticale	Question 6b (8)	Question 9 (8)	Présentation systématique ?
---------------------------------	-----------------	----------------	-----------------------------

	nombre	%	nombre	%	N > I	I > N
indicateur de la catégorie grammaticale	3	37,5	5	67,5	+	+
forme des noms au pluriel	4	50,0	0	0	+	+
genre des noms	5	67,5	7	87,5	+	+
utilisation du singulier ou du pluriel	0	0	2	25,0	+	+
types de verbes	0	0	0	0	+	+
conjugaison des verbes	5	67,5	2	25,0	+	+
restrictions concernant l'usage	0	0	0	0	+	-
positions en fonction des autres mots	0	0	1	12,5	-	-
flexion des autres catégories grammaticales	2	25,0	0	0	-	+
formation de dérivés	0	0	1	12,5	+	+
combinaison avec d'autres mots	1	12,5	6	75,0	+	+
divers	4	50,0	0	0		

Comme on peut le constater dans les deux dernières colonnes, les deux volumes du dictionnaire Prisma contiennent une grande variété d'informations grammaticales. Cependant, et le tableau ne le montre pas, les types d'information donnés le plus souvent dans les deux volumes concernent les indicateurs des catégories grammaticales, la forme plurielle des noms, les marqueurs de genre, et les différentes conjugaisons possibles (dans le précis grammatical, avec un système de numéros donnés après les verbes). Si l'on considère que ces quatre catégories font partie des types d'information grammaticale les plus courants et qu'on les compare avec les catégories mentionnées à la question 6b et cochées à la question 9, on peut tirer deux conclusions générales.

Tout d'abord, les réponses données par les étudiants à la question 6a se recourent avec les types d'informations grammaticales les plus fréquemment présentés dans les dictionnaires. La seule disparité concerne la catégorie «divers». Cette catégorie contenait les remarques suivantes : «articles» (deux fois), «exemples», et «utilisation des différents types de mots».

Ensuite, lorsqu'ils se sont trouvés en présence d'une liste d'options à la question 9, les étudiants ont transféré leurs choix concernant l'utilité de certaines catégories vers d'autres types d'information. Personne n'a coché la catégorie «formes des noms au pluriel», alors qu'elle avait été mentionnée par quatre sujets à la question 6a, la catégorie «genre» a obtenu des résultats encore plus importants que pour la question 6a (respectivement 67,5 % et 87,5 %), la catégorie «conjugaison des verbes», au contraire a obtenu de moins bons résultats à la question 9 (25 % au lieu de 67,5 % à la question 6a); et bien qu'un seul étudiant ait mentionné la catégorie «combinaison avec d'autres mots» à la question 6a, la majorité des sujets (75 %) ont considéré cette catégorie comme un aspect très important des entrées de dictionnaires.

Les deux résultats les plus frappants concernent les catégories «conjugaison des verbes» et «combinaison avec d'autres mots». Tout comme dans les deux volumes du dictionnaire Van Dale français ↔ anglais, les numéros renvoyant aux types de conjugaison des verbes italiens (et que l'on trouve dans les deux volumes) font de Prisma un véritable dictionnaire de production. Les numéros devraient empêcher les utilisateurs d'utiliser les mauvaises formes verbales lorsqu'ils écrivent en italien.

**nadagen\*** in zijn - zijn *essere* (46) in *declino, essere* (46) in *decadenza*  
**uitlachen** *deridere* (25), *beffarsi* (1) *di\** iem. in zijn gezicht - *ridere* (25) in *faccia a ql.*

Bien que les réponses données à la question 6a montrent que les étudiants savent que leur dictionnaire contient les indications nécessaires pour utiliser les bonnes conjugaisons, les résultats obtenus à la question 9 semblent indiquer que les étudiants considèrent ces indications comme moins importantes que d'autres types d'information (voir également le tableau n° 16). Ce résultat est étonnant, si l'on sait que les sujets n'étudiaient l'italien que depuis un an et demi lorsqu'ils ont rempli le questionnaire. Cela signifie que les étudiants savent déjà comment conjuguer un verbe en italien et n'ont donc pas besoin de vérifier dans leur dictionnaire, ou bien qu'ils considèrent la conjugaison des verbes comme moins importante que ce à quoi pourraient s'attendre leurs professeurs. Une troisième explication pourrait être que les étudiants n'ont pas envie de prendre la peine de vérifier la signification de codes numériques peu transparents. En effet, les types d'information qui sont présentés de façon explicites dans le dictionnaire, comme les flexions et les abréviations claires obtiennent d'excellents scores à la question 9.

Les collocations représentent un type d'information que les sujets estiment très important. Sur les 8 sujets, 6 ont coché cette catégorie à la question 9. L'un des

deux étudiants qui n'avaient pas choisi cette catégorie a indiqué : «Malheureusement, Prisma ne fournit pas cette information», en soulignant l'adverbe.

A la question 7a, quatre des sujets (50 %) ont indiqué que le dictionnaire Prisma ne contenait pas assez d'informations grammaticales. La catégorie des «collocations» est mentionnée à trois reprises. Selon les étudiants, les types d'information suivant devraient être plus largement utilisés dans le dictionnaire Prisma :

- (i) «plus d'exemples et plus de significations»;
- (ii) «utilisation du singulier/pluriel, combinaison avec par exemple les prépositions; restrictions concernant l'usage; et tous les autres points mentionnés à la question 9, parce que Prisma ne donne aucune information dans la plupart des entrées»,
- (iii) «collocations; utilisation du mot; phrases d'exemples; combinaison avec d'autres mots et d'autres catégories grammaticales», et
- (iv), «plus d'exemples montrant quelles prépositions doivent être utilisées».

De toute évidence, les étudiants d'italiens savent ce qu'ils veulent.

Le tableau n° 16 montre que les sujets savent également comment la grammaire est présentée dans le dictionnaire Prisma.

Tableau n° 16 : Comment l'information grammaticale se présente t-elle ?

Forme de l'information grammaticale	Nombre de sujets (8)	
	nombre	%
flexions complètes (par exemple forme des pluriels)	6	75,0
abréviations (par exemple «adv.» pour «adverbe»)	7	87,5
nombres (par exemple «122» pour «tableau 122»)	3	37,5
codes (par exemple «mv14» pour «pluriel, paragraphe 14»)	0	0
exemples	5	67,5

Les quatre types d'information contenus dans le dictionnaire Prisma ont obtenu des scores relativement élevés. Les trois premiers types d'information qui sont énumérés dans le tableau sont utilisés le plus fréquemment dans les deux volumes du dictionnaire. Le nombre d'étudiants qui ont mentionné la catégorie «nombres» est relativement important. Bien que les professeurs et les lexicogra-

phes puissent considérer que le système des nombres mentionnés après les verbes est une composante utile des dictionnaires, peu d'utilisateurs semblent l'apprécier. La situation est semblable pour le Van Dale néerlandais <-> français.

Le tableau n° 17 explique pourquoi les huit étudiants en italien connaissent si bien leur dictionnaire : ils en utilisent la grammaire très fréquemment.

Tableau n° 17 : Selon quelle fréquence utilisez-vous l'information grammaticale dans votre dictionnaire ?

Fréquence	Nombre de sujets (8)	
	nombre	%
une fois par jour	4	50,0
une fois par semaine	4	50,0
une fois par mois	0	0
moins d'une fois par mois	0	0
jamais	0	0

80 % des étudiants en anglais et français affirment qu'ils utilisent la grammaire dans leur dictionnaire au moins une fois par semaine. Les étudiants de seconde année en italien totalisent 100 %.

Quatre des sujets ont répondu à la question 10a, alors qu'aucun n'avait à le faire. Bien qu'ils aient tous indiqué que leur dictionnaire d'italien devrait contenir plus d'informations grammaticales, aucun des quatre sujets n'a pu détailler précisément ce qui manquait. Trois d'entre eux ont simplement noté «voir question 6», référence infructueuse, puisque la question 6 ne demandait pas quel type d'information grammaticale manquait.

Malgré leur prix modeste, les deux volumes du dictionnaire Prisma italien ↔ néerlandais contiennent un nombre important d'informations grammaticales. En ce qui concerne la grammaire, ils peuvent même, en fait, rivaliser avec les Van Dale, beaucoup plus gros et beaucoup plus chers. Il y a pourtant une exception : les étudiants d'italien ont critiqué leur dictionnaire pour le peu de renseignements sur les prépositions à utiliser dans les collocations, alors que le Van Dale fournit des informations très claires à ce sujet.

Il y a une autre différence bien plus importante entre les dictionnaires Van Dale et les dictionnaires Prisma. Parce qu'ils sont abordables, ces derniers sont achetés par un lectorat plus varié. Par conséquent, une partie de l'information de grande valeur qui s'y trouve (autrement dit les codes numériques renvoyant aux

conjugaisons des verbes) peut paraître trop énigmatique, trop difficile d'accès à beaucoup d'utilisateurs. Si des étudiants d'italien de seconde année, motivés, n'utilisent pas ou n'apprécient pas totalement ces références codées, d'autres types d'utilisateurs peuvent ne pas les utiliser du tout.

Le dictionnaire Prisma contient beaucoup d'informations grammaticales essentielles, même pour l'utilisateur ordinaire. Si les étudiants en traduction sont un peu critiques à son égard, c'est peut-être parce qu'ils attendent trop d'un dictionnaire bilingue à vocation généraliste.

## 8 Les résultats pour le Wehr arabe > anglais

Le *Dictionary of Modern Written Arabic* arabe > anglais (1960) est basé sur un dictionnaire arabe > allemand publié en 1952. Le dictionnaire utilisé par les étudiants d'arabe à l'École de Traduction et d'Interprétation de Maastricht en est la troisième édition (1976). C'est un dictionnaire qui «sera le bienvenu non seulement pour les utilisateurs anglais et américains, mais aussi pour les orientalistes de par le monde qui sont plus à l'aise en anglais qu'en allemand» (Préface). La grammaire joue un rôle important dans la présentation des entrées du dictionnaire. Le premier principe qui règle la disposition des mots vedettes concerne les radicaux arabes, qui sont classés selon l'ordre alphabétique. Le principe suivant consiste à regrouper les formes apparentées, toujours selon des principes grammaticaux.

Pour un radical donné, la disposition des entrées se fait comme suit. Le verbe au parfait du radical de base, s'il existe, est placé en premier avec la translittération indiquant la vocalisation. Il est suivi par la voyelle à l'imparfait et, entre parenthèses, les noms verbaux ou *masadir*. Ensuite viennent les radicaux dérivés, indiqués par les chiffres romains en caractère gras allant de II à X. (Introduction : XIII).

Les entrées donnent également des informations sur les cas, les formes du pluriel, les prépositions à utiliser dans les collocations et le genre.

Sept sujets ont répondu au questionnaire pour le dictionnaire Wehr : six étudiants de quatrième année et un professeur d'arabe. Malgré l'importance de la grammaire dans le dictionnaire, un étudiant, parmi les six, ne savait pas qu'il en contenait. Le tableau n° 7 a déjà montré que les 6 étudiants avaient consulté les exemples d'entrées données à la page 4 du questionnaire pour vérifier si leur conception de ce qu'est la composante grammaticale dans un dictionnaire était correcte. Le tableau n° 18 montre à quel point les sujets connaissent bien leur dictionnaire. L'étudiant qui ne savait pas si le dictionnaire contenait de la grammaire ou non n'a pas répondu aux questions 6b à 9. Seuls les résultats de 6 participants apparaissent donc dans le tableau.

Tableau n° 18

Type d'information grammaticale	Question 6b (6)		Question 9 (8)		Présentation systématique ?
	nombre	%	nombre	%	AR > AN
indicateur de la catégorie grammaticale	2	33,3	0	0	+
forme des noms au pluriel	4	66,7	6	100,0	+
genre des noms	2	33,3	2	33,3	+
utilisation du singulier ou du pluriel	0	0	0	0	+
types de verbes	0	0	1	16,7	+
conjugaison des verbes	5	83,3	6	100,0	+
restrictions concernant l'usage	0	0	1	16,7	+
positions en fonction des autres mots	0	0	0	0	-
flexion des autres catégories grammaticales	1	16,7	0	0	-
formation de dérivés	0	0	3	50,0	+
combinaison avec d'autres mots	2	33,3	1	16,7	+
divers	1	16,7	1	16,7	

Il n'est pas surprenant que les deux catégories «formes des noms au pluriel» et «conjugaison des verbes» totalisent un maximum de 100 % à la question 9. Elles représentent en effet des types d'informations grammaticales essentielles pour ceux qui apprennent l'arabe. Cependant, deux résultats sont frappants. Le premier concerne la catégorie «formation des dérivés» qui totalise 0 % à la question 6b et seulement 50 % à la question 9. On aurait pu s'attendre à un score maximal ici aussi, puisque les entrées du dictionnaire sont organisées selon les formes radicales communes. Le second résultat étonnant concerne la catégorie

«combinaison avec d'autres mots». Seul le professeur d'arabe a mentionné cette catégorie à la fois à la question 6b et à la question 9. Aucun des étudiants ne l'a cochée à la question 9.

Tableau n° 19 : Comment l'information grammaticale se présente-t-elle ?

Forme de l'information grammaticale	Nombre de sujets (6)
-------------------------------------	-------------------------

	nombre	%
flexions complètes (par exemple forme des pluriels)	4	66,7
abréviations (par exemple «adv.» pour «adverbe»)	3	50,0
nombres (par exemple «122» pour «tableau 122»)	1	16,7
codes (par exemple «mv14» pour «pluriel, paragraphe 14»)	0	0
exemples	3	50,0

La plupart des informations grammaticales sont présentées assez explicitement dans le Wehr. Les prépositions, les notes sur l'usage et les formes du pluriel des noms apparaissent en entier (les formes du pluriel des noms sont également précédées de l'abréviation *pl*). Les catégories grammaticales des mots et les notes occasionnelles sur les genres apparaissent sous formes d'abréviations, comme *prep.*, *conj.*, ou *f*. Les chiffres (romains) ne sont utilisés que pour les radicaux dérivés. Par conséquent, les sujets auraient dû cocher les catégories «flexions pleines», «abréviations» et peut-être aussi la catégorie «exemples». Tous les sujets étudiaient l'arabe à un niveau avancé et savaient donc ce qui est important pour eux (voir tableau n° 18), mais beaucoup d'entre eux ne semblaient pas savoir sous quelle forme la grammaire est présentée dans le dictionnaire.

Le tableau n° 20 montre avec quelle fréquence les sujets utilisent les informations grammaticales de leur dictionnaire. Le nombre de réponses données est de 7 au lieu de 6 (l'étudiant qui avait coché NE SAIS PAS, à la question 6a, a répondu JAMAIS à la question 8).

Tableau n° 20 : Selon quelle fréquence utilisez-vous l'information grammaticale de votre dictionnaire ?

Fréquence	Nombre de sujets (7)	
	nombre	%
une fois par jour	4	57,1
une fois par semaine	2	28,6
une fois par mois	0	0
moins d'une fois par mois	0	0
jamais	1	14,3

Il n'est pas surprenant de constater que l'information grammaticale du dictionnaire arabe > anglais est utilisée très fréquemment. En règle générale, les étudiants et le professeur d'arabe utilisent la grammaire de leur dictionnaire bien plus souvent que les utilisateurs des autres dictionnaires mentionnés dans ce texte.

*A Dictionary of Modern Written Arabic* est un dictionnaire unidirectionnel. Bien que ce soit un dictionnaire de décodage plus que d'encodage, l'information grammaticale qu'il contient est utilisée bien plus souvent que celle contenue dans n'importe quel autre dictionnaire mentionné dans cette enquête, car la connaissance de la grammaire est essentielle pour les étudiants d'arabe. Les 7 sujets le savent, mais la plupart d'entre eux n'ont pas encore pris pleinement conscience de toute l'aide qu'ils peuvent trouver dans leur dictionnaire. Seul un petit nombre d'entre eux a mentionné une ou deux caractéristiques importantes du dictionnaire Wehr.

## 9 Conclusion

Un des objectifs principaux de l'enquête était de voir, à travers les réponses à un questionnaire soumis à 129 sujets, ce que les utilisateurs de dictionnaires bilingues, d'un niveau relativement avancé savent de l'information grammaticale contenue dans leurs dictionnaires. Une comparaison a été faite entre les déclarations des sujets et la réalité telle qu'elle apparaît dans les dictionnaires examinés. Etant donné que les sujets étaient interrogés sur des descriptions, plutôt que sur des exemples, ils ont pu trouver les questions plutôt abstraites. C'est pourquoi il faut considérer les résultats de cette enquête comme fournissant une

vision d'ensemble de la perception de l'information grammaticale contenue dans les dictionnaires bilingues.

La plupart des étudiants et des professionnels qui ont participé à l'enquête ont une idée approximative des types d'informations grammaticales présents dans leurs ouvrages de référence, bien que pour plusieurs étudiants de première année la grammaire dans un dictionnaire regroupe tout ce qui va du proverbe à la prononciation. Cependant, les types d'information que les utilisateurs néerlandais de dictionnaires bilingues disent utiliser le plus souvent se réduisent dans l'ensemble aux catégories traditionnelles : indicateurs de catégorie grammaticale, marqueurs de genre, formes du pluriel et flexions verbales. Et lorsqu'on leur demande comment l'information grammaticale est présentée dans les dictionnaires, beaucoup de sujets mentionnent un nombre encore plus restreint de catégories. Ils se rappellent la présence des formes entières et des formes abrégées, mais beaucoup d'entre eux ont oublié que les entrées verbales, par exemple, dans beaucoup de dictionnaires, contiennent des renvois chiffrés aux appendices et aux précis grammaticaux contenus dans les ouvrages. Ce que les partisans de la grammaire dans les dictionnaires considèrent comme de l'information grammaticale utile et importante semble donc passer inaperçu aux yeux d'un grand nombre d'utilisateurs. On peut même tirer une conclusion encore plus pessimiste. Les utilisateurs qui devraient bénéficier le plus de l'aide apportée par les dictionnaires (c'est-à-dire les étudiants de première année), semblent apprécier le moins les informations sur le comportement des mots. Les codes et les chiffres qui servent de renvois peuvent apparaître trop compliqués pour la plupart des utilisateurs, et si l'on veut que la grammaire dans les dictionnaires soit utile pour la majorité des utilisateurs, il faut qu'elle soit aussi explicite que possible.

L'enquête révèle aussi des points positifs. Les utilisateurs néerlandais débutants et avancés de dictionnaires bilingues semblent s'attendre à trouver de la grammaire dans leurs ouvrages et ils l'utilisent relativement fréquemment. En fait, pour certaines langues (comme l'arabe), la grammaire se révèle être un élément essentiel dans l'emploi d'un dictionnaire de production. De plus, certains utilisateurs ont une idée très précise de ce qui manque dans leur dictionnaire : ils demandent plus d'informations sur les collocations et sur la combinaison des mots entre eux. Cependant, ils ne demandent pas davantage de tableaux de conjugaison et de notes d'usage et ils ne semblent pas non plus demander des codes et des structures syntaxiques comme celles qui sont données dans les dictionnaires monolingues d'anglais pour apprenants.

## Remerciements

J'aimerais remercier Geer Hoppenbrouwers qui m'a aidé à convertir de longues listes de réponses écrites en données statistiques valables, et Martin Custers qui m'a expliqué l'organisation du *Dictionary of Modern Written Arabic*.

Tableau N° 21 : Nombre de sujets ayant répondu à l'enquête

Question	Nombre de questions auxquelles les sujets ont répondu	Questions auxquelles il fallait répondre	Différence
6a	52	69	-17
6b	39	52	-13
6c	53	52	+1
6d	42	52	-10
6e	10	11	-1
7a	53	52	-16
7b	10	11	-1
8	52	52	0
9	52	52	0
10a	0	44	+44
10b	25	39	-14
10c	5	5	0

# ÉLABORATION ET RÉUTILISATION DES EXEMPLES DANS LA LEXICOGRAPHIE BILINGUE

Xavier BLANCO

### 1 Introduction

La presque totalité des dictionnaires généraux français > espagnol, espagnol > français prévoient l'utilisation d'exemples dans leurs programmes de micro-structure. D'une part, on en vante la multifonctionnalité : l'exemple permettrait d'offrir des renseignements précieux sur toutes les facettes du mot-vedette (aspects morphologiques – e.g. pluriels irréguliers –, syntaxiques – e.g. régime, changement de diathèse –, lexicaux – e.g. cooccurrence restreinte, champs analogiques –, sémantiques – e.g. discrimination entre différentes acceptions du mot-vedette –, pragmatiques – e.g. présentation de *Routineformeln* –). D'autre part, l'exemple excellerait par sa capacité à "illustrer" une information déjà présentée par d'autres catégories d'information lexicographique, en aidant l'utilisateur d'abord à la décoder et ensuite à s'en servir convenablement (Blanco, 1995b:305).

Il serait cependant nécessaire de se demander si l'exemple est toujours conçu de façon à pouvoir assurer la lourde tâche qu'il est censé remplir. En effet, la comparaison de différents dictionnaires permet de constater que, très souvent, l'exemple n'est pas élaboré en vue de compléter ou d'illustrer la description linguistique proposée dans la microstructure qui va l'accueillir, mais qu'il est repris d'autres compilations déjà existantes qui affichent des objectifs et des procédés parfois très éloignés de ceux du dictionnaire hôte. Adopté, mais pas toujours adapté, l'exemple n'est plus à même de jouer le rôle qu'il se voit pourtant toujours attribuer et devient soit un simple agrément de l'article, soit, à la limite, un élément étranger à la prédication métalinguistique portant sur le mot-vedette.

Nous nous proposons de montrer quelles sont les dimensions du phénomène de la réutilisation de l'exemple, ainsi que de sa traduction, dans la lexicographie bilingue français > espagnol, espagnol > français. Nous pensons que l'étude des mécanismes de récupération de l'exemple peut permettre, d'une part, de préciser

le statut de cette catégorie d'information lexicographique et, d'autre part, de mettre en évidence des sérieux défauts dans l'exemplification qui découlent d'une réutilisation inadéquate ou d'une adaptation forcée de certaines séquences-exemples.

Pour cela, nous travaillerons à partir d'un corpus de base formé de 4070 exemples et de leurs traductions tirés de cinq dictionnaires bilingues français > espagnol et de cinq dictionnaires bilingues espagnol > français. Il s'agit de la totalité des exemples présents sous la lettre «f» dans les ouvrages signalés ci-dessous.

- LFE-LEF<sup>1</sup> GARCÍA-PELAYO R.; GROSS, R. et J. TESTAS (1992) : *Grand Dictionnaire Français-Espagnol, Espagnol-Français*, Paris, Larousse.
- BFE-BEF VIDAL, J.-P. (1989) : *Dictionnaire Français-Espagnol / Espagnol-Français Hispano Bordas*, Paris, Bordas.
- HFE-HEF DENIS, S.; MARAVALL, M. et L. POMPIDOU (1976/1960) : *Dictionnaire Espagnol-Français et Français-Espagnol*, Paris, Hachette.
- VFE-VEF V.V.A.A. (1990) : *Diccionario Manual Francés-Español Español-Francés*, Barcelona, Vox Biblograf.
- MFE-MEF GARCÍA NAVARRO, A.M. et J. CLERC (1988) : *Diccionario moderno Francés*, Barcelona, Herder.

Précisons, également, que nous partirons de la définition opératoire suivante (Blanco, 1995a : 131) pour le terme «exemple lexicographique (dans le dictionnaire bilingue)» :

Un exemple est un élément lexicographique de microstructure (optionnel et, normalement, pluriverbal) qui

- contient une des formes comprises dans le paradigme du mot-vedette;
- est caractérisé par certaines marques typographiques (italique, gras, caractères séparateurs à valeur métalinguistique...);
- comporte une traduction.

---

<sup>1</sup> Par souci de brièveté, nous nous servirons désormais des abréviations indiquées ici pour nous référer aux différents ouvrages. Nous distinguons entre les dictionnaires français > espagnol (FE) et les dictionnaires espagnol > français (EF).

## 2 D'un dictionnaire à l'autre

Le parcours des exemples présents sous le traitement du «f» dans les dictionnaires cités permet de déceler la présence de certaines séquences répétées. On constate, en effet, pour les dictionnaires français > espagnol (2580 séquences considérées), l'existence de 216 cas d'exemples (identiques)<sup>2</sup> présents dans plus d'un ouvrage. Concrètement, nous avons la situation suivante :

- séquences répétées dans 5 dictionnaires : 2<sup>3</sup>
- séquences répétées dans 4 dictionnaires : 10
- séquences répétées dans 3 dictionnaires : 30
- séquences répétées dans 2 dictionnaires : 174

Quant aux dictionnaires espagnol > français (1490 séquences considérées), il existe 91 cas d'exemples répétés; à savoir :

- séquences répétées dans 5 dictionnaires : 0
- séquences répétées dans 4 dictionnaires : 4
- séquences répétées dans 3 dictionnaires : 16
- séquences répétées dans 2 dictionnaires : 71

Etant donné les contraintes d'espace, nous n'introduisons pas la liste des séquences ni les calculs des taux de répétition entre les différentes paires de dictionnaires. Nous nous bornerons aux deux constatations suivantes :

- la grande majorité des séquences répétées correspond à des unités lexicales complexes (surtout à des noms composés). Le caractère codifié de ces séquences fait que leur réapparition est moins significative d'un point de vue lexical (reprise d'une unité appartenant au système de la langue), mais non d'un point de vue lexicographique. En effet, des exemples identiques figurent dans des microstructures fort différentes, ce qui met en question l'adaptation de ces séquences aux descriptions lexicographiques proposées par les différents dictionnaires. Par ailleurs, il convient de remarquer qu'un traitement correct de composés tels que *jeux floraux* (LFE et BFE) ou *fiesta nacional* (LEF et MEF) voudrait que ceux-ci fussent considérés comme des sujets métalinguistiques et non comme une partie de la prédication métalinguistique portant sur les lemmes **floral**, e *adj* et **fiesta** *n.f.*;
- certains des exemples répétés présentent un caractère marqué, comme *il a un faible pour sa fille cadette* (BFE et VFE), *sólo le faltaba convencer a sus padres* (BEF, VEF) ou *fumarse la clase* (LEF et HEF). Ces séquences semblent indiquer l'existence de certains courants de filiation d'exemples entre les dictionnaires.

2 Nous considérons uniquement les séquences identiques. Des paires comme HFE *faire une farce à* vs MFE *faire une farce à qqn* ou BFE *c'est (de) ma faute* vs VFE *c'est de ma faute, c'est ma faute* ne seront donc pas retenues ici malgré leur proximité.

3 Il s'agit de *fosse commune* et de *prendre le frais*.

Soulignons que les dictionnaires qui présentent un taux de répétition plus élevé entre leurs corpora respectifs de séquences sont aussi ceux qui présentent des exemples répétés qui choquent par leur caractère non codifié. Il s'agit, d'une part, des paires BFE-VFE (38 séquences communes) et BEF-VEF (33 séquences communes), qui partagent des suites comme celles citées ci-dessus ou encore *le poisson est en train de frire* (BFE, VFE); *hablar en tono festivo* (BEF, VEF), *no me fío de él* (BEF, VEF).

D'autre part, nous avons la paire LEF-HEF (12 séquences communes), avec des coïncidences comme *una noticia fiambre* ou *reloj fiel*. Dans le premier cas, il faut signaler que les séquences repérées aussi bien dans le Bordas 1989 que dans le Vox Biblograf 1992 (BFE-BEF et VFE-VEF, deux ouvrages sans relation apparente) apparaissent déjà dans un dictionnaire français > espagnol, espagnol > français édité conjointement par Biblograf SA (Barcelone) et Bordas (Paris) en 1976. Par contre, elles sont absentes du Vox Biblograf de 1974 (sixième édition du *copyright* de 1969). On remarquera également que Jean-Paul Vidal, auteur du Bordas 1989, signe aussi le Vox Biblograf de 1976. Cette situation pourrait expliquer la répétition de certains exemples très typés dans le BFE-BEF et le VFE-VEF.

Dans le deuxième cas, il n'est pas possible de déceler un ancêtre commun (bilingue) au LEF et au HEF. Cependant, il est tout à fait révélateur que les séquences répétées dans ces deux ouvrages se trouvent aussi dans le *Diccionario de la Real Academia Española*<sup>4</sup>, ce qui permet de considérer ce dernier dictionnaire comme la source d'inspiration commune au LEF et au HEF.

### 3 Du dictionnaire français > espagnol au dictionnaire espagnol > français : une réversibilité problématique

La répétition de l'exemple dans l'univers lexicographique qui nous occupe ne se borne pas au réemploi des exemples d'un dictionnaire français > espagnol ou espagnol > français par un dictionnaire traitant de la même paire de langues et dans le même sens de traduction. En effet, étant donné que tous les exemples du corpus abordé sont dotés d'une traduction, il se trouve que le tout formé par l'exemple et sa traduction dans un dictionnaire donné français > espagnol apparaît inversé dans le dictionnaire espagnol > français qui forme un ensemble avec le premier.

4 Aussi bien dans l'édition de 1970 que dans celle de 1984, éditions immédiatement antérieures au HEF (1976) et au LEF (1992). Cependant, et étant donné que ces deux ouvrages reprennent à leur tour ces exemples du Hachette *Espagnol-Français* 1968 et de la série de dictionnaires modernes Larousse (e.g. *Dictionnaire moderne Larousse* 1967), il se peut que la source première de ces séquences soit l'édition de 1956 du *Diccionario de la Real Academia Española*.

Ainsi, 54 séquences L1-L2 du LFE (11,06 % du total des exemples présents sous la lettre «f») apparaissent inversées dans le LEF (e.g. LFE *distinguer le faux du vrai, distinguir lo falso de lo verdadero*; LEF *distinguir lo falso de lo verdadero, distinguir le faux du vrai*). Les chiffres correspondant aux différents dictionnaires de notre corpus sont :

– LFE-LEF :	54	(11,06 %)
– BFE-BEF :	34	(5,48 %)
– HFE-HEF :	2	(0,45 %)
– VFE-VEF :	7	(4,43 %)
– MFE-MEF :	18	(2,06 %)

En dehors des cas qui correspondent à des composés techniques (comme LFE *aldéhyde formique, aldehído fórmico*; LEF *aldehído fórmico, aldéhyde formique*), où la réversibilité peut être considérée comme totale aussi bien d'un point de vue lexicologique que lexicographique, le phénomène observé semble comporter des risques. Ainsi, si un exemple comme LFE *conclusion qui présente des faiblesses, conclusión que ofrece puntos flacos* s.v. **faiblesse** n.f. peut se révéler convenable pour l'illustration du sens de *faiblesse* correspondant à 'défaut de qualité (d'une œuvre d'art, d'une production de l'esprit)<sup>5</sup>, la séquence LEF *conclusión que ofrece puntos flacos, conclusion qui présente des faiblesses*, située s.v. **flaco**, ca adj comme illustration du composé *punto flaco* semble forcée en tant qu'exemple, surtout si l'on songe au fait que *punto débil* est beaucoup plus commun dans le sens de 'aspecto o parte más vulnerable de alguien o de algo'<sup>6</sup>.

De la même façon, d'autres inversions de séquences comme LFE *les flashes de l'actualité, los fogonazos de la actualidad*; LEF *los fogonazos de la actualidad, les flashes de l'actualité* ou LFE *les voitures se sont heurtées de plein fouet, los coches chocaron de frente*; LEF *los coches chocaron de frente, les voitures se sont heurtées de plein fouet* résultent d'une pertinence bien contestable.

Si le procédé consistant à présenter des exemples inversés peut apparaître dans certains cas comme délicat, le fait d'adapter la séquence L2 d'un dictionnaire donné afin de pouvoir réexploiter l'exemple dans le dictionnaire inverse s'avère déjà tout à fait critiquable. Nous avons des cas comme LFE *franchir un obstacle, salvar un obstáculo* face à LEF *franquear un obstáculo, franchir un obstacle*. Le LEF ignore l'équivalence *franchir* ↔ *salvar* et a recours au gallicisme *franquear*, afin de pouvoir profiter d'un exemple tout fait. Il ne s'agit pas, par ailleurs, d'un cas isolé, puisque nous trouvons aussi LFE *problème facile, problema sencillo* face à LEF *problema fácil, problème facile*; LFE *fermeté de caractère, entereza de carácter* face à LEF *firmeza de carácter, fermeté de caractère*; LFE *le foyer de la rébellion, el centro de la rebelión* face à LEF *el foco de la rebelión, le foyer de la rébellion*; LFE *la fureur du jeu, la pasión del juego* face à LEF *el furor del*

5 Cf. **faiblesse** (PR).

6 DRAE 1992. Précisons que *punto flaco* n'apparaît pas dans ce dictionnaire ni s.v. **punto** ni s.v. **flaco**.

*juego, la fureur du jeu*, etc. Dans tous ces cas, le LEF semble désavouer les traductions du LEF. En fait, il s'agit bien de réexploiter des exemples déjà construits, même au détriment de la description lexicographique.

Signalons que, bien qu'il s'agisse d'un phénomène quantitativement plus important dans le LFE-LEF que dans les autres dictionnaires de notre corpus, il n'est nullement l'apanage de ce seul ouvrage. Ainsi, nous avons BFE *tomber de fatigue, caerse de cansancio* face à BEF *caerse de fatiga, tomber de fatigüe*; HFE *froncer les sourcils, fruncir* ("ou" *arrugar el ceño*) ("ou" *el entrecejo*) face à HEF *fruncir las cejas, froncer les sourcils*; MFE *fonts baptismaux, pila bautismal* face à MEF *fuenta bautismal, fonts baptismaux*.

Précisons que nous trouvons le nombre de cas suivant dans chacun des dictionnaires examinés et toujours pour le traitement de la lettre «f» :

– LFE-LEF :	9	(1,84 %)
– BFE-BEF :	4	(0,64 %)
– HFE-HEF :	3	(0,67 %)
– VFE-VEF :	0	(0 %)
– MFE-MEF :	15	(1,72%)

Dans les cas envisagés ci-dessus, la séquence L2 d'un dictionnaire français > espagnol était adoptée, avec certaines modifications, comme séquence L1 par un dictionnaire espagnol > français. Il faudrait également rendre compte des cas où c'est la séquence L2 du dictionnaire espagnol > français qui se retrouve, de nouveau avec certains changements, comme séquence L1 dans un dictionnaire français > espagnol. Nous avons donc : LFE *conséquence forcée, consecuencia forzosa* vs LEF *consecuencia forzosa, conséquence inévitable*; BFE *feindre la surprise, fingir sorpresa* vs BEF *fingir sorpresa, feindre l'étonnement*; HFE *toile de fond, telón de foro* vs HEF *telón de foro, décors du fond*; MFE *tenter fortune, probar fortuna* vs MEF *probar fortuna, tenter sa chance*. L'impression d'incohérence est évidente. Signalons, cependant, que le phénomène décrit est quantitativement peu important :

– LFE-LEF :	2	(0,40 %)
– BFE-BEF :	2	(0,32 %)
– HFE-HEF :	3	(0,67 %)
– VFE-VEF :	0	(0 %)
– MFE-MEF :	7	(0,80 %)

Signalons aussi l'existence de deux autres types particuliers de répétition de séquences. Le premier consiste en la reprise totale d'un exemple par un même dictionnaire sous différents lemmes<sup>7</sup>. Il s'agit de cas tels que BFE *fer forgé, hierro forjado* s.v. **fer** n.m. et s.v. **forger** v.tr.; HFE *le fond et la forme, el fondo y la forma*

7 Il existe, dans notre corpus, un seul cas où une séquence L1 soit répétée dans un même dictionnaire et traduite de façon différente : MFE *en fer forgé, de hierro forjado* s.v. **fer** n.m. vs MFE *en fer forgé, en hierro forjado* s.v. **forger** v. tr.

s.v. **fond** *n.m.* et s.v. **forme** *n.f.*; MFE *faute de frappe, error de máquina* s.v. **faute** *n.f.* et s.v. **frappe** *n.f.*; MFE *feu follet, fuego fatuo* s.v. **feu** *n.m.* et s.v. **follet** *adj.*

Ces cas illustrent un phénomène de caractère plus général motivé par la relation bijective qui s'établit, dans un dictionnaire-papier, entre exemple et lemme. Soulignons que la rigidité de cette relation découle des caractéristiques physiques du support papier et non tant de la nature de la catégorie d'information lexicographique de l'exemple. En effet, une séquence peut être utile sous plusieurs articles du dictionnaire qui la contient, bien qu'elle n'ait été consignée que sous un lemme donné.

Précisons que le fait de remarquer l'existence de cette entropie lexicographique n'implique pas, de notre part, l'affirmation qu'elle puisse être facilement éliminée, au moins pour ce qui est des dictionnaires-papier, où les mécanismes de récupération de l'information sont limités et les renvois coûteux.

Le deuxième type de reprise auquel nous faisons référence consiste en la répétition, dans un même dictionnaire, de deux séquences L2 identiques associées à des séquences L1 différentes<sup>8</sup>. Nous trouvons des cas comme LFE *la frigidéz del mármol, la frigidité du marbre* vs LFE *la frialdad del mármol, la frigidité du marbre*; BFE *je file à la mairie, voy volando al ayuntamiento* vs BFE *je fonce à la mairie, voy volando al ayuntamiento*; VFE *ficher à la porte, echar a la calle* vs VFE *flanquer à la porte, echar a la calle*; MFE *santé fragile, salud delicada* vs MFE *santé frêle, salud delicada*.

Bien que ce procédé semble parfois bien justifié (e.g. MFE *faisceau code, luz de cruce*; MFE *feu de croisement, luz de cruce*), il donne bien souvent l'impression d'un certain appauvrissement par rapport à L1.

#### 4 En guise de conclusion

Il existe, entre les différents dictionnaires bilingues français > espagnol, espagnol > français, des courants de filiation d'exemples plus ou moins importants d'après les cas. Nous avons voulu montrer que ce phénomène, dans la forme qu'il revêt actuellement, nuit autant à la pertinence qu'à la cohérence de l'emploi de l'exemple dans la prédication lexicographique.

Parallèlement, de nombreuses séquences qui pourraient être utiles à la description de plus d'un lemme restent inaccessibles à l'utilisateur du fait de la rigidité du lien entre le mot-vedette et les exemples qui lui sont associés.

8 Nous ne considérons pas ici les cas d'homonymie accidentelle comme MEF *finca rústica, ferme* vs MEF *en firme (com), ferme*.

La complexité de l'exemple comme catégorie d'information lexicographique, ainsi que la relation étroite qu'il garde avec la microstructure concrète pour laquelle il a été créé ou sélectionné, font que sa réutilisation dans d'autres articles s'avère délicate. La réflexion sur les conditions de réemploi de ces séquences, outre son intérêt pratique (exploitation optimale des données linguistiques déjà codifiées) nous semble donc être d'un intérêt théorique non négligeable, puisqu'elle débouche, d'une part, sur l'étude des possibilités et des limites de la réversibilité du dictionnaire bilingue et, d'autre part, sur l'analyse de l'interdépendance entre les différents éléments de la microstructure.

## Chapitre 6

# PROBLÈMES D'ÉQUIVALENCE DANS LES DICTIONNAIRES BILINGUES

Thomas SZENDE

*«Übersetzte Wörter lügen immer, übersetzte  
Texte nur, wenn sie schlecht übersetzt sind.»*

H. Weinrich, *Linguistik der Lüge*, Heidelberg,  
L. Schneider, 1966, p. 24.

### 1 L'anisomorphisme des unités lexicales

Certains concepts trouvent une expression lexicale dans toutes les langues. Il existe en effet des domaines où c'est notre univers commun qui impose le découpage lexical : tous les êtres humains habitent la même planète, se trouvent confrontés à la réalité de la terre et du ciel, du froid et du chaud, aux notions d'espace et de temps, etc. Mais les zones de réalités les plus diverses peuvent se découper différemment selon les langues, a fortiori d'une famille de langues à l'autre.<sup>1</sup>

La réalité extralinguistique ne connaît que des gradations imperceptibles. C'est la langue qui crée des oppositions au sein d'une réalité sans limites précises. L'oeil distingue, dans l'échelle des couleurs, des centaines de variantes, mais les langues naturelles ne se servent que d'un nombre limité de dénominations fondamentales. Loin d'être des nomenclatures distinctes, les langues sont plutôt

---

1 L'équipe que je dirige, au sein du Centre Interuniversitaire d'Etudes Hongroises (Université de la Sorbonne Nouvelle - Paris 3), est chargée d'élaborer un nouveau dictionnaire hongrois > français (un dictionnaire français > hongrois étant réalisé parallèlement à l'Université de Szeged, en Hongrie, sous la direction du Professeur Miklós Pálffy). Les exemples utilisés dans cet article s'inspirent ainsi d'une comparaison entre le français et le hongrois, langue finno-ougrienne, apparentée à une série de langues parlées en Europe et en Asie dont les plus connues sont le finnois, l'estonien et le lapon.

Je tiens à remercier les membres de l'équipe lexicographique (Joëlle Dufcuilly, Viktória Eröss, Emilie Molnos, Chantal Philippe et Dominique Radanyi) pour l'aide constante qu'elles apportent, grâce à leurs compétences linguistiques, dans l'établissement des équivalences entre les deux langues. J'adresse mes remerciements au Professeur Jean Perrot, directeur du CIEH, initiateur du projet auquel il apporte un soutien précieux.

des réseaux de signification qu'organise de différentes manières le monde expérimenté. La langue n'est pas constatation mais délimitation de frontières, à l'intérieur de l'expérimenté.

Si le passage d'une langue à l'autre est possible c'est probablement parce qu'il existe des traits communs dans le domaine sémantique également. Toutes les langues comportent ainsi les principaux types de discours (narration, démonstration, dialogue) et elles peuvent toutes exprimer les relations logiques (cause/effet, condition/résultat), le temps, l'espace, le mode, etc. De même, elles établissent la distinction entre des catégories sémantiques de base : objets (êtres, choses, entités nommables susceptibles de participer d'une manière ou d'une autre à des événements); événements (actions), etc.

La présence dans toutes langues de mots-tabous, d'expressions entourées d'une auréole d'interdit est aussi probablement un trait sémantique universel, de même qu'il existe dans toutes les langues des «méthodes» visant à rendre les bruits du monde, ce qui ne veut nullement dire que le même son est perçu de façon identique par toutes les oreilles. Il arrive certes que les onomatopées soient dans deux langues rigoureusement identiques : par exemple, les Hongrois et les Français qui ont froid disent tous *brrr...*; mais pour inciter un cheval à avancer, les Français emploient l'interjection *hue !*, alors que les Hongrois crient *gyí !*; si les vaches françaises font *meuh !*, celles de Hongrie font *muu !*, etc.

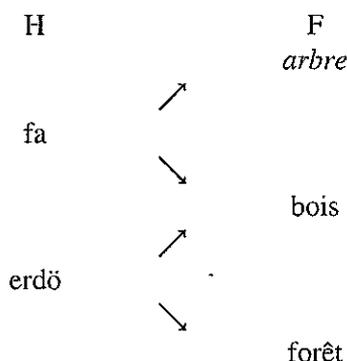
Le fait que l'on trouve dans toutes les langues des mots «opaques» et des mots «transparents», c'est-à-dire des mots pour lesquels le rapport son-sens est purement arbitraire (le terme hongrois *unoka*) ou sensiblement figuratif (ses équivalents français : *petit-fils* ou *petite-fille*) peut être également perçu comme un universel sémantique.

La création linguistique dans une langue peut correspondre, mais sans nulle obligation, à des délimitations qui paraissent «objectives» dans un autre environnement linguistique. Le lexique d'une langue peut posséder deux ou plusieurs mots pour exprimer le même concept. On a ainsi en français *alpin* et *alpestre* (*plantes alpines* ou *alpestres*, mais *paysage alpestre*) contrairement au hongrois qui fera appel dans tous les cas à l'adjectif *alpesi*. En revanche, le substantif hongrois peut exprimer la diathèse, neutralisée dans le substantif français : *felszabadítás* (ACTION DE LIBÉRER) / *felszabadulás* (FAIT D'ÊTRE LIBÉRÉ) = *libération*. En face de termes français désignant à la fois un objet et une action, le hongrois peut séparer les deux sens et les désigner par des termes différents : *zuhany* = *douche* (DOUCHE INSTRUMENT) et *zuhanyozás* = *douche* (FAIT DE PRENDRE UNE DOUCHE).

Une langue peut habituellement user d'un mot unique là où une autre use d'une périphrase. Il y a ainsi des nuances hongroises que le français ne peut rendre que d'une manière périphrastique : *magyarázkodás* = «action de faire l'entêté», *mérgezettség* = «le fait d'être empoisonné», etc. Contrairement au hongrois, le français ne fait pas de distinction entre le concept ÉCLAIR et le fait de considérer

le moment de son apparition : *villám/villámlás*.

Y a-t-il une raison objective pour que se différencient dans beaucoup de langues les concepts ARBRE et BOIS ou BOIS et FORÊT ? Le hongrois, de son côté, fonctionne sans faire ces distinctions :



Tout lexique comporte des incohérences, des redondances et des vides. Mais bien qu'incomplets et dissymétriques, les systèmes lexicaux coïncident avec les besoins majeurs de leurs locuteurs. Le fait que telle construction conceptuelle porte un nom, qu'elle est en quelque sorte solidifiée par un nom, prouve l'intérêt que lui porte la communauté linguistique, alors que telle autre ne peut s'exprimer que de façon analytique et indirecte. Si on accepte de concevoir le lexique comme un filet, les mailles en sont serrées sur certains points, et sur d'autres extrêmement lâches. En français, il n'existe aucun terme spécifique pour désigner le JEUNE MULET, comme le jeune âne est appelé *ânon* ou le jeune cheval *poulain*. Toute langue souffre de lacunes dans son vocabulaire et dans une perspective contrastive il y a lacune chaque fois qu'un signe de la langue de départ ne trouve pas d'équivalent dans la langue d'arrivée.

Si chaque langue a ses «trous» c'est que dans la civilisation d'en face il y a des particularités dont elle tient moins compte, ou pas du tout. L'expérience de l'«insuffisance» du vocabulaire est intimement liée à la pratique quotidienne de la lexicographie bilingue ! Ainsi les trois concepts ci-dessous manquent en français alors qu'ils sont lexicalisés en hongrois :

H		F
<i>láb</i>	∅	(= <i>pied + jambe</i> )
<i>lábbeli</i>	∅	(= <i>chaussures + bottes + pantoufles</i> , etc.: tout se qui se met aux pieds)
<i>ágynem</i>	∅	(= <i>draps + couvertures</i> )

L'abondance lexicale dans une langue peut aussi causer bien des embarras pour l'utilisateur non natif. Faisant allusion à l'existence en hongrois de deux mots désignant HISTOIRE (*történet/történelem*), A. Sauvageot surprend en déclarant qu'«il ne sert à rien de vouloir à toute force distinguer l'histoire conçue

comme la succession des événements et l'histoire conçue comme la description des événements successifs», d'autant plus, continue-t-il, qu'aucune distinction correspondante n'existe ni en allemand, ni en français, ni dans aucune des langues de civilisation avec lesquelles les Hongrois sont en relation. Sa conclusion : «il y aura une difficulté de plus à se servir de ces mots, et cela sans bénéfice aucun pour la précision de la pensée ou pour la clarté de la communication»<sup>2</sup>.

La capacité d'une communauté linguistique à distinguer (ou à ne pas distinguer) tels ou tels concepts peut effectivement devenir source de multiples confusions pour l'utilisateur étranger. Mais les mots et formules d'acception «équivalente» enrichissent nécessairement la langue et dotent la communication de multiples nuances stylistiques. Les noms de parenté constituent un domaine où la structuration du lexique est facile à percevoir; ils offrent ainsi des matériaux particulièrement instructifs pour la comparaison des langues au point de vue de ce que L. Hjelmslev appelle «commutation» et «substitution» : deux membres d'une série lexicale sont «commutables» si le remplacement de l'un par l'autre peut entraîner un remplacement analogue dans l'expression.

Les quatre grandeurs sémantiques FRÈRE AÎNÉ, FRÈRE CADET, SŒUR AÎNÉE, SŒUR CADETTE sont toutes mutuellement commutables en hongrois, tandis qu'en français comme dans la plupart des langues européennes il y a substitution entre «aîné» et «cadet» :

GRANDEUR SEMANTIQUE	H		F
FRÈRE AÎNÉ	<i>báty</i>	\	<i>frère</i>
FRÈRE CADET	<i>öcs</i>	/	
SŒUR AÎNÉE	<i>nővér</i>	\	<i>soeur</i>
SŒUR CADETTE	<i>húg</i>	/	

Quant au terme hongrois *testvér* (littéralement : «corps-sang»), il n'a aucun équivalent en français puisqu'il s'applique indistinctement à un *frère* ou à une *sœur* :

GRANDEUR SEMANTIQUE	H	F
«avoir les mêmes parents»	<i>testvér</i>	Ø

Dans le sens inverse, la traduction de *il* ou *elle* se heurte à une lacune en hongrois qui ne connaît pas la distinction des genres. Il y a donc commutation dans le pronom personnel du français (*il* : *elle*), mais substitution dans celui du hongrois. En hongrois, le remplacement de l'une de ces grandeurs sémantiques par l'autre dans le pronom personnel ne peut pas entraîner un remplacement analogue dans l'expression. La distinction de sexe trouve toujours son expression

2 Sauvageot (1964 :210-211)

dans toute langue, dès que cette distinction devient utile. Si le hongrois ne possède pas de marque du genre grammatical, il dispose cependant de plusieurs procédés pour indiquer le sexe. Dans les deux langues, les noms peuvent exprimer la distinction du genre. Mais l'opposition de genre est loin de s'appliquer à tous les noms. Elle revêt un aspect formel dans les exemples ci-dessous aussi bien en hongrois qu'en français :

H	F
<i>színész / színésznő</i>	<i>acteur / actrice</i>
<i>eladó / eladónő</i>	<i>vendeur / vendeuse</i>
<i>tanító / tanítónő</i>	<i>instituteur / institutrice</i>

Dans d'autres cas, il y a asymétrie entre les deux langues :

<i>tanár / tanárnő</i>	<i>professeur / Ø</i>
<i>bíró / bírónő</i>	<i>juge / Ø</i>
<i>igazgató / igazgatónő</i>	<i>directeur / directrice</i> (ce dernier terme ayant un usage plus restreint que <i>igazgatónő</i> en hongrois), etc.

Un des domaines où la distinction du genre grammatical devrait théoriquement refléter le plus fidèlement l'opposition de sexe est celui des noms d'animaux. Or, on est frappé de constater que l'opposition morphologique n'est ni générale ni symétrique. Dans certains cas les deux langues peuvent recourir à des expédients :

H	F
<i>hím / nőstény elefánt</i>	<i>un éléphant mâle / un éléphant femelle</i>
<i>hím / nőstény szúnyog</i>	<i>un moustique mâle / un moustique femelle, etc.</i>

Dans d'autres cas, seul le hongrois recourt à des expédients :

H	F
<i>hím / nőstény oroszlán</i>	<i>lion / lionne</i>
<i>hím / nőstény tigris</i>	<i>tigre / tigresse, etc.</i>

Un dictionnaire bilingue doit comporter des renseignements grammaticaux cohérents sur l'expression du sexe, catégorie sémantique incontournable, y compris les contraintes morphologiques et syntaxiques correspondantes (voir Lemmens, dans ce volume).

Il existe bien peu de signes équivalents d'une langue à l'autre, même entre deux langues «proches» comme le sont le français et l'anglais. En dehors de quelques termes techniques ou scientifiques à l'usage strictement circonscrit à un domaine (par exemple : *bioszféra*, *bojár*, *kultusz* ont respectivement pour équivalents français *biosphère*, *boyard*, *culte*). La grande majorité des signes de deux langues ne sont pas équivalents; ils peuvent désigner des réalités multiples et souvent fort différentes.

On constate entre les exemples cités ci-dessus une analogie formelle. Mais il ne faut jamais se fier aux similitudes entre les langues. A côté de quelques vrais amis (ressemblance formelle + identité de sens), il existe un nombre important de faux amis entre le hongrois et le français, langues pourtant fort éloignées l'une de l'autre sur le plan typologique. Nous pensons notamment aux mots étrangers empruntés massivement par le hongrois qui malgré la ressemblance formelle ont subi généralement un certain glissement de sens par rapport au modèle de départ :

H	F
<i>szolid</i>	<i>sérieux ; bien élevé</i>
<i>figura</i>	<i>curieux personnage</i>
<i>artista</i>	<i>artiste de cirque</i>
<i>laikus</i>	<i>profane ; incompétent</i>
<i>bizsu</i>	<i>bijou (de) fantaisie/en toc ; etc.</i>

Dans chaque langue, la convention a imposé aux mots des acceptions résultant de montages conceptuels qui s'édifient spontanément, au fur et à mesure que tel mot trouve tel emploi dans telles circonstances.

## 2 Signes arbitraires et signes motivés

Le dictionnaire bilingue confronte deux systèmes linguistiques et notamment deux systèmes lexicaux. Il est notoire que le hongrois, aussi bien dans son lexique que dans sa structure, est très différent du français. Dans le domaine sémantique, on a le sentiment que très souvent plusieurs mots hongrois correspondent à un mot français, mais l'extension des mots français est plus vaste que celle des mots hongrois.

Pour les substantifs concrets, où les choses et les êtres sont nommés dans leur singularité, les égalités devraient être fréquentes. En effet à *kutya* correspond *chien* et au hongrois *kanál* correspond le français *cuillère*. Mais voici qu'au terme *musique* correspondent à la fois *zene* et *muzsika*, tandis que *dal* en hongrois signifie à la fois *chant* et *chanson*. Le hongrois possède deux mots pour dire *père* : *atya* et *apa*, le premier étant plus noble, plus respectueux. Cette fantaisie des aires sémantiques ne doit pas nous surprendre, même dans le domaine du concret. Et lorsqu'on passe du sens propre au sens figuré, les disparités deviennent encore plus flagrantes.

Le hongrois paraît plus près du réel, développant dans son lexique de fortes structures plus motivées. En effet, certains linguistes tentent de distinguer entre vocabulaires arbitraires et vocabulaires motivés, les premiers ayant une prédilection pour les mots ne possédant aucun lien – naturel ou logique – avec la « chose » ou le concept qu'ils désignent. Par rapport à *repülögép* (« machine à voler »), *avion* serait arbitraire. Font partie des termes « arbitraires » ceux dont la motivation s'est obscurcie : le francophone non linguiste ne pourra pas identifier les éléments qui

composent des mots comme *banqueroute*, *s'arc-bouter* ou *calfeutrer*. En français, dans la création lexicale on a souvent recours à des racines gréco-latines : ce qui est littéralement «médecin des yeux» (= *szemorvos*) en hongrois, deviendra *ophthalmologiste* en français.

La part de la motivation est plus forte en hongrois, grâce aux divers procédés de dérivation et de composition, contrastant avec le caractère immotivé du lexique français. Par la richesse de son vocabulaire en termes (notamment des verbes) de sens concrets, le hongrois est souvent à son aise pour trouver une construction là où le français est plutôt embarrassé en raison de son vocabulaire plutôt abstrait.

La langue hongroise peut s'enorgueillir d'un système dérivationnel extrêmement original. Si l'on ne considère que les catégories grammaticales de base, on aboutit aux types de dérivation suivants (V = verbe; V' = suffixe formateur de verbe; N = substantif; N' = suffixe formateur de substantif; A = adjectif; A' = suffixe formateur d'adjectif; D = adverbe; D' = suffixe formateur d'adverbe) :

V	<i>üt</i>	<i>frapper</i>	N	<i>ágy</i>	<i>lit</i>
VV'	<i>ütöget</i>	<i>frapper à plusieurs reprises</i>	NV'	<i>ágyaz</i>	<i>faire le lit</i>
VN'	<i>ütés</i>	<i>coup (de poing)</i>	NN'	<i>ágyacska</i>	<i>petit lit</i>
VA'	<i>üt</i>	<i>celui qui frappe</i>	NA'	<i>ágyas</i>	<i>pourvu d'un lit</i>
A	<i>fekete</i>	<i>noir</i>	D	<i>távol</i>	<i>loin</i>
AV'	<i>feketít</i>	<i>noircir</i>	DV'	<i>távolít</i>	<i>éloigner</i>
AN'	<i>feketeség</i>	<i>noirceur</i>	DN'	<i>távolság</i>	<i>distance</i>
AA'	<i>feketés</i>	<i>noirâtre</i>	DA'	<i>távoli</i>	<i>lointain</i>

Par sa morphologie, le hongrois assemble partout où il le peut; le français en revanche aime à séparer. Le hongrois exprime souvent en un seul mot composé les rapports permanents établis entre deux objets ou deux notions. En face de ces mots composés, le français présente souvent des éléments liés par une copule (à ou de) :

H		F
	<i>kézitáska</i> (= «main-sac»)	<i>sac à main</i>
	<i>hálószoba</i> (= dormant-pièce»)	<i>chambre à coucher</i>
	<i>vízesés</i> (= «eau-chute»)	<i>chute d'eau, etc.</i>

Il arrive évidemment qu'en face d'un mot composé hongrois il y ait un mot unique en français :

H		F
	<i>jegygyűrű</i> (= «fiançailles-bague»)	<i>alliance</i>
	<i>összeadás</i> (= «ensemble-action de donner»)	<i>addition</i>
	<i>hashajtó</i> (= «ventre-relâchant»)	<i>purge, etc.</i>

Mais dans ces cas-là, en face du mot hongrois qui, grâce à sa composition même, est un signe motivé, on trouve en français un signe plus arbitraire dont la

motivation sous-jacente ne sera retrouvée que lorsque le lien avec d'autres lexèmes sera perçu (*alliance* avec *allier*, par exemple).

Le procédé de composition lexicale qui consiste à juxtaposer deux mots sans marque, c'est-à-dire sans indication du rapport qui existe entre eux, est très fréquent en hongrois. Les adjectifs hongrois formés ainsi auront souvent une extension moins grande que les adjectifs français équivalents. En revanche, ils auront tendance à tirer leur précision de leur forme même :

H	F
<i>életvidám</i> (= «vie-gai»)	<i>gai</i>
<i>vérszegény</i> (= «sang-pauvre»)	<i>anémique</i>
<i>szűkszávú</i> (= «étroit-parole»)	<i>taciturne, etc.</i>

Le mot hongrois semble s'insérer davantage dans la réalité concrète, contrairement au mot français qui traduit une représentation plus abstraite : le signe tend à se substituer à l'image. Le signe arbitraire supporte plus de significations que le signe motivé accroché à une partie précise du réel. Il existe de nombreux verbes hongrois qui conservent l'image, le caractère de l'action exprimée et qui n'ont en face d'eux qu'un verbe français «signe» dans lequel ne revit que d'une façon lointaine l'image originale. Il est tentant d'établir des séries de verbes concrets du hongrois en face des verbes français qui n'ont leur valeur précise que par leur contexte. Pour traduire en hongrois le verbe français *entrer*, il faut se demander qui entre, comment et pourquoi, bref, quelle est la nature de l'entrée. Pauvreté, richesse ou miracle, un seul petit verbe «signe» du français, le verbe *passer* peut correspondre à une très longue série de verbes hongrois; l'énoncé nuance ou change son sens selon les mots auxquels le verbe *passer* est associé :

H	F
quelqu'un <i>passa</i> devant qch	valaki <i>elhalad</i> valami előtt
quelque chose ne <i>passera</i> pas	valami nem fog <i>felülkerekedni</i>
<i>passer</i> une ligne	<i>kihagy</i> egy sort
<i>passer</i> chez qqn	<i>benéz</i> valakihez
la pièce <i>passera</i> bientôt	a darab hamarosan <i>színre kerül</i>
toute sa fortune y <i>passera</i>	<i>rámegy</i> az egész vagyona, etc.

### 3 Le traitement lexicographique des différents types d'unités lexicales

Le dictionnaire bilingue a pour objectifs de rendre compte :

- des sens et des emplois des mots de la langue de départ et
- des traductions possibles pour chaque sens et emploi dans la langue d'arrivée.

L'article du dictionnaire bilingue présente les sens et les emplois de la vedette selon une certaine hiérarchisation de l'information. Il est construit, ainsi, selon

une structure clairement reflétée dans la délimitation typographique des subdivisions.

– subdivisions grammaticales

H	F
<i>abszolút</i>	I adj <i>absolu</i> II adv <i>absolument</i>
<i>akadémikus</i>	I adj <i>académique</i> II n <i>académicien</i>

– subdivisions de sens

H	F
<i>ablakmosó</i>	1 ( <i>személy</i> ) <i>laveur de carreaux</i> 2 ( <i>eszköz</i> ) <i>produit pour les vitres</i>
<i>dirigál</i>	1 <i>péj donner des ordres</i> 2 <i>Mus diriger</i>

L'article est construit de manière à orienter l'utilisateur vers le choix de la traduction exacte dans le contexte requis, d'où l'importance de l'appareil d'indicateurs grammaticaux, sémantiques, stylistiques, etc. permettant de choisir l'équivalent le plus juste possible (voir ci-dessus l'article *ablakmosó* comportant les indications *személy* = «personne» et *eszköz* = «outil» ou l'article *dirigál* comportant les abréviations *péj* = «péjoratif» et *Mus* = «musique»).

Dans l'esprit de l'utilisateur, la traduction est toujours possible et une équivalence existe nécessairement. Or, les problèmes d'équivalence se posent sur deux plans : le plan du réel et le plan de la langue. Le réel existe-t-il ou non dans la culture des locuteurs ? Le mot qui le désigne existe-t-il ou non dans la langue des locuteurs ?

Si les signifiés de la langue de départ et de la langue d'arrivée renvoyaient toujours à une même réalité culturelle et que les signifiants étaient représentés par un seul élément du lexique dans les deux langues, les dictionnaires bilingues seraient infiniment plus simples. Mais le cas le plus courant n'est pas celui-là. La pratique lexicographique quotidienne montre qu'il n'existe généralement pas de concordance directe entre les unités linguistiques de deux systèmes différents<sup>3</sup>.

Lorsque le rédacteur ne dispose d'aucun moyen pour fournir une équivalence satisfaisante, il proposera une glose contextuelle. C'est ce qui se passe lorsque le réel n'existe que dans l'univers culturel et le lexique du locuteur de la langue de départ :

3 «The task of the bilingual dictionary is to provide words and expressions in the source language with counterparts in the target language which are as near as possible, semantically and as regards style level (register). But a complete equivalence of this kind between words and expressions in two languages is quite unusual, in the same way as it is unusual to find exact synonyms within one language.» Svensén (1993:140)

H	F
<i>ballagás</i>	= «monôme et chants par lesquels les élèves de terminale prennent congé de leur école»
<i>abszolutórium</i>	= «certificat de fin d'études universitaires»
<i>akó</i>	= «ancienne mesure de capacité pour les liquides et les grains, égale à environ 55 litres»

Il arrive que le mot existe bien en tant qu'élément du lexique dans les deux langues, mais qu'il n'occupe pas tout à fait la même place dans l'univers culturel des locuteurs des deux communautés. Le mot *vache* a sans doute le même signifié en Inde qu'en France (femelle du taureau), mais il n'a pas la même charge culturelle : la vache est protégée, parce que sacrée en Inde, exploitée, parce que nourricière en France. Aussi, les dictionnaires bilingues doivent-ils s'employer à combler le handicap des locuteurs non natifs face aux diverses connotations partagées par les natifs. Le Normand et le refus de trancher (réponse «normande») sont étroitement liés dans l'esprit des Français, de même que les adjectifs *nyakas* (= *têtu*) et *kálvinista* (= *calviniste*) dans l'esprit des Hongrois. Chaque fois qu'il y a une connotation particulière, le dictionnaire doit alerter le traducteur, afin que celui-ci évite dans la langue d'arrivée les mots susceptibles d'éveiller des associations peu désirables. Mais, avant de les signaler, le lexicographe doit s'assurer qu'il s'agit bien de connotations collectives et non pas de connotations individuelles.

A noter à ce propos que l'auteur du dictionnaire bilingue s'efface au maximum derrière les équivalents qu'il fournit. Son rôle est de recenser des faits de langue. Néanmoins, le libre arbitre du rédacteur ne peut jamais être exclu. Il fait des choix afin de ne garder que les sens les plus productifs, les contextes les plus pertinents, les traductions les plus naturelles, etc. Il peut arriver, notamment dans le cas des expressions ligées et métaphoriques, que le rédacteur soit amené à créer des termes qui lui sont personnels, et à tenter de les imposer à l'usager.

La richesse d'un dictionnaire bilingue ne se mesure pas forcément au nombre d'entrées ni non plus au nombre d'équivalents. Certes, un dictionnaire doit aider le traducteur à diversifier son style, sans pour autant se laisser entraîner dans une synonymie cumulative. Si la richesse est loin d'être un leurre, du moins peut-elle susciter une certaine confusion. Si on propose, dans les dictionnaires bilingues, pour une même unité lexicale ou un même cas d'emploi, plusieurs équivalents, c'est pour élargir les possibilités de choix du traducteur, mais aussi pour stimuler le jeu associatif dans la recherche de variantes personnelles. Ces équivalents peuvent parfois être plus riches, plus idiomatiques que l'unité lexicale originale, ce qui permet de jouer au maximum des ressources de la langue d'arrivée et de mettre en pratique le principe de compensation.

Lorsqu'il n'y a pas de rapports synonymiques entre les équivalents possibles, on doit procéder à une segmentation très nette à l'intérieur de l'article.

H	F
<i>alak</i>	1. <i>forme ; aspect</i> 2. <i>taille ; silhouette</i> 3. <i>individu ; type</i>

En revanche, il est indispensable de tenir compte des éventuelles affinités synonymiques entre les équivalents d'un mot dans une autre langue. Quels sont les rapports entre les équivalents enregistrés les uns après les autres (séparés par une virgule, un point-virgule ou par une barre oblique, suivant les dispositions du protocole de rédaction) au sein des articles de dictionnaires bilingues ? Au moins deux possibilités théoriques s'offrent en ce qui concerne leur emploi :

- ils peuvent être utilisés indifféremment;
- ils ne sont interchangeables que dans un certain nombre de constructions.

En effet, deux mots de forme totalement différente peuvent avoir, pour une partie de leurs emplois, une acception à peu près identique, alors que dans tous les autres cas ils ont habituellement des sens divergents.

On sait que la synonymie n'est pas perpétuelle, que deux mots ne sont pas synonymes en toute occasion. Ainsi, dans un dictionnaire bilingue, les équivalents proposés parallèlement ne sont interchangeables qu'à titre exceptionnel. Une certaine synonymie ou interchangeabilité fonctionnelle s'installe néanmoins lorsque

- deux ou plusieurs mots sont de valeur approximativement identique dans une situation précise et peuvent donc être employés dans des fonctions assimilables;
- la substitution entre eux ne risque de causer aucun trouble dans la communication quotidienne.

La question principale que se pose le lexicographe sélectionnant les équivalents est la suivante : y a-t-il des situations où deux mots peuvent se substituer l'un à l'autre. Le traitement du mot hongrois *év* (= *an*; *année*) doit mettre en évidence les principales règles qui se rapportent à l'usage des équivalents respectifs :

H	F
<i>év</i>	<i>an</i>
<i>húsz éves</i>	<i>il a trente ans</i>
<i>év</i>	<i>année</i>
<i>több évig</i>	<i>(pendant) plusieurs années</i>
<i>év</i>	<i>an/année</i>
<i>tavaly</i>	<i>l'an dernier ou l'année dernière</i>

Les équivalents *an* et *année* n'interfèrent pas complètement, chacun des deux termes possédant son aire particulière d'emploi. C'est leur somme qui est appelée à rendre pleinement le sens de la vedette hongroise.

En français, l'existence de nombreux doublets est liée à la coexistence de deux sortes de traditions lexicales : celle des mots hérités (qui ont subi les changements phonétiques par lesquels les parlers latins sont devenus du français) et celle des mots empruntés (qui ont fait l'objet d'adaptations plus ou moins réduites). Le hongrois aussi fait coexister pour de nombreux concepts un terme autochtone (dont seule la recherche étymologique peut déceler l'origine) et un terme étranger, ce qui ne veut nullement dire qu'il s'agit de synonymes parfaits. Ainsi, en ce qui concerne les mots désignant des objets ou des institutions de toute évidence importés, c'est le terme étranger qui est le plus courant; le terme hongrois, réservé au langage officiel, est alors senti comme guindé, artificiel et est réservé à l'administration :

<i>bicikli / kerékpár</i>	= «roue paire»
<i>telefon / távbeszélő</i>	= «parlant à distance»
<i>távíró / telex</i>	= «écrivain à distance»

Pour ces mots, la différenciation est relativement simple; dans d'autres cas, elle implique des considérations plus complexes. Chaque langue peut éprouver le besoin de s'appuyer sur des supports spécifiques pour distinguer les différents aspects – aussi ténus fussent-ils – d'un même concept. Les stratégies adoptées en hongrois et en français pour s'adresser à un interlocuteur ou le saluer sont très différentes. Le locuteur hongrois dispose, dans ce domaine ritualiste de la communication quotidienne, d'un choix d'expressions étonnamment vaste et diversifié dont certaines trouvent difficilement une expression lexicale équivalente en français. La formule *kezét csókolom* (= *je vous baise la main*), pratiquement obligatoire pour tout homme s'adressant à une femme, peut s'abréger en *csókolom* dans le langage familier quand c'est un homme «supérieur» qui s'adresse à une femme «inférieure»; en *csókollak* (= *je t'embrasse*) quand c'est un homme «inférieur» qui s'adresse à une femme «inférieure» et, chose encore plus surprenante, la formule *kezét csókolom* est parfois utilisée par l'homme «inférieur» s'adressant à un homme «supérieur».

Par ailleurs, le terme *néni* (= *tante*) s'applique à toutes les femmes d'un certain âge, de même que *bácsi* (= *oncle*) s'applique à tous les hommes respectables. De plus, pour les enfants hongrois, tous les adultes sont des *bácsi* ou des *néni*. A noter également que ces termes, couplés le plus souvent avec un nom patronymique ou avec un prénom, peuvent s'employer indifféremment dans un rapport de tutoiement ou de vouvoiement.

Les langues comprennent aussi des énoncés qui, à défaut d'être traduits, doivent être définis à l'aide de paramètres pragmatiques (rapport d'égalité, de supériorité, d'infériorité; sexe, âge, appartenance socio-professionnelle des locuteurs; degré d'intimité; etc.). Les énoncés déclenchés par une situation analogue ne sont pas toujours transposables (traduisibles) directement. Les conversations téléphoniques, par exemple, sont particulièrement riches en clichés. Associés à telle ou telle phase du contact téléphonique, les énoncés ci-dessous échappent aux règles de la traduction «directe» :

H

*ne quittez pas*  
*occupé*  
*de la part de qui?*

F

*tessék maradni* (= «veuillez rester»)  
*mással beszél* (= «il/elle parle à qqn d'autre»)  
*ki keresi?* (= «qui le/la cherche?»)

Il s'agit de formules, imposées par l'usage et, par conséquent, peu prévisibles si l'on se fonde uniquement sur des connaissances linguistiques ou encyclopédiques.

Une langue ne s'apprend pas seulement par ses règles et par ses mots, mais par tous les éléments «pré-codés» du discours : locutions, termes complexes et figés, «manières de dire» liées aux stratégies de la parole, éléments qui pour la plupart sont maîtrisés spontanément par le locuteur natif, alors que le locuteur étranger doit les apprendre comme si c'étaient des mots.

L'association de plusieurs unités significatives est un phénomène commun à toutes les langues, indispensable au fonctionnement même du langage. En règle générale, c'est la cohésion des éléments constitutifs et le degré d'autonomie que ceux-ci conservent dans la nouvelle unité qui servent de critère de délimitation et de classement (voir Roberts, dans ce volume). Les difficultés de l'apprentissage et de la traduction révèlent l'importance de ces groupes plus ou moins figés. En effet, pour maîtriser une langue, il est indispensable d'acquérir les règles de «montage» ou de «combinaison» qui permettent de produire ces énoncés qui, pour les locuteurs natifs, sont déclenchés automatiquement par la situation. Grâce à une description plus systématique de ces règles, on laisse moins de place à l'intuition du locuteur étranger.

Paradoxalement, ce qui rend insatisfaisants et inadéquats les dictionnaires bilingues traditionnels c'est qu'ils sont attachés surtout à réunir des équivalents, c'est-à-dire à traduire des mots par d'autres, alors que les mots apparaissent généralement parmi d'autres mots; pris isolément ils peuvent n'avoir que des virtualités de signification. Sans contexte, il est généralement difficile ou carrément impossible de les interpréter avec une certitude absolue. La signification pertinente des mots est engendrée généralement par leur interdépendance contextuelle.

Dans un dictionnaire bilingue, le mot ne peut pas constituer l'unité de base universelle dans l'établissement des équivalents, surtout si l'on se rappelle que certaines catégories grammaticales sont présentes à travers des formes lexicalisées dans une langue et absentes dans une autre. Ainsi, le pronom adverbial *en* du français (*il y en a*) peut correspondre en hongrois à une formulation extrêmement économique qui est celle de l'équivalent zéro : *van* (= *il y a*). De même que l'équivalent grammatical de l'article partitif français (*c'est du vin*) est l'article zéro hongrois : *ez* (= *ceci*) *bor* (= *vin*).

Il y a dans chaque langue grosso modo deux sortes d'éléments lexicaux :

- des «mots à part entière», qui supportent un signifié plus ou moins définissable, même lorsque cette définition ne se dégage que des conditions de leur emploi

(ordre des mots, associations avec d'autres termes, signification globale de la séquence, etc.) et,

- des mots qui n'ont pas de valeur sémantique propre, qui sont pour ainsi dire subordonnés aux autres, servant à signaler les liaisons et qui sont des outils pour la construction des syntagmes ou pour l'expression des catégories grammaticales ou sémantiques.

Il n'est pas toujours facile de tracer une ligne de démarcation entre les deux catégories. Tout mot à signification générale tend à devenir outil, selon la fréquence où il apparaît dans certaines constructions. Ainsi, le participe présent *maintenant*, vidé de sa signification, a donné un adverbe de temps qui entre même dans une locution conjonctive sous la forme de *maintenant que*. Dans beaucoup de cas, l'emploi plein et l'emploi grammatical coexistent. Ainsi, le verbe *aller*, vidé de son propre sens, est devenu en combinaison avec un infinitif, une sorte d'indice de futur (*il va manger*), mais dans d'autres locutions et selon les circonstances extra-linguistiques, il continue à être utilisé avec son sens premier (*il va – «de ce pas» – manger*).

Il y a des mots dans le lexique de chaque langue qui ne semblent pas avoir de frontières sémantiques bien délimitées. Hélas, même les meilleurs dictionnaires bilingues ne donnent qu'un faible aperçu des multiples effets de sens de ces mots «fourre-tout». La difficulté de leur traduction réside dans le fait qu'ils expriment des rapports trop divers. Voici à titre d'exemple un extrait de l'article consacré à *csak* dans le nouveau dictionnaire hongrois > français en cours d'élaboration :

**csak**

**I**adv **1** (*restriction*) ~ **egy hétre jön** il vient seulement pour une semaine; **már** ~ **néhány perc, és éjjel lesz** plus que quelques minutes, et il fera nuit; ~ **két ablak tört be** il n'y a que deux fenêtres cassées; ~ **ekkorára nőtt** il n'est pas plus grand que ça **2** (*renforcement*) ~ **egy kicsit maradj még !** reste donc encore un petit peu !; ~ **nem akarsz elmenni ?** ne me dis pas que tu veux partir ?; ~ **semmi izgalom !** allons, du calme ! /pas d'effolement !; **amíg ~ él, nem felejtí** tant qu'il vivra, il ne l'oubliera pas; ~ **azért, mert** pour la seule raison que; ~ **úgy kérdeztem** j'ai juste demandé comme ça (histoire de poser la question); ~ **elintézem, és máris itt vagyok** je termine ça et j'arrive tout de suite; **jó, ~ bajod ne legyen !** c'est bon, mais surtout qu'il ne t'arrive rien ! **4** (*comparaison*) **míntha ~ látnám** c'est (tout) comme si je le voyais **5** (*ironie*) **te már ~ tudod** et toi, bien entendu, tu sais cela (*alors qu'en fait, il n'en sait rien*) **6** (*étonnement*) **egyszer ~ nyílt az ajtó** et voilà que la porte s'ouvre **7** (*espoir*) **majd ~ lesz valahogy** il en sortira bien quelque chose **8** (*encouragement*) ~ **rajta ! (ön/ti) allez-y !; (te) vas-y !; (mi) allons-y !** **9** (*consentement*) (**hát**) ~ **csinálja** eh bien, d'accord, qu'il le fasse **10** (*défi*) **no várj** ~ ! attends voir un peu ! **11** (*continuité*) **egész nap ~ ül** il ne fait rien d'autre que rester assis toute la journée; ~ **ment, ment** il a marché, marché (*contes populaires*) ...

Il est rare que la traduction puisse se réduire à un mot-à-mot littéral; de par sa nature le mot déborde sur le contexte syntagmatique, au-delà des deux «blancs» qui le délimitent. Certes, l'adverbe hongrois *csak* peut être traduit par *seulement*

(c'est d'ailleurs son équivalent «scolaire»), mais pour aller beaucoup plus loin dans la description de son sémantisme, on ne peut guère éviter de le mettre en situation dans des expressions et locutions.

Autre exemple qui illustre bien l'importance accrue du contexte : la pauvreté du système temporel hongrois. Le hongrois ne possède, en face des dix temps français de l'indicatif, que deux temps, présent et passé. Certain rapports temporels, fondamentaux en français (l'opposition duratif/non-duratif et l'antériorité), sont ainsi exprimés en hongrois par des moyens lexicaux, préverbes et adverbes.

L'analyse du contenu sémantique est liée à l'étude de la distribution du mot dans la phrase (voir aussi Lépinette, dans ce volume). Il suffit de changer le régime du verbe *hallgat* pour distinguer les acceptions suivantes :

H	F
<i>hallgat vkire</i> (= «sur qqn»)	<i>écouter qqn</i>
<i>hallgat vmiröl</i> (= «de qqch»)	<i>taire qqch</i>
<i>adózik vmi után</i> (= «après qqc»)	<i>payer un impôt/une taxe sur qqch</i>
<i>adózik vmivel</i> (= «avec qqc»)	<i>témoigner de qqch</i>

Dans les dictionnaires bilingues, on ne traduit pas des mots, ni des phrases, mais des «contextes suffisants». Ce que nous appelons «contexte suffisant» est une portion de texte de longueur variable fournissant l'information nécessaire pour rechercher une équivalence tout en permettant de réduire à un seul sens la multiplicité des sens virtuels de ce mot :

H	F
<i>áldozat 1</i>	<i>victime</i>
<i>áldozatul esik vminek</i>	<i>être victime de qqch</i>
<i>áldozat 2</i>	<i>sacrifice</i>
<i>áldozatot vállal</i>	<i>accepter un sacrifice</i>
<i>áldozat 3 Rel</i>	<i>sacrifice</i>
<i>áldozatot mutat be</i>	<i>faire un sacrifice</i>

Ceux qui apprennent ou enseignent des langues étrangères savent bien que ce qui est particulièrement problématique, c'est l'acquisition de l'aptitude à fabriquer des énoncés, à produire et à reproduire des structures syntaxiques correctes. Or, les fautes en langues étrangères sont fortement conditionnées non seulement par le lexique mais aussi par le système syntaxique de la langue de départ, qui est souvent «projeté» sur celui de la langue d'arrivée. Ainsi, plus un dictionnaire bilingue propose, autour de ses vedettes, des couples de syntagmes corrects, structurellement symétriques ou asymétriques, plus il sera apprécié par l'utilisateur<sup>4</sup>. Car celui-ci ne cherche pas seulement de simples équivalents dans un

4 «The dictionary should offer not explanatory paraphrases or definitions, but real lexical units of the target language which, when inserted into the context, produce a smooth translation...» Zgusta (1984 :147)

dictionnaire bilingue; il est aussi à la recherche de séquences susceptibles d'être insérées dans son discours et formant des énoncés naturels.

## 4 Conclusions

On tombe à chaque pas, dans les langues étrangères, sur des termes qui correspondent à des concepts qu'on ne possède pas dans son vocabulaire d'origine. Et inversement, on éprouve le même type d'embarras à interpréter la majorité des termes de sa propre langue en langue étrangère. Aussi la notion d'équivalence est-elle un concept clé de la lexicographie bilingue. Le rédacteur d'un dictionnaire bilingue est confronté à la tâche complexe de trouver dans la langue d'arrivée des équivalents révélant les sens et les emplois du terme de la langue de départ en veillant à ce que ces équivalents impliquent les mêmes valeurs connotatives et stylistiques.

La rubrique «équivalent» du dictionnaire bilingue doit permettre un choix conforme aux normes et à l'usage, tout en indiquant les degrés d'équivalence, sans oublier de mettre en garde l'utilisateur contre tout risque de généralisation abusive.

D'une langue à l'autre, la désignation d'une même réalité est généralement obtenue par des cheminements différents. Les mots des diverses langues ne sont pas des étiquettes différentes collées sur les mêmes cases. Il n'existe pas deux langues dont les vocabulaires se recouvrent exactement, mot pour mot, toutes les acceptions d'un mot de la langue de départ correspondant à toutes les acceptions d'un mot de la langue d'arrivée. Il y a dans toutes les langues des éléments qui sont comme le dénominateur commun de notre humaine condition – en cela, toutes les langues sont traduisibles – mais chacune porte également l'empreinte lexicale d'une vision du monde particulière, – et en cela, elles sont intraduisibles. Le rédacteur du dictionnaire, étant le plus souvent incapable de proposer une équivalence absolue, se dote donc d'un ensemble de moyens lui permettant de fournir des équivalences relatives, notamment à l'aide d'exemples authentiques.

Ce qui est souvent vrai pour les mots pris isolément l'est moins lorsque ces mots sont insérés dans un contexte. Tout concepteur de dictionnaire bilingue constatera ainsi que si, en règle générale, beaucoup de mots ne peuvent guère être traduits, les phrases peuvent l'être. D'où la citation mise en exergue de ce texte : «les mots traduits mentent toujours, les textes traduits, seulement s'ils sont mal traduits».

## Chapitre 7

# L'UTILISATION D'UN CORPUS DANS LA REDACTION DU DICTIONNAIRE BILINGUE

Valerie GRUNDY

Au cours des dix dernières années, nous avons pu assister à une véritable révolution en ce qui concerne les processus de rédaction et de fabrication des dictionnaires. Un certain nombre des progrès qui ont été accomplis s'insèrent dans le contexte plus large de l'industrie de l'édition en général, tandis que d'autres sont spécifiques à la lexicographie même. L'informatique a joué un rôle primordial dans cette révolution, tant au niveau de la rédaction même d'ouvrages lexicographiques qu'à celui de la typographie et de la fabrication. La technologie des bases de données existante, ainsi que celle des logiciels de rédaction conçus pour les dictionnaires, ont permis à tous ceux qui travaillent dans le domaine de la lexicographie, que ce soit dans le cadre d'un projet de recherche très spécialisé ou dans celui de la fabrication d'un dictionnaire commercial, un accès plus facile et plus souple au texte du dictionnaire en cours et aux informations qui sont à la base de ce texte, toutes choses qui étaient impensables il y a seulement vingt ans. L'utilisation du système d'étiquetage SGML<sup>1</sup> permet un contrôle systématique de la cohérence de la présentation des faits linguistiques à l'intérieur des articles du dictionnaire, c'est-à-dire de la microstructure. Ce système permet également de gérer les systèmes de renvois et la cohérence de la macrostructure du dictionnaire avec une rapidité et une fiabilité bien supérieures à celles de la méthode traditionnelle. Il est en outre possible d'opérer des changements de contenu ou de typographie de dernière minute pour répondre aux besoins du marché visé.

Le présent chapitre ne concerne qu'un aspect de cette évolution : la possibilité d'utiliser des corpus de textes électroniques dans un domaine bien spécifique, celui de la création de dictionnaires bilingues.<sup>2</sup> Je m'efforcerai donc de montrer comment l'utilisation d'un corpus facilite la tâche du lexicographe, et les

---

1 Standard Generalized Markup Language.

2 La méthode décrite ici a été mise au point au cours de la compilation de l'*Oxford-Hachette French Dictionary* (1994), en collaboration avec mes collègues de travail participant au projet et en particulier Marie-Hélène Corréard et Sue Atkins, que je tiens à remercier ici.

conséquences de cette utilisation sur la qualité du dictionnaire final, telle que celle-ci peut être perçue par l'utilisateur.

## 1 L'utilisation d'un corpus : tradition et innovation

Au sein d'une maison d'édition, ce sont les contraintes commerciales qui déterminent, dans une large mesure, les méthodes de travail utilisées par les lexicographes. C'est l'éditeur et non le lexicographe qui décide quel marché le dictionnaire visera et par conséquent quel seront son envergure, son format, son nombre de pages, et son prix (voir Atkins, 1993:11; Hanks 1993:110). Le contexte de la fabrication du dictionnaire est inévitablement celui de l'investissement financier et du rendement. Cela a pour conséquence que les coûts induits par la composition du corpus, la mise en place des ressources nécessaires pour stocker de larges quantités de texte, le développement et l'entretien de logiciels permettant de stocker ce texte et de l'exploiter doivent être rapportés aux avantages que l'on peut retirer du marché.

Le fait que les lexicographes, les linguistes et les métalexigraphes soient tous d'accord pour faire l'éloge de l'utilisation du corpus dans la rédaction de dictionnaires est en lui-même insuffisant dans le contexte de l'édition commerciale. La plupart des utilisateurs – et donc des acheteurs – de dictionnaires n'ont jamais rencontré le mot «corpus» et il serait illusoire de penser que la présence d'un corpus électronique constitue pour eux un critère de choix quand ils achètent un dictionnaire. Le fait qu'un dictionnaire soit composé à partir d'un corpus électronique ne peut donc pas être considéré *a priori* comme une vertu. Les avantages doivent être perceptibles dans les résultats, et la supériorité du dictionnaire basé sur un corpus doit pouvoir être démontrée clairement à l'utilisateur.

Le corpus électronique à objectif linguistique existe depuis quelque temps déjà et l'histoire de son évolution a été écrite à plusieurs reprises (voir Leech : 1991). L'utilisation d'un corpus dans la rédaction de dictionnaires publiés par des maisons d'édition est, cependant, un phénomène assez récent. Le dictionnaire COBUILD a fait figure de pionnier en la matière, et a montré, au moins aux linguistes et aux lexicographes professionnels, la nécessité de s'appuyer sur des faits de langue attestés pour rédiger des dictionnaires – surtout lorsque ces dictionnaires sont destinés à être utilisés par des locuteurs non natifs, qui ont besoin d'avoir accès à la langue telle qu'elle est réellement utilisée tous les jours. Depuis la création du premier corpus COBUILD (7 millions de mots), l'intérêt pour la création de corpus et leur utilisation en lexicographie s'est accru. La technologie s'améliore constamment pour tenter de satisfaire les demandes des utilisateurs de corpus. Il est de plus en plus facile d'avoir accès à des documents lisibles par informatique, grâce à la révolution qui s'est opérée dans les méthodes

de production de livres, de journaux et de revues au cours des dix à quinze dernières années, ce qui permet d'envisager l'accumulation et le stockage d'énormes quantités de textes. Le progrès le plus important pour les lexicographes réside dans la quantité – et donc dans l'objectivité – des données auxquelles ils peuvent désormais avoir accès et dans la qualité des outils dont ils disposent pour traiter l'information.

Le corpus anglais auquel il est fait référence dans ce chapitre est le *Oxford Pilot Corpus* (maintenant intégré, avec quelques modifications, dans le *British National Corpus*<sup>3</sup>) qui contient à peu près 60 millions de mots au moment où nous écrivons ces lignes. Le corpus de français moderne auquel il sera fait référence a été créé pour le projet de l'*Oxford-Hachette French Dictionary* (désormais OHFD), par Oxford University Press. Il contient environ 10 millions de mots provenant de sources variées (journalisme, articles, rapports, monographies, romans<sup>4</sup>) et se veut représentatif de la prose contemporaine. Je ne parlerai pas des problèmes soulevés par la création d'un corpus, dont le plus intéressant est sans doute celui de l'équilibre entre les sources (voir aussi Knowles, dans ce volume).

La possibilité d'accéder à un large corpus de textes électroniques ne représente pas l'unique moyen d'observer la langue telle qu'elle est employée, et la lexicographie bénéficie d'une longue et respectable tradition en matière de collecte de faits de langue à partir desquels les articles de dictionnaires sont construits. La cueillette de citations de langue "réelle" et l'utilisation de ces citations dans les dictionnaires monolingues vise depuis longtemps deux objectifs assez différents. Pour Samuel Johnson, qui par l'importance qu'il accordait à l'authenticité, à la syntaxe et aux collocations, est considéré à juste titre comme le père de la lexicographie anglaise telle qu'elle se pratique aujourd'hui, les citations avaient pour but essentiel d'illustrer le sens et l'usage des mots – objectif empirique plus que normatif :

«Certains exemples sont pris chez des auteurs qui ne sont considérés ni comme des maîtres de l'élégance, ni comme des modèles pour le style, mais on doit aller chercher les mots là où on les utilise, et dans quels écrits réputés pour leur pureté peut-on trouver des termes relatifs à l'industrie et à l'agriculture ? De nombreuses citations n'ont d'autre objectif que de prouver l'existence pure et

---

3 Le British National Corpus a été mis en place, avec l'aide de fonds gouvernementaux, par un consortium dirigé par Oxford University Press, et comprenant les universités d'Oxford et de Lancaster, les éditeurs Chambers, Longman, et la British Library. Ce corpus atteint maintenant 100 millions de mots, conformément aux objectifs de départ.

4 Le présent chapitre évoque l'utilisation d'un corpus dans le processus de création d'un dictionnaire et décrit une méthodologie pour exploiter deux corpus entièrement indépendants, l'un en anglais et l'autre en français. Aucun de ces corpus ne contient de texte traduit de l'autre. Un grand nombre de travaux très intéressants ont été effectués dans le domaine du traitement automatique des langues naturelles et de la traduction assistée par ordinateur utilisant un corpus «bilingue», contenant des textes traduits. (Voir par exemple Warwick et Russel, 1990 et Church et Gale, 1991).  
Pour la rédaction de dictionnaires, il est sans doute préférable d'utiliser des textes dans leur langue et leur version d'origine.

simple des mots et sont de ce fait sélectionnées avec moins d'attention que celles qui ont pour but *d'enseigner les structures et les relations*. (Johnson: 1755)  
(italiques ajoutées par VG)

Par contre, lorsqu'une centaine d'années plus tard, James Murray et ses collègues conçurent leur vaste programme de collecte de citations pour ce qui s'appelait encore à l'époque le *New English Dictionary* et qui est devenu depuis l'*Oxford English Dictionary*, ils n'avaient pas pour objectif principal d'étudier l'usage courant, mais plutôt l'évolution du sens au cours de l'histoire. Leur programme de lecture et de collecte de citations visait à repérer les dates d'entrée des mots dans la langue et ensuite les "dates de naissance", pour ainsi dire, des nouveaux sens. La collecte de citations est toujours très importante pour les lexicographes et l'*Oxford English Dictionary* continue de rassembler 40 000 citations par mois. L'accent n'est plus exclusivement mis sur l'historique de la langue, et une série impressionnante de livres et de journaux, britanniques, américains et autres, sont passés au peigne fin chaque année.

Une telle procédure, même lorsqu'elle n'a pas pour but de faire apparaître l'évolution historique, reflète inévitablement ce qui est inhabituel et qui n'a pas encore été enregistré, et ce faisant, permet de suivre l'évolution du vocabulaire et les changements de sens. La périphérie de la langue, qui va de l'argot et du jargon à la terminologie spécialisée, change très rapidement et il y a de bonnes raisons de penser que le domaine de la création des mots – qui passe souvent plus par la langue parlée que par la langue écrite – peut être abordé encore aujourd'hui plus efficacement par l'étude systématique et détaillée de publications et d'émissions de télévision que par l'utilisation d'un corpus (voir Thompson, 1992). Le *British National Corpus* contient des textes d'anglais parlé transcrits mais ces documents, précieux par ailleurs, ne sauraient être, pour des raisons techniques, suffisamment à jour pour servir de référence pour les termes les plus récents et expressions nouvelles.

Il ne fait aucun doute que la lexicographie bilingue devrait s'intéresser à ce qui est nouveau, pittoresque, exotique. Les listes de «mots nouveaux» qui apparaissent sur la jaquette de pratiquement tous les dictionnaires commercialisés actuellement sembleraient indiquer que le public se laisse volontiers tenter par la promesse de trouver dans le dictionnaire les néologismes ainsi que, dans le cas d'un dictionnaire bilingue, leurs équivalents dans l'autre langue. Cependant, il serait faux de penser que la présence de ces mots prouve que le dictionnaire est moderne et exhaustif. Johnson était déjà conscient de l'attrait que représente tout ce qui est rare ou nouveau mais il savait aussi que ceci ne doit constituer ni le principal atout du dictionnaire ni le but principal du lexicographe :

«Celui qui ne retient que ce qui se produit rarement négligera ce qu'il s'attend à rencontrer régulièrement, celui qui est à la recherche de mots rares et lointains négligera ceux qui sont évidents et familiers». (Johnson, 1755)

On utilise les dictionnaires bilingues comme des outils pratiques, comme des aides à la traduction, active ou passive. De ce point de vue, le critère essentiel

selon lequel l'utilisateur juge le dictionnaire sera : «Est-ce que j'y trouverai ce que je cherche ?». Dans la majorité des cas, l'utilisateur sera à la recherche non pas de ce qui est récent et éphémère mais plutôt des moyens d'expression essentiels. Quand il s'agit de l'enregistrement et de l'analyse de la langue générale, le principe de la collecte de citations, basé sur le travail de centaines de lecteurs qui parcourent différentes publications, malgré sa valeur inestimable dans d'autres domaines, n'est pas à la hauteur de la tâche.

D'un autre côté, aucune équipe de lexicographes ne peut espérer mener à bien ce travail herculéen de documentation et d'analyse sans avoir accès à des textes. Quelle est la fréquence de telle ou telle unité lexicale et quelle est son importance pour la communication ? Quelle place doit-on lui accorder dans un dictionnaire ? Quels sont les modèles syntaxiques de base qu'elle exploite ? Quelle acception est la plus fréquente ? Quels sont les exemples les plus typiques de son utilisation dans chacune de ses acceptions ? Quelles acceptions sont devenues vieilles ou obsolètes ? Qu'est-ce qui constitue un véritable changement de sens et qu'est-ce qui relève simplement d'une préférence contextuelle ? Un travail sur la langue qui se fait uniquement par introspection produira inévitablement un grand nombre d'exemples d'usages potentiels (par opposition à des usages réels), généré par la peur bien compréhensible des lexicographes d'oublier quelque chose d'important. Il est pratiquement impossible, par exemple, pour un lexicographe de dresser, sans recours à des textes, la liste des substantifs qui sont *typiquement* qualifiés par tel ou tel adjectif, par rapport à ceux qui le sont *potentiellement*. Cela est vrai aussi des compléments d'objet typiques associés à un verbe (voir Hanks, 1998 et Fontenelle, 1992).

Les méthodes basées uniquement sur l'introspection risquent donc de produire un excès d'informations linguistiques, inutiles parce qu'elles sont plus imaginaires que réelles. Inversement, de telles méthodes conduisent inévitablement à des lacunes, puisque même le plus brillant des lexicographes ne peut produire instantanément une description systématique du comportement caractéristique d'une unité lexicale donnée. Ce qui frappe le plus quand on utilise un corpus de manière quotidienne, c'est l'écart – littéralement énorme – entre la "connaissance" que peut avoir un locuteur du comportement d'une unité lexicale et le comportement réel de cette unité tel qu'on l'observe dans le corpus. Ce décalage ne devrait pas nous surprendre, puisqu'il correspond parfaitement à ce que constatent les psycholinguistes à propos de l'incapacité de l'être humain d'analyser ou de décrire les processus de production du discours. Ce que décrit Sinclair correspond sans aucun doute à l'expérience de tout lexicographe :

«(...), il y a une différence énorme et systématique entre les impressions des gens notent sur les détails de la langue et les faits de langue réels que l'on peut observer dans les textes. Cela laisse supposer que les êtres humains ont une idée tout à fait particulière de la langue et que cette idée ne correspond pas du tout à ce qui se passe effectivement lorsqu'ils utilisent cette langue». (Sinclair, 1991:4)

Une chose est claire : l'utilisateur est en droit d'attendre de son dictionnaire qu'il répertorie les usages réels et non ceux que le lexicographe a imaginés. Quelle que soit l'ingéniosité avec laquelle on peut prévoir les besoins de cet utilisateur, il est difficile d'envisager que ces besoins, dans le cadre de la compréhension de la langue source ou celui de la communication en langue cible, puissent être fondés sur autre chose que des exemples d'utilisation du langage en situation réelle.

Comment le lexicographe bilingue s'y prend-il pour utiliser un corpus de façon à répondre aux besoins des utilisateurs dans le contexte hautement formalisé et condensé d'un article de dictionnaire ? Pour pouvoir répondre à cette question il faut d'abord examiner comment on peut identifier et décrire les besoins des utilisateurs pour un dictionnaire particulier. Ensuite je décrirai brièvement les processus éditoriaux utilisés dans la rédaction d'un dictionnaire bilingue, avant d'étudier comment le corpus peut être utilisé dans chacun de ces processus.

## 2 Un dictionnaire pour qui et pour quoi faire ?

Nous avons vu que dans le cas d'un gros dictionnaire publié par une maison d'édition, c'est l'éditeur qui décide à quels groupes d'utilisateurs le dictionnaire est destiné. En ce qui concerne les dictionnaires bilingues, un certain nombre de critères fondamentaux déterminent le type ou les types d'utilisateurs visés. Les ventes du dictionnaire seront-elles de la même importance dans chacune des deux communautés linguistiques ? Quel sera l'âge des utilisateurs et leur niveau dans la langue étrangère ? Dans le cas d'un gros dictionnaire général français > anglais / anglais > français, lorsque chacune des deux langues impliquées est parlée par une vaste communauté, et lorsque chaque communauté comporte une proportion relativement importante d'habitants qui apprennent l'autre langue du dictionnaire ou s'en servent dans le cadre de leur travail, les besoins des utilisateurs sont complexes.

Dans l'introduction de l'OHFD, quatre groupes distincts d'utilisateurs sont identifiés :

- le locuteur anglophone qui veut comprendre le français,
- le locuteur anglophone qui veut écrire ou parler le français,
- le locuteur francophone qui veut comprendre l'anglais,
- le locuteur francophone qui veut écrire ou parler l'anglais.

Ainsi, dans chaque groupe de locuteurs, il y a deux types d'utilisateurs : celui qui décode (de L2 vers L1), et celui qui encode (de L1 vers L2). Les besoins de ces groupes sont différents, à tel point que les métalexigraphes en ont souvent conclu qu'aucun article de dictionnaire ne pouvait les satisfaire tous (voir par exemple Snell-Hornby, 1984). Le but de ce chapitre n'est pas de polémiquer sur

ce sujet. L'idéal serait sans doute soit un dictionnaire conçu pour une seule communauté linguistique, soit un dictionnaire en deux volumes, l'un pour décoder et l'autre pour encoder. Cependant, l'investissement nécessaire pour créer un dictionnaire vraiment fiable et vraiment nouveau est tel qu'aucun éditeur ne s'y lancerait sans être assuré qu'il pourra vendre le produit final. En outre, plus le marché visé est restreint, plus le prix du dictionnaire sera élevé. Quelle que soit l'utilité d'un outil, il sera peu bénéfique pour ceux qui en ont besoin si seule une petite minorité d'entre eux peut se l'offrir. Toute politique éditoriale doit donc viser à satisfaire les besoins des utilisateurs moyens pour l'encodage et pour le décodage. Pour le lexicographe, cela implique un énorme travail d'analyse et de sélection car il est évident que l'information nécessaire à l'encodage – la production dans la langue étrangère – ne peut trouver sa place dans le dictionnaire qu'au détriment de la quantité de termes et des expressions nécessaires pour le décodage. Le lexicographe devra faire des choix déchirants et il aura besoin pour cela de toute l'aide que peut lui apporter le corpus, pour être sûr que le dictionnaire réponde aux besoins du plus grand nombre possible d'utilisateurs.

Le lexicographe a pour responsabilité de faire en sorte que le dictionnaire soit :

- 1 complet : l'utilisateur, conscient de l'envergure du dictionnaire qu'il est en train de consulter, doit trouver ce qu'il cherche;
- 2 riche en informations : l'utilisateur qui veut encoder a besoin d'aide pour utiliser la langue étrangère, alors que l'utilisateur qui veut décoder cherche des renseignements sur des mots et des expressions inconnus qu'il rencontre dans la langue étrangère;
- 3 fiable : les traductions proposées par le dictionnaire ne doivent ni tromper l'utilisateur ni le faire douter.

Ce sont donc ces motivations, de même que les différents profils des utilisateurs, qui vont guider les processus de mise au point de la politique éditoriale générale du dictionnaire, et qui continueront par la suite à motiver les prises de décisions quotidiennes des lexicographes. Comment ceux-ci peuvent-ils utiliser le corpus de façon à répondre aux différents besoins des utilisateurs ? Comment l'utilisation du corpus peut-elle les aider à choisir ce qu'il faut inclure dans le dictionnaire, et ce qui peut-être laissé de côté ? Comment l'information doit-elle être présentée ?

### **3 Les processus de la lexicographie bilingue**

La méthodologie mise en œuvre pour la création d'un dictionnaire varie considérablement en fonction de l'ampleur du projet, de la nature des données linguistiques disponibles, et des ressources, essentiellement financières et matérielles, qui ont été allouées. Néanmoins, quelle que soit l'ampleur ou la complexité de l'entreprise, trois processus distincts entrent en jeu. Les deux langues sources doivent d'abord être analysées séparément. Dans un second temps, un

lexicographe dont la langue maternelle est la langue cible doit procéder à une analyse des faits et des exemples d'usage de la langue source et les transférer dans la langue et la culture cibles. C'est un processus difficile. Une bonne partie des données sélectionnées en langue source ne pourront pas être traduites directement en langue cible : les structures et les idiotismes donnés sous forme canonique en langue source n'auront bien souvent pas d'équivalent direct en langue cible et devront donc être intégrés dans des contextes sélectionnés pour être le plus représentatif et productif possible. Des groupes de collocations, servant d'indicateurs de contexte, donnés pour une unité lexicale en langue source, devront être séparés en plusieurs équivalents en langue cible pour une même unité, etc. Inversement, de nombreuses nuances qui posent problème dans le processus de transfert d'une langue à l'autre peuvent ne pas avoir été repérées par le lexicographe qui travaille dans la langue source.

Dans un troisième temps, pour mettre au point l'article de dictionnaire définitif, les deux lexicographes doivent collaborer à l'organisation de l'information, à la sélection de ce qui est le plus important pour le groupe d'utilisateurs de leur langue de travail et à la résolution de toutes les ambiguïtés entre les différents équivalents de traduction. Ces trois étapes sont les processus que Atkins désigne respectivement par *analyse, transfert et synthèse (analysis, transfer and synthesis)* (Atkins, 1990).

Le premier de ces processus, celui de l'*analyse*, concerne le lexicographe de langue source, qui doit fournir au lexicographe de langue cible un cadre structuré en langue source. La première tâche est d'assembler la macrostructure de la langue source selon des principes établis lors de la conception du dictionnaire. Ces principes auront déterminé le type de vocabulaire qu'il faut traiter (général, spécialisé, littéraire, jargon, encyclopédique etc.), les types d'unités lexicales auxquelles on doit donner le statut de mot vedette et ceux qui pourraient éventuellement être relégués en sous-rubriques (noms composés, verbes à particules, idiotismes, formes dérivées, etc.). Le lexicographe de langue source doit ensuite procéder à une analyse structurée de chaque unité lexicale retenue. Les informations données dans les dictionnaires monolingues ne suffisent pas à identifier les faits de la langue source qui sont essentiels dans un dictionnaire bilingue. Les dictionnaires monolingues généraux tendent à favoriser la sémantique au détriment de présentations systématiques de la syntaxe et de l'usage (voir Lépinette, dans ce volume). Vu l'importance de ces faits dans un dictionnaire bilingue le lexicographe doit aller chercher des renseignements ailleurs, c'est-à-dire dans sa propre intuition sur la langue ou, dorénavant, dans un corpus<sup>5</sup>.

Nous avons déjà dit qu'un dictionnaire bilingue général doit être complet, moderne et fiable. Pour être complet, il doit également proposer un traitement

5 De plus, les renseignements sur les distinctions de sens que l'on trouve dans un dictionnaire monolingue ne sont pas forcément utiles pour les dictionnaires bilingues (voir Winter, 1992:42).

adéquat du vocabulaire spécialisé et semi-spécialisé, sans négliger ce qui est inhabituel, les formes et les emplois littéraires qui pourraient rester opaques pour un lecteur étranger. Mais le vocabulaire hautement technique et spécialisé et le vocabulaire littéraire doivent être considérés comme des sous-groupes en ce qui concerne le corpus. Il est difficile de mettre au point des définitions et des équivalences de traduction pour les termes spécialisés, et un corpus de textes de langue générale ne peut pas être très utile. Nous ne parlerons pas de ce problème car la mise au point de corpus de langues spécialisées est un sujet à part. Pour le lexicographe monolingue général, ce sont les mots quotidiens qui sont les plus difficiles à analyser et il utilisera donc le corpus, comme d'autres spécialistes, pour observer non pas le vocabulaire rare et exotique, mais le vocabulaire typique (voir Hanks 1988, et Sinclair 1991).

Le rôle du lexicographe en langue source est d'établir les faits essentiels du comportement des unités lexicales de la langue dans laquelle il travaille. En gros, on peut distinguer les informations nécessaires à l'encodage et les informations nécessaires au décodage. A la première catégorie appartiennent toutes les informations concernant la façon dont une unité lexicale s'insère dans la syntaxe d'une phrase pour créer du sens. L'utilisateur qui encode cherche à exprimer du sens dans la langue cible, et cette catégorie comprend donc également les informations sur la structure, les collocations et les compléments, de même que tous les moyens fixes, mais imprévisibles, de combiner les lexèmes pour créer du sens : *you have only yourself to blame, you wouldn't catch me doing that, don't tell a soul, manger à sa faim, avoir la parole facile*.

Dans la seconde catégorie, on trouve toutes les combinaisons dont le sens est différent ou plus vaste que la somme de leurs différentes parties, et qui ne sont donc pas accessibles au locuteur de la langue cible par une simple opération de déduction : noms composés (*blister pack, couch potato*), verbes à particules (*carry out, hold forth*) et syntagmes fixes (*to give the game away, to be out to lunch; c'est l'hôpital qui se moque de la charité, rouler les mécaniques, être fleur bleue*). Par contre l'utilisateur qui décode, une fois qu'il a compris le sens d'un terme ou d'une expression dans l'autre langue, se retrouve tout de suite dans l'univers de sa propre langue, qu'il connaît bien, et il n'aura aucune difficulté à retrouver les articulations syntaxiques, prépositions et autres, qui posent un grave problème de traduction pour l'encodeur.

#### 4 L'accès au corpus

Les lexicographes travaillant sur corpus utilisent en général des outils de recherche qui affichent les résultats à l'écran sous forme de lignes de concordances *KWIC* (*keyword in context*), qui font normalement apparaître toutes les occurrences d'une forme dans une colonne au centre de l'écran. Pour chacune de

ces formes le programme de recherche fournit la liste des occurrences ainsi que le contexte immédiat. La taille d'un corpus se mesure au nombre de «mots-occurrences», c'est-à-dire de «chaînes de caractères bordées par un espace de chaque côté : mots *stricto sensu*, mais aussi abréviations, et même fautes de frappe etc.»

Les programmes de concordances utilisés actuellement varient considérablement selon les besoins auxquels ils correspondent. Les lexicographes et les linguistes disposent d'outils de plus en plus sophistiqués pour analyser les textes volumineux qu'ils exploitent et pour adapter leurs recherches à leur objectif et au type d'unité lexicale étudié. Nous n'aborderons pas ici la description de tous les outils disponibles, mais il est utile de procéder à un bref inventaire de ceux qui peuvent aider le lexicographe dans l'analyse des données d'un corpus.

Pour explorer un corpus, on peut rechercher par préfixe et par suffixe, avec caractères «joker», par une approche booléenne, par listes de formes etc. Un analyseur morphologique intégré au système facilite la recherche des formes fléchies. Si le corpus contient des étiquettes désignant les parties du discours, le lexicographe pourra en rechercher certaines et en éliminer d'autres. Il peut par exemple vouloir étudier le verbe *urge* séparément du substantif. Cela est particulièrement important en anglais, où il y a beaucoup plus de formes correspondant à plusieurs parties du discours qu'en français. Si le programme de recherche permet de distinguer entre majuscules et minuscules, le lexicographe pourra éliminer, par exemple, toutes les occurrences d'un nom propre ayant la même forme qu'un nom commun. Si le corpus a des étiquettes syntaxiques, le lexicographe pourra effectuer des recherches pour établir, par exemple, la liste des compléments d'objets directs typiques de tel ou tel verbe.

Il est également utile de pouvoir spécifier le nombre de lignes de concordances à afficher ou bien de pouvoir sélectionner un sous-corpus en fonction du type d'unité lexicale étudié. La présentation des informations à l'écran peut varier. Par exemple, le lexicographe travaillant sur les compléments d'un verbe anglais gagnera du temps si les lignes de concordances sont organisées par ordre alphabétique en fonction de la lettre initiale du mot qui suit immédiatement le mot clef, c'est-à-dire par ordre alphabétique à droite. Il verra donc apparaître sous forme de blocs sur son écran les suites d'objets les plus typiques. Par contre, s'il veut étudier les sujets du verbe, il aura besoin de contextes ordonnés à gauche. Mais il existe maintenant des outils qui vont encore plus loin et qui, à partir d'une analyse statistique, présentent la liste des collocations les plus fréquentes, ainsi que leur taux de fréquence. La mise au point d'outils d'analyse des corpus est actuellement l'un des domaines les plus actifs de la linguistique informatique, pour le plus grand bénéfice à la fois des lexicographes et des spécialistes du traitement automatique du langage naturel.

Il arrive souvent que le contexte affiché à l'écran pour une occurrence soit insuffisant. Ce problème se présente souvent pendant la phase de transfert, car le

lexicographe de langue cible peut avoir besoin d'un contexte plus large que celui qui est à l'écran pour déterminer quelle est la traduction la plus appropriée. Il devrait donc être possible de sélectionner n'importe quelle ligne de concordance et de faire afficher un contexte plus large, c'est-à-dire plusieurs lignes de texte, une phrase entière, ou même un paragraphe entier.

Il est également essentiel de connaître les coordonnées des textes sources : le lexicographe a besoin de savoir dans quel contexte apparaît une unité lexicale ou un usage (écrit ou oral, texte scientifique, littéraire, journalistique, etc.). Un usage apparemment intéressant peut correspondre à un nombre élevé d'occurrences, mais si toutes ces occurrences sont issues de la même source leur nombre perd toute signification pour le lexicographe.

## 5 L'utilisation du corpus dans l'analyse en langue source

Je ne parlerai pas ici de l'utilisation du corpus pour constituer la nomenclature du dictionnaire. Disons simplement qu'il peut être intéressant de comparer les index de nomenclatures de dictionnaires avec ceux d'un ou de plusieurs corpus. L'exploration du corpus peut aussi aider au repérage d'unités lexicales complexes. Une exploration rapide des lignes de concordance de *brand* permet d'obtenir les noms composés suivants : *brand acceptance*, *brand awareness*, *brand identification*, *brand image*, *brand leader*, *brand loyalty*, *brand management*, *brand name*, *brand recognition* et *brand switching*. Avec *market*, on obtient *market analysis*, *market analyst*, *market-based*, *market-led*, *market capitalization*, *market cross*, *market day*, *market economy*, et des dizaines d'autres.

L'analyse du corpus produit également des résultats intéressants en ce qui concerne les collocations. Le contexte a une importance primordiale pour la mise en correspondance des deux langues (voir Szende, dans ce volume), et pour procéder au transfert, les lexicographes de langue cible doivent pouvoir disposer des caractéristiques essentielles de chaque mot vedette en matière de collocation. Ces informations seront également utilisées pendant la phase de synthèse, et pourront aider à ordonner les traductions, en utilisant les collocatifs les plus typiques comme indicateurs de contexte. L'analyse statistique des données du corpus peut fournir des listes de sujets et de compléments d'objet des verbes ordonnés selon leur fréquence d'apparition dans le corpus, des listes des substantifs les plus fréquemment modifiés par tel ou tel adjectif et celles des verbes et des adjectifs modifiés par tel ou tel adverbe (voir Church, Gale, Hanks et Hindle, 1991). La différence entre les collocations typiques choisies intuitivement par le lexicographe et celles qui apparaissent dans un corpus a été démontrée (Fontenelle, 1992).

Un bon article de dictionnaire bilingue doit impérativement fournir tous les moyens d'aide dont l'utilisateur peut avoir besoin pour insérer la traduction d'un

lexème isolé dans un syntagme de langue cible. L'identification des informations nécessaires à cette opération est particulièrement problématique. La consultation systématique du corpus produit des résultats étonnants en ce qui concerne les compléments, les structures, les collocations et les caractéristiques sémantiques d'un contexte. L'analyse du corpus avec contexte ordonné à droite en ce qui concerne le substantif *advantage*, très fréquent en anglais, produit ce qui suit :

advantage + PREP *for*  
 advantage + PREP *in* + V-ing  
 advantage + PREP *of*  
 advantage + PREP *over*  
 advantage + CONJ *that*  
 advantage + V *be* + CONJ *that*  
 advantage + V-INF  
 advantage + PREP *to*

L'analyse du contexte ordonné à gauche produit :

V *be* + ART INDEF + advantage  
 V *be* + PREP *at* + ART INDEF + advantage  
 V *have* + ART INDEF + advantage  
 ART DEF + advantage  
 V *be* + PREP *to* + POSS + advantage  
 PREP *to* + POSS + ADJ + advantage  
 PREP *to* + POSS + ADJ *good/full/best/better* + advantage  
 V *take* + advantage

Si l'on combine les résultats de ces deux recherches pour le nom *advantage*, on obtient :

DET(ART INDEF) + advantage + PREP *for* (*this might be an advantage for some..., it was also an advantage for the government that few country boroughs had..., that is a psychological advantage for us*).

DET + advantage + PREP *in* + V-ing (...*if they perceive an advantage in doing so, we don't see any advantage in staying small, there is no advantage in using hardwood*)

DET + advantage + PREP *of* + NP (*the advantage of Palin's unshakeable charm is..., you can have the advantage of a bank account*)

DET + advantage + PREP *of* + V-ing (*one advantage of accepting this conclusion is that..., a small holiday park which has the advantage of being quiet, there is an electoral advantage in keeping this quiet*)

V *have* + DET + advantage + PREP *over* (*the presenter of a television series has some advantage over a writer, it has no advantage over other investors*)

V *have/offer* + DET + advantage + THAT CLAUSE (*this has the obvious advantage that all countries match their local time..., the first has the advantage that no time is wasted*)

DET + advantage + V *be* + THAT CLAUSE (*its major advantage is that it trains students to deal with..., the second advantage is that the music was written by an American, the advantage over slides and home movies is that...*)

V be + advantage + to-INF (it can be an advantage to have a low background heat..., it may be an advantage to lower the flap)

V be + PREP to + POSS + advantage + to-INF (if it is clearly to people's advantage to be on AZT..., it could be to her advantage to be perceived as a celluloid creation)

PREP to + ADJ good + advantage (the main parts of the system can usually be prototyped to good advantage)

PREP to + POSS + best/better/full/fuller/fullest advantage (show off the clothes to their best advantage, utilizing the energy to better advantage, exploit the green, green grass of SW19 to its fullest advantage)

V be + PREP at + DET (ART INDEF) + advantage (clearing houses with a computerized database are at an advantage here)

V take advantage.

Nous avons ainsi ce que l'on peut considérer en toute confiance comme un schéma complet des structures syntaxiques d'*advantage* dans des documents très variés. A vrai dire, le mot *advantage* lui-même ne présente pas de réels problèmes de traduction vers le français. Il est presque toujours traduit par *avantage*, parfois par *intérêt*. Les problèmes viennent précisément du fait que *advantage* est utilisé dans plusieurs schémas syntaxiques différents, dont la plupart sont difficilement prévisibles et de ce fait difficile à traduire, surtout pour l'utilisateur de langue source. Une liste des structures et des schémas comme ceux que nous avons mentionnés ci-dessus fournit le cadre dans lequel le lexicographe de langue cible travaillera pour le transfert de la langue source vers la langue cible et pour la mise au point des équivalences. Il est clair qu'aucun lexicographe ne peut obtenir une telle quantité d'informations en se basant uniquement sur la connaissance intuitive qu'il a de la langue. On entend souvent dire «tout le monde le sait», mais cela est tout à fait différent de la capacité d'un locuteur natif – et même d'un lexicographe – de produire ces faits de langue lorsqu'on lui demande.

Cependant, ce cadre est insuffisant en lui-même. Le lexicographe de langue cible ne peut pas fournir des exemples de traduction pour des structures syntaxiques présentées dans «un vide sémantique». Par conséquent, le lexicographe de langue source doit fournir une liste complète des caractéristiques sémantiques typiques associées à chacune des structures syntaxiques. Après cela le lexicographe de langue cible devra encore revenir au corpus en langue source pour travailler plus à fond sur différents types de contextes typiques de façon à être capable de fournir pour chaque structure des traductions à la fois fiables et naturelles.

Chaque structure est ensuite divisée en groupes sémantiques ordonnés, trop volumineux pour être reproduits ici. Je prendrai un seul exemple pour illustrer le processus. Dans le *Oxford Pilot Corpus* (60 millions de mots), il y a 3 951 occurrences d'*advantage* au singulier seulement et 5 361 occurrences du singulier et du pluriel combinés. Une recherche pour *take/takes/taking/took advantage* permettant deux mots maximum entre le verbe et le nom (*take full advantage, take the fullest advantage* etc.) produit 1 108 occurrences. Cela signifie que sur toutes les occurrences d'*advantage(s)*, 20,68 % correspondent à *take advantage* et ses

formes fléchies. De plus, si l'on considère la forme singulière d'*advantage* seulement, *take advantage* et ses formes fléchies représentent 28 % des occurrences.

Ces résultats ont des conséquences importantes en ce qui concerne l'article *advantage*. Nous avons vu que l'objet principal d'une entrée de dictionnaire bilingue est de fournir des équivalents en langue cible aux structures et aux expressions que l'utilisateur est susceptible d'avoir besoin de traduire. Une simple analyse comme celle que nous avons décrite ci-dessus permet au lexicographe d'avoir accès aux informations essentielles sur ce qui peut être conservé et ce qui peut être rejeté.

L'article du dictionnaire doit accorder une attention toute particulière au traitement de *take advantage*, puisque cette expression joue un rôle important dans le discours. Bien sûr, il peut apparaître au cours de la phase de transfert que, malgré sa haute fréquence, le lexème en langue source pose peu de problèmes de traduction et par conséquent ne mérite pas un traitement approfondi dans l'article final du dictionnaire. Néanmoins, avant d'en arriver à cette conclusion, le lexicographe de langue cible devra examiner tous les aspects du lexème de la langue source pour être certain que ce qui est proposé comme traduction est utilisable dans tous les cas.

En matière de lexicographie bilingue, le fait de distinguer des catégories sémantiques discrètes dans un article consacré à un mot permet moins de fournir une «anatomie» satisfaisante du mot que d'offrir à l'utilisateur des points d'accès à la traduction et aux informations sur l'usage. Ce qui est important, c'est d'identifier le contexte sémantique que le mot vedette attire ou génère et toutes ses structures syntaxiques essentielles, de façon à repérer les équivalences de traduction qui fonctionneront dans le même contexte. Cette méthode essentiellement pragmatique d'aborder l'analyse linguistique en passant de la syntaxe au sens est la plus efficace pour appréhender les faits de langue essentiels contenus dans le corpus. Nous disposons d'outils pour faciliter l'analyse syntaxique, mais ceux qui pourraient aider à l'analyse sémantique semblent appartenir à un avenir plus lointain. Le type d'analyse syntaxique détaillée que l'on a présentée ci-dessus permet de repérer les distinctions de sens qui sont les plus utiles dans un dictionnaire bilingue ou pour apprenants étrangers, voire dans un dictionnaire monolingue de langue maternelle. Dans le cas de *take advantage*, on peut extraire du corpus les caractéristiques sémantiques essentielles suivantes :

### 1 *PERSON or HUMAN GROUP takes advantage of FACILITY*

Cette catégorie est de loin la plus vaste et peut être répartie en deux domaines, GENERAL et COMMERCIAL.

Il fallait différencier ces deux domaines, puisque de très nombreuses citations repérées dans les lignes de concordance sont issues du domaine commercial et

financier. Exemples d'unités lexicales du domaine GENERAL : *facilities, roadways, beaches, runway space, school, training scheme, scholarship, tennis court, sauna and solarium, technique, technology, water power, Channel Tunnel.*

On peut résumer ce SENS par : *derive benefit by using.*

Exemples d'unités lexicales du domaine COMMERCIAL : *offer, credits, discount, cheap flights, service, deal, increase in value, lower prices, tax allowance, tax relief, special arrangement.* Il y a également un certain nombre d'exemples issus du domaine de la publicité (*take advantage of a special offer* etc).

On peut résumer ce SENS par : *derive financial benefit by using.*

## 2 **PERSON or HUMAN GROUP takes advantage of**

Exemples d'unités lexicales de la catégorie EVENT : *opportunity, change, liberalization, visit, discussion, talks, diversion, victory, disturbance, disaster, uprising, recession.* Une étude minutieuse des lignes de concordance de cette catégorie révèle la structure syntaxique suivante qui lui est souvent associée :

PERSON or HUMAN GROUP *takes advantage of* EVENT *to do something*

On peut résumer ce SENS par : *react to (in order to do something one wants to do).*

## 3 **PERSON or HUMAN GROUP takes advantage of SITUATION**

Exemples d'unités lexicales de la catégorie SITUATION : *situation, circumstances, liberty, conditions, moment, weather, climate, vie, turmoil.*

On peut résumer ce SENS par : *derive benefit from.*

## 4 **PERSON or HUMAN GROUP takes advantage of PERSON or HUMAN GROUP.**

Exemples d'unités lexicales de cette catégorie : *parents, workers, senior citizens, lonely women, vulnerable patients.*

On peut résumer ce SENS par : *exploit unfairly.*

## 5 **PERSON or HUMAN GROUP takes advantage of HUMAN QUALITY, EMOTION or STATE.**

Exemples d'unités lexicales dans cette catégorie : *person's absence, generous nature, weakness, trust, euphoria, discontent, grief, good humour, ignorance.*

On peut résumer ce SENS par : *exploit unfairly.*

Cette analyse du contexte syntaxique et sémantique fournit le matériel nécessaire au lexicographe de langue cible pour mettre en place la phase de transfert et l'identification d'équivalents productifs en langue cible.

## 6 Le transfert

La première tâche du lexicographe est de proposer des traductions générales pour chacune des classes du mot ou de l'expression vedette. Cela n'est pas toujours possible, mais dans le cas de *take advantage*, la tâche est relativement simple et peut se résumer comme suit :

1. GENERAL utiliser, profiter de
1. COMMERCIAL profiter de
2. EVENEMENT profiter de
3. SITUATION utiliser, profiter de
4. PERSONNE utiliser
5. QUALITE, EMOTION, ETAT profiter de

Le lexicographe de langue cible dispose donc désormais d'un ensemble d'hypothèses qu'il peut comparer avec les données du corpus en langue source. Il vérifiera les traductions qu'il propose en fonction des compléments d'objet typiques de chaque classe et également dans le contexte plus vaste des citations considérées comme normales et typiques par le lexicographe de langue source. Ce procédé permet parfois de trouver des traductions plus adaptées à certains contextes particuliers, et de faire apparaître des domaines d'usage pour lesquels il n'est pas possible de trouver une traduction directe. Dans ce dernier cas, le lexicographe de langue cible fournira des traductions pour toutes les lignes du corpus qui sont problématiques, en s'efforçant d'identifier la nature du problème de traduction et de parvenir à un moyen de le résoudre.

Jusqu'à ce stade du processus, le lexicographe de langue cible ne travaillait qu'avec son intuition : il donnait spontanément des équivalents de traduction pour la langue source dans un grand nombre de contextes différents. Cependant, lorsqu'il a des difficultés pour trouver une traduction satisfaisante, il peut rechercher des contextes parallèles dans le corpus de langue cible. Dans de nombreux cas, il devra également vérifier dans le corpus de langue cible les traductions proposées initialement. Cette opération ne fait parfois que confirmer les traductions proposées mais est néanmoins nécessaire pour rassurer le lexicographe et garantir la fiabilité du dictionnaire.

Le corpus de langue cible peut également s'avérer très utile pour identifier les zones à problèmes. J'en évoquerai ici deux qui depuis toujours ont empoisonné l'existence des lexicographes et ont donné prise aux critiques justifiées des métalexigraphes. La première concerne le degré d'équivalence sémantique et/

ou syntaxique des mots apparentés dans les deux langues, les “faux amis” potentiels. La deuxième concerne les quasi-synonymes proposés comme traductions et la nécessité d’une analyse contrastive permettant de les distinguer clairement. Le verbe français *inciter* illustre ces deux cas. L’analyse des citations du corpus montre qu’il est utilisé dans un grand nombre de contextes différents.

### 1 **PERSONNE ou PAROLE ou ATTITUDE incite à une ACTION ou ATTITUDE**

(des écrits incitant à la haine raciale, on titille l’intérêt tout en incitant à la prudence, une incompréhension qui incite à l’angoisse).

On peut classer les collocations correspondant à la catégorie ACTION/ ATTITUDE selon qu’elles s’accompagnent de connotations «neutres/bonnes» ou «mauvaises». Exemples de connotations «neutres/bonnes» : *dynamisme, modération, épargne, économies, activité*. Exemples de connotations «mauvaises» : *émeutes, violence* (les connotations «mauvaises» sont très peu nombreuses).

### 2 **PERSONNE ou GROUPE HUMAIN incite PERSONNE ou GROUPE HUMAIN à + V-INF**

(Michel Cavaille incite plusieurs grandes villes à étudier l’aménagement de nouvelles lignes, la presse locale incite la police à intervenir).

Les verbes sont associés à un grand nombre de champs sémantiques différents : *embaucher, acheter, réfléchir, se réfugier, découvrir, améliorer, tirer les conséquences, voter, répondre, etc.*

### 3 **ÉVÈNEMENT ou SITUATION incite PERSONNE ou GROUPE HUMAIN à + V-INF**

(Ce phénomène incite les Américains à essayer de se protéger, la croissance incite les entreprises à embaucher, la conjoncture économique incite les industriels à supprimer des emplois).

Il est difficile d’imaginer à partir des dictionnaires monolingues que les mots apparentés *inciter* et *incite* diffèrent en termes de fréquence, de sens ou d’emploi : «conduire (qqn) à un sentiment, un comportement, par une influence morale (surtout avec sujet de chose)» (*Le Petit Robert*); « I. Urge, spur on, (a person); stir up, animate, stimulate to do something» (*Shorter Oxford English Dictionary*), «to stir up or provoke to action» (*Collins English Dictionary*). Cependant, si l’on observe les fréquences comparées dans les corpus, on constate qu’un corpus français de 10 millions de mots donne 297 exemples d’*inciter* (si l’on inclut seulement le verbe et ses formes fléchies et non pas les dérivés tels qu’*incitation*)

alors qu'un corpus anglais de 60 millions de mots contre, ne fait apparaître que 112 exemples de *incite*. Pourtant, *incite* est souvent cité comme traduction générale de *inciter* : *Le Harraps Shorter* (HS) mentionne *to incite, to urge (on)* comme traduction générale – ce qui peut en outre laisser croire que *incite, urge et urge on* peuvent être employés indifféremment. *Le Collins Robert* (CR) donne la structure *inciter qqn à faire*, montrant ainsi que le complément d'objet est fréquemment humain et *to incite ou urge sb to do*. Le HS donne cette dernière structure après les traductions générales mais la traduit par *to incite ou prompt s.o to do something*, laissant ainsi croire au lecteur que ces traductions sont interchangeables. *Le Grand Dictionnaire Larousse* (GDL) propose *prompt* et *encourage* comme traductions de la structure *inciter qqn à faire qqch*, alors que *incite* est donné comme traduction seulement pour un autre sens du mot, appartenant au domaine juridique (*Jur*). Le CR propose *encourage* seulement dans le cas de deux exemples bien particuliers, et à partir desquels il est difficile d'extrapoler : *ça n'incite pas au travail* et *ça vous incite à la paresse*. Le HS mentionne encore un équivalent, *inclîne*, pour *incite* dans les deux exemples suivants : *cette réponse m'incite à la prudence* et *son attitude n'incite pas à l'aider*. Le GDL donne *encourage* et *prompt* comme équivalents dans l'exemple *son succès l'incita à continuer*, renversant ainsi l'ordre des équivalents tel qu'il avait été donné dans la structure précédente *inciter qqn à faire qqch* plus haut dans l'article.

Que comprendra l'utilisateur dans tout cela ? Les trois dictionnaires proposent quatre traductions générales pour le verbe *inciter*, et on ne peut pas dire qu'il y a consensus sur les distinctions d'emplois. Il va sans dire que le but de l'exercice n'est pas de critiquer les dictionnaires en question, mais plutôt de montrer la complexité du problème. Il est clair que l'utilisateur de langue source a peu de chances de trouver la traduction dont il a besoin.

On a critiqué la tendance du dictionnaire bilingue à «*présenter des lexèmes isolés alors qu'il est censé fonctionner pour des mots pris dans des textes individuels et dans des contextes variés.*» (Snell-Hornby, 1984). On s'attend à ce que le dictionnaire bilingue propose «*des unités lexicales réelles de la langue cible, qui, lorsqu'elles sont insérées dans le contexte, produisent une traduction qui semble naturelle*» (Zgusta : 1984). Si les équivalences proposées ne sont pas toujours utilisables dans tous les contextes, c'est parce que les lexicographes ne disposaient pas jusqu'à une date très récente de données satisfaisantes sur l'usage linguistique réel. Il se peut que le lexicographe de langue cible soupçonne que les traductions proposées pour *inciter* ne sont pas synonymes, et qu'il se sente en mesure de formuler des hypothèses basées sur les nuances de sens. Mais plusieurs lexicographes n'arriveront jamais à un consensus. Dans le passé, la solution était de rester dans le vague; on donnait donc plusieurs propositions en espérant que l'utilisateur saurait trouver celle qui lui convient. Les articles rédigés selon cette méthode ne peuvent satisfaire les besoins des utilisateurs de langue source qui s'aventurent dans le domaine relativement inconnu de la langue cible en espérant trouver des indications susceptibles de les mettre sur la bonne voie. Les lexico-

graphes étaient bien conscients de ce problème mais, ne disposant d'aucun moyen d'observer le comportement des quasi-synonymes dans le discours, ils ne pouvaient pas repérer les nuances de sens et les exposer sans risque de se tromper.

Au point où nous sommes arrivés dans le processus de transfert du verbe *inciter* vers l'anglais, il s'agit donc de trouver un moyen efficace de distinguer entre les différentes solutions proposées comme équivalents de traduction. L'analyse de *incite* dans le corpus de langue cible montre, comme nous l'avons vu, que le verbe est beaucoup moins utilisé que son équivalent français. On peut supposer qu'il y a entre eux soit une différence sémantique soit une différence de domaine d'utilisation. Si l'on examine le corpus anglais, il semble bien les deux phénomènes sont en cause. Le mot est utilisé dans les contextes juridiques, comme son équivalent français : on le trouve dans des formulations d'actes d'accusation (*inciting a riot, inciting arson*). Mais, contrairement à ce que laisse entendre le GDL, le mot n'est pas spécifique du contexte juridique; il semble, en fait, que son utilisation dans ces contextes ait eu des conséquences sur ses autres utilisations. Ainsi le verbe anglais *incite* dans des contextes non juridiques est utilisé presque exclusivement avec des noms et des verbes qui désignent des actions condamnables par la société, ce qui n'est pas du tout le cas pour *inciter*. Il n'est pas sans intérêt de noter qu'il s'agit là d'une évolution récente : jusqu'à la fin du dix-neuvième siècle, *incite* était aussi neutre que son équivalent français. Johnson le cite comme l'exemple même d'un mot «neutre» (Johnson, 1747).

Si l'on revient aux structures lexico-sémantiques d'*inciter* citées ci-dessus, il est clair que l'on peut proposer *incite* comme équivalent de traduction seulement pour un sous-ensemble des schémas 1 et 3 où l'ACTION ou l'ATTITUDE ou le V-INF est considéré comme «mauvais». L'analyse de *encourage*, cependant, montre qu'il peut être utilisé avec presque tous les schémas et les contextes couverts par *inciter* en français. Il y a 3246 occurrences du verbe *encourage*, 1987 du verbe *urge* et 1641 du verbe *prompt* dans le corpus de 60 millions de mots.

Un examen du contexte de *urge*, montre que le sujet est toujours une PERSONNE ou un GROUPE HUMAIN, alors que le sujet de *encourage* peut être un EVENEMENT, une CIRCONSTANCE, ou un PHENOMENE tout comme il peut être une PERSONNE. Il est également évident d'après les lignes de concordances que le verbe *urge* ne peut être utilisé que si la personne supposée accomplir l'action est en mesure de le faire. C'est pourquoi on utilise beaucoup *urge* dans des contextes comme les reportages sur les débats parlementaires, les discours officiels et les négociations. Mais ceci est également vrai d'*inciter*.

Les données sur le verbe *prompt* (1614 occurrences) sont très différentes. *Prompt* est fréquemment utilisé comme verbe permettant de rapporter au discours direct, mais dans ses autres emplois il fonctionne avec des sujets qui sont des EVENEMENTS ou des CIRCONSTANCES (ainsi qu'avec des compléments d'objet direct qui sont des PERSONNES ou des EVENEMENTS). *Prompt* s'oppose sur le premier point au verbe français *inciter*, qui n'est jamais utilisé pour rapporter du discours.

La consultation des deux corpus de textes nous a permis à la fois d'analyser la langue source et de comparer le comportement des différents équivalents proposés. Certes, il aurait été possible de demander au lexicographe de langue cible de traduire chaque citation du corpus de langue source qui contient *inciter* en l'invitant à se demander pourquoi telle traduction convient mieux que telle autre, en espérant qu'une certaine logique finirait par se dessiner. Cela demande beaucoup de temps et risque de donner des résultats très différents selon les lexicographes dans un cas comme *inciter* dont on se rend compte qu'il se révèle être un faux ami. Dans le cas d'*inciter*, il était indispensable d'utiliser les deux corpus pour pouvoir arriver à une analyse aussi fine des différences, mais fort heureusement tous les cas ne sont pas aussi complexes, et le lexicographe peut souvent se contenter du corpus en langue source, et n'avoir recours au corpus en langue cible que pour y vérifier ses hypothèses.

## 7 Synthèse

L'utilisation du corpus aux stades de l'analyse et du transfert permet d'essayer chaque traduction dans un contexte réel, et de tirer des conclusions générales à la fois sur la langue source et sur les équivalents de traduction, sans jamais perdre de vue l'utilisation en contexte.

«De nombreux spécialistes critiquent les équivalents proposés dans les dictionnaires bilingues du point de vue de la traduction des texte continus. Il y a une différence considérable entre l'équivalent qui convient parfaitement bien au contexte spécifique dont a besoin le traducteur et l'équivalent hors contexte que le dictionnaire bilingue se doit de proposer. Le rôle principal d'un dictionnaire est de ne pas induire l'utilisateur en erreur. Lorsque l'espace imparti ne permet pas de donner des informations métalinguistiques détaillées, le devoir du lexicographe est de proposer une traduction générale, dont l'utilisation ne peut pas être totalement fautive plutôt qu'une traduction qui serait parfaite dans certains contextes mais impossible dans d'autres.» (Atkins 1992/1993)

C'est ce souci de garantir l'applicabilité des équivalents qui guide le processus de *synthèse*. Nous avons vu dans le cas d'*inciter* que les équivalents de traduction proposés sont tous plus ou moins déterminés par le contexte. Lorsque les lexicographes ne disposaient pas de corpus leur permettant d'établir avec un degré raisonnable de certitude quelles sont les traductions qui peuvent fonctionner dans le plus grand nombre de contextes, il était particulièrement difficile de rédiger les articles des petits dictionnaires. La seule solution possible était d'omettre tout métalangage et de donner une suite de traductions en laissant l'utilisateur en tirer le meilleur parti possible. Comme nous l'avons montré, cette solution peut éventuellement se révéler suffisante pour l'utilisateur qui veut décoder, mais pas pour celui qui veut encoder. Cela est particulièrement dange-

reux car ce sont souvent les débutants qui utilisent les plus petits dictionnaires. L'utilisation d'un corpus permet de proposer un nombre plus restreint d'équivalences de traduction mais une quantité plus grande d'informations sur l'utilisation de chacune. Ainsi nous avons vu avec l'analyse des équivalences d'*inciter* que l'utilisateur de langue source peut utiliser sans risque *encourage* comme équivalent dans presque tous les cas «neutres». Il serait donc bien plus efficace dans un tel dictionnaire de ne proposer que *encourage* comme traduction directe, et de donner *incite* avec les collocations *violence*, *haine* pour indiquer les restrictions du contexte.

Dans les dictionnaires plus volumineux, il est plus facile de donner plus d'informations sur les différences entre les différents équivalents de traduction. Les lexicographes de langue source et de langue cible doivent se mettre d'accord sur la meilleure façon d'établir les généralisations les plus aptes à répondre aux besoins des utilisateurs, à la fois pour l'encodage et pour le décodage. C'est le lexicographe de langue source qui doit finalement décider des lexèmes qui doivent être conservés en termes de fréquence et de typicalité. Le lexicographe de langue cible peut intervenir pour inclure un usage relativement obscur qui pose des problèmes au décodage, mais si l'analyse du corpus montre que l'utilisateur de langue cible n'a pas plus de chance de rencontrer cet usage que d'autres usages similaires et peu courants qui ont été écartés, il n'est alors pas nécessaire de l'inclure dans le dictionnaire. Les traducteurs disent, non sans raison, qu'ils ne trouvent jamais les usages «rares» qu'ils recherchent dans les dictionnaires. Ceci résulte peut-être d'un malentendu sur la nature du dictionnaire bilingue, et sur ce qu'on peut en attendre. Même la perspective de disposer d'un espace quasiment illimité dans les dictionnaires électroniques n'est qu'une chimère de ce point de vue, puisque les particularités du contexte sont elles aussi quasiment infinies et aucun traducteur ne peut sérieusement envisager la perspective de parcourir des centaines de lignes de concordances pour essayer de trouver la solution qui correspond le mieux à son contexte. On peut espérer, néanmoins, que le traducteur qui a besoin d'une traduction pour un contexte bien précis aura dorénavant plus de chances de trouver ce qu'il cherche, puisque c'est précisément ce que l'utilisation du corpus permet de faire en proposant, lors de la phase de synthèse, des généralisations qui sont basées sur l'observation de l'usage.

Un dictionnaire construit à partir d'un corpus peut donc proposer *encourage* avec certitude comme traduction de *inciter*, lorsque le sujet est une personne (ou un groupe de personnes), une situation ou une attitude humaine et *prompt* lorsque le sujet est un événement. Si l'on dispose d'assez de place, on peut aussi indiquer que *événement* est en général un *changement* ou une *décision*. Il s'agit donc de distinctions qui sont à la fois suffisamment précises pour être clairement comprises par l'utilisateur et suffisamment générales pour être utilisées dans un grand nombre de contextes. Le processus de transfert aura également fait apparaître des combinaisons lexico-syntaxiques particulières qui peuvent poser des problèmes pour la traduction. Encore une fois, c'est le lexicographe de langue source qui

décide de retenir ou de ne pas retenir ces combinaisons selon qu'elles représentent des schémas qui aident à la production ou qu'elles sont trop particulières pour être mentionnées. Dans le cas d'*inciter*, certains groupes sémantiques construits sur le schéma *inciter à + NP* sont difficiles à rendre en anglais. On peut donc proposer, comme le fait le OHFD, *la récession incite à la prudence : the recession is making people cautious* et *inciter à l'espoir/au découragement : to be hopeful/discouraging*. Ces traductions ne sont pas proposées comme étant liées à des contextes très spécifiques, mais plutôt comme illustrant les stratégies qui peuvent être déployées dans des situations qui sont – et nous le découvrons en étudiant le corpus – à la fois fréquentes et problématiques du point de vue de la manipulation syntaxique.

Si l'espace disponible le permet, les traductions du mot vedette peuvent être illustrées par des exemples et on peut même proposer des traductions moins générales en indiquant les restrictions d'usage. Ainsi *incite* peut être proposé soit dans un exemple soit avec des collocations qui montrent que l'usage du verbe en anglais se limite à l'incitation des actions condamnables par la société. Par ailleurs, une analyse plus poussée du corpus en langue source a montré que *inciter vivement* était relativement fréquent surtout dans des contextes "officiels". Vu les résultats de l'analyse du verbe *urge* en anglais (voir ci-dessus), nous pouvons donc proposer ce dernier comme traduction de l'expression en français.

Nous avons vu à quel point le corpus pouvait être utile pour faire apparaître les différences significatives dans les fréquences des équivalents entre la langue source et la langue cible et pour repérer les restrictions syntaxiques et sémantiques. Il y a beaucoup d'autres informations dans un article de dictionnaire bilingue, en particulier pour l'utilisateur qui a besoin d'encoder. Peut-on utiliser un adjectif indifféremment comme épithète ou comme attribut dans la langue cible ? Y a-t-il une différence de registre entre le mot en langue source et son équivalent en langue cible ? Un substantif comptable au pluriel doit-il être traduit par un substantif non comptable ? L'équivalence en langue cible peut-elle être utilisée à tous les temps du verbe en langue source ? Peut-elle être utilisée au passif ? etc. La consultation du corpus peut aider à répondre à ces questions de manière bien plus fiable que l'intuition du lexicographe même si celle-ci garde toute sa valeur dans la pratique quotidienne du métier.

L'accès à un corpus permet désormais de rassembler facilement toute une série d'informations dont les lexicographes bilingues ont besoin. Les éditeurs, qui craignaient que les lexicographes ne se perdent dans les méandres du corpus, qu'ils ne soient déconcertés et découragés par la masse de faits de langue mise soudainement à leur disposition, n'ont pas vu leurs craintes se réaliser. Ils constatent également qu'il y a des changements perceptibles dans la qualité et l'utilité des ouvrages qu'ils publient. Bien entendu, les lexicographes s'interrogent toujours autant sur ce qu'ils devraient retenir ou omettre. Mais grâce au corpus, ils peuvent désormais voir, analyser – et donc ordonner plus clairement – les informations qu'ils proposent aux utilisateurs. Ils peuvent également

vérifier leur intuition et s'assurer qu'ils n'oublient rien d'essentiel. Les utilisateurs de dictionnaires ont donc plus de chances de trouver – et plus facilement – ce qu'ils cherchent. L'analyse de corpus, qui fait de plus en plus partie des ressources de base mises à la disposition du lexicographe, a beaucoup amélioré une situation qui avait peu changé depuis Johnson :

«Lorsque j'ai considéré mon projet pour la première fois, j'ai trouvé notre discours abondant mais sans ordre, énergique mais sans règles : de tous côtés, il n'y avait qu'ambiguïtés à clarifier et confusions à éclaircir; il fallait choisir parmi un nombre infini, sans principe clair de sélection». (Johnson, 1755).



## Chapitre 8

# L'INFORMATISATION DE LA FABRICATION DES DICTIONNAIRES BILINGUES

Frank KNOWLES

### 1 Introduction

Les progrès accomplis depuis une dizaine d'années dans le domaine des technologies de la communication ont permis aux spécialistes de nombreuses disciplines, et notamment à ceux qui travaillent dans les secteurs de pointe, d'améliorer constamment leurs produits. Il semble que ces produits aient ainsi pu, presque sans interruption, devenir meilleurs et plus performants. On peut en fait affirmer que ces progrès ont déjà révolutionné de nombreux types de tâches, parmi lesquelles le travail intellectuel. Cette révolution, aussi surprenant que cela puisse paraître, n'est pas seulement un bref épisode mais une évolution qui s'auto-alimente, et que l'on peut s'attendre à voir continuer pendant un certain temps encore. On affirme parfois que ce qui s'est produit dans ce domaine n'est rien d'autre que l'automatisation de procédés anciens, une amélioration notable des capacités de traitement, de production et de qualité du travail induite par des besoins ergonomiques, toutes choses qui sont certes importantes mais n'annoncent pas forcément une nouvelle conception du travail, une nouvelle éthique ou une nouvelle maîtrise des conditions existentielles. Dire cela revient pratiquement à dire que les ordinateurs ont pu accomplir des tâches qui, pour des êtres humains, auraient été trop difficiles, auraient demandé trop d'attention, ou auraient pris trop de temps. Ceux qui disent cela pensent aussi souvent, mais sans le dire, qu'au fond peu de choses ont changé dans la qualité du travail et la façon de l'envisager. Il revient sans doute à ceux dont les tâches ont été radicalement transformées par les technologies de l'information de se demander comment ils répondraient à ces affirmations. La question n'est pas seulement rhétorique, elle a des rapports directs avec le contenu de ce chapitre.

Le dictionnaire figure parmi les produits, à la fois culturels et utilitaires, de notre société moderne qui ont été le plus profondément modifiés par la révolution des technologies de l'information. La plupart des utilisateurs de dictionnaires traditionnels, même ceux qui ont peu l'habitude de s'en servir, peuvent à leur

propos non seulement dire à quel point ils en sont parfois déçus, mais exprimer des critiques révélatrices. La plus remarquable, celle qui fait le plus mal, est l'expression d'une aversion à l'utilisation du dictionnaire due à la crainte que l'effort de la consultation ne soit pas récompensé : les utilisateurs, au lieu d'être conduits efficacement et directement vers l'endroit où se trouve le mot qu'ils recherchent, sont attirés sur de fausses pistes, à tel point que leur énergie et leur enthousiasme s'y épuisent, laissant ainsi place, dans leur souvenir, à des impressions négatives. Il est en fait assez difficile, et sans doute superflu, heureusement, de quantifier le sentiment de frustration qui naît à cette occasion chez l'utilisateur certain que ce qu'il cherche «doit être quelque part dans le dictionnaire» mais qui n'arrive pas à le trouver. La morale de cette histoire est que les compilateurs de dictionnaires informatiques ne peuvent pas se contenter de prévoir des voies d'accès directes pour l'utilisateur moyen tel qu'ils se le représentent : ils doivent également prendre du recul par rapport à leur travail et se poser, de façon plus fondamentale, les questions concernant l'ampleur, la nature et la diversité des tâches qui leur incombent.

Dès lors, on peut peut-être interpréter l'informatisation – en l'occurrence des dictionnaires bilingues – de deux façons différentes. Tout d'abord, et c'est le plus évident, cela peut signifier que l'on prend le texte d'un dictionnaire bilingue traditionnel existant en version papier et qu'on le convertit, par exemple en un produit CD-ROM. Cela est relativement simple à réaliser si une version électronique du texte du dictionnaire existe déjà. C'est généralement le cas de nos jours, pour la simple raison que les versions finales des dictionnaires sont souvent données aux imprimeurs sous cette forme. Il est important de distinguer entre le «texte de travail» et le texte à imprimer. Le texte de travail est plutôt une base de données qu'un texte : il contient toutes les informations lexicographiques stockées selon une hiérarchie de structures de champs et de catégories, mise au point pour rendre compte fidèlement des décisions préalablement prises par l'équipe éditoriale en matière de macrostructure et de microstructure du dictionnaire<sup>1</sup>. Cette hiérarchie est rendue complexe par les réseaux d'interdépendances, de liens multiples et réciproques qui existent dans le stock de mots-vedettes, découpés et «atomisés» pour leur traitement. Malgré cette atomisation, un article consacré à un mot n'est jamais une entité complètement indépendante des autres. Le texte de travail contient aussi, évidemment, des informations destinées explicitement à la gestion de la préparation du texte du dictionnaire. Les marqueurs désignant les divers champs contenus dans l'article du mot-vedette sont visibles constamment et il faut qu'ils le soient en raison de la complexité du travail de préparation. Ces informations et les données qui les motivent sont considérées comme strictement privées et destinées au seul personnel participant à la préparation du dictionnaire. Le texte à imprimer, au contraire, est un fichier texte d'un type

<sup>1</sup> Pour une description détaillée du travail d'une équipe de rédaction, voir Atkins, B.T.S., (1992).

particulier, contenant toute une série d'étiquettes nécessaires pour affecter aux différentes rubriques les fontes et mises en formes désirées. Par exemple, ces groupes d'étiquettes permettent de gérer la sélection et le changement du format de la page, de la colonne, de la ligne et d'activer des polices de caractères variées.

## **2 L'informatique dans la production d'un grand dictionnaire**

Si l'on examine la production d'un grand dictionnaire par une maison d'édition moderne, on est amené à constater que les dictionnaires sont des produits qui exigent des normes de qualité extrêmement élevées. Ceci explique et justifie les niveaux de coût et de complexité des projets qui aboutissent aujourd'hui et qui impliquent l'utilisation de moyens informatiques considérables. Cet effort en faveur de l'informatique a deux objectifs : avant tout, fournir le travail commandé, en respectant les critères les plus sévères de qualité; ensuite – et c'est très important – saisir l'occasion pour innover, augmenter l'efficacité et la productivité. Bien sûr, ceux qui conçoivent et mettent en œuvre les améliorations du travail par l'informatique doivent s'assurer que leurs efforts seront récompensés, si possible sans trop hypothéquer l'avenir. On peut comprendre que la façon la plus raisonnable de procéder soit de mettre au point des moyens génériques modulables tels que ceux auxquels fera nécessairement appel tout projet d'envergure dans le domaine de la lexicographie commerciale.

Le terme «moyens» regroupe tout un ensemble hétéroclite de logiciels et de matériels. Le matériel doit disposer d'une mémoire centrale importante, d'une mémoire de travail extensible, d'une capacité de stockage multiple et, par dessus tout, d'une grande rapidité de transfert de données. Les écrans de travail doivent être d'une grande qualité ergonomique et doivent pouvoir afficher sans déformations des textes utilisant un grand nombre de systèmes d'écritures alphabétiques et non-alphabétiques. Tout ceci a des conséquences pour les techniques de stockage des données : il semblerait que l'on ait un grand besoin du système qui a été proposé sous le nom d'«Unicode», même s'il est très gourmand en espace de stockage.

En ce qui concerne les logiciels, la liste des outils indispensables est assez conséquente. Ils correspondent aux trois stades du processus de fabrication d'un dictionnaire informatisé, pré-lexicographique, lexicographique et post-lexicographique. La période pré-lexicographique concerne principalement le corpus (voir Grundy, dans ce volume). A ce stade, on a besoin de logiciels permettant la saisie des données et leur préparation à l'emploi. La liste de tâches qui demandent des moyens logiciels à la fois fiables et adaptables dans leur utilisation, avec une intervention humaine ponctuelle, pourrait s'établir comme suit :

1. Mettre l'intégralité du texte sous une forme standard en ce qui concerne les majuscules/minuscules. Ou bien, autre solution, s'assurer que tous les algo-

rithmes permettant la production fonctionnent selon un principe tel que la distinction est annulée. Pour ce faire, on peut :

- a) changer toutes les minuscules en majuscules. Cette option est utilisée assez fréquemment;
  - b) changer toutes les majuscules en minuscules. Cette option est aussi assez souvent utilisée. L'inconvénient est que les noms propres sont alors «déformés» et deviennent impossibles à distinguer des autres membres des paradigmes d'homographes apparus au cours de la phase de transformation. Les acronymes sont encore plus gravement déformés, et constituent de réels problèmes pour l'analyseur.
  - c) changer toutes les majuscules en début de phrase, ou en position équivalente, en minuscules, en veillant à conserver les majuscules dans les noms propres, etc.
- 2 Supprimer tous les traits d'union introduits pour couper les mots en fin de ligne. Les traits d'union qui font partie intégrante de mots doivent être conservés, même si par hasard, ils se trouvaient en fin de ligne !
  - 3 Indiquer la division en paragraphes, même si les moyens utilisés ne sont pas les mêmes que dans le document source. Une des solutions communément utilisées est de laisser un interligne entre les paragraphes et de débiter chaque paragraphe avec un retrait. Dans la plupart des projets de dictionnaires de dimensions respectables ces indications sont presque automatiquement données par des étiquettes.
  - 4 Noter la ponctuation. Dans certains projets il importe, par exemple, de marquer la différence entre les guillemets qui ouvrent une citation et ceux qui la ferment. Il peut également être extrêmement utile de distinguer les points de fin de phrase des points utilisés dans les nombres décimaux, et les virgules ordinaires des virgules qui, en anglais, séparent les chiffres dans les grands nombres pour en faciliter la lecture.
  - 5 Remplacer les guillemets simples utilisés dans les citations en anglais par des guillemets doubles de façon à éviter toute confusion avec les apostrophes.
  - 6 Mettre au point si nécessaire un système qui permette d'identifier les noms propres, susceptibles de créer des problèmes pendant l'analyse. Une solution consiste à regrouper tous les noms propres sous une même catégorie appelée NOM PROPRE.
  - 7 Créer un système d'indexation fiable pour les formes qui se trouvent dans le corpus. Il est également nécessaire de noter l'emplacement de toutes les occurrences de chaque forme.
  - 8 Mettre au point une méthode adéquate pour regrouper les occurrences qui «vont ensemble», même si l'opération n'est pas une vraie lemmatisation. S'assurer que l'on peut retrouver et déduire la fonction de chaque élément d'un lemme. C'est un sujet que les lexicographes évitent délibérément depuis

trop longtemps; c'est pourquoi la codification et la présentation des informations qui se trouvent à droite du mot-vedette sont nécessairement et parfois très gravement faussées par rapport à la réalité observable.

- 9 Lier entre eux les éléments discontinus, comme dans *he ran a big bill up*.
- 10 Séparer toutes les concaténations orthographiques pour faire apparaître chaque partie constituante. Par exemple *prisonservice – prison service*.
- 11 Repérer un par un tous les «intrus» provenant de langues étrangères et dont la fréquence est faible ou douteuse.
- 12 Repérer toutes les entités numériques et les formules, et préparer tous les outils graphiques pour qu'ils deviennent utilisables comme éléments du corpus.
- 13 Repérer tous les éléments lexicaux sans contexte syntaxique<sup>2</sup>.

Etant donné ce qui vient d'être dit ci-dessus, notamment au point 6, il convient d'approfondir la réflexion sur la conception et l'enchâssement dans le texte, lu de gauche à droite, des différents «marqueurs» ou étiquettes destinés à faciliter le traitement ultérieur. On peut se demander si l'on a besoin de deux versions du même corpus : une version de «surface», que l'on peut «voir» (dans les deux sens du terme) comme la version fondamentale et originelle, et une version de «travail», dont la configuration et le format correspondent aux besoins des lexicographes. Il serait extrêmement contraignant du point de vue logistique de conserver deux corpus en parallèle, et c'est pourquoi, au lieu d'essayer de travailler avec un corpus préparé on note les indications de préparation dans la base de données issue du corpus. Il se peut que le corpus conserve des étiquettes enchâssées, mais celles-ci n'affectent pas son intégrité fondamentale puisque le logiciel d'analyse grammaticale préserve les traces de l'environnement et peut faire la différence entre ce qui est d'ordre linguistique et ce qui est d'ordre métalinguistique. Si ces étiquettes gênent la lecture, elles peuvent être cachées facilement par une commande du logiciel.

Il n'est pas nécessaire que nous nous attardions sur les logiciels utilisés au cours du stade post-lexicographique : on a surtout besoin des moyens logiciels et matériels pour la gestion générale du corpus lorsqu'il est archivé plutôt que consulté, ce qui se résume à des opérations de sauvegarde et de maintenance. L'essentiel de l'activité lexicographique nécessite un soutien logiciel de tous les instants, ce qui souligne le besoin de programmes souples et adaptables ainsi que de méthodes permettant de mettre en place aisément une synergie entre deux logiciels ou plus. Si tous les outils ont été mis au point de manière interne et dans le but d'assurer un fonctionnement optimal du point de vue de la modularité, il faut postuler une interface, un environnement et un macro-langage communs à

---

2 Il est évident que ces recommandations sont basées sur une expérience de travail sur des corpus écrits anglais, et qu'elles doivent être modifiées de façon à prendre en compte la nature particulière de textes écrits en d'autres langues.

tous les logiciels, permettant en particulier la résolution des conflits. Il est clair que les meilleures conditions ne seraient pas réunies si l'on n'utilisait pas toutes les potentialités logicielles d'affichage sur écran, l'échange dynamique de données («dynamic data exchange» ou DDE), la fonction d'orientation, de liaison et d'enchâssement de l'objet («object orientation» ou OO, «object linking and embedding» ou OLE), et l'interface graphique («graphical user interface» ou GUI). Ces derniers permettent non seulement de mettre en place une interface «invisible» entre les fonctions utilitaires et le logiciel pour le corpus et la gestion de la base de données.

Il est bien connu que le développement de l'informatique a considérablement augmenté les exigences du consommateur moyen. Ces exigences ne sont pas toujours réalistes mais elles représentent une force qui fait évoluer le marché, et elles joueront sans aucun doute un rôle important dans les progrès techniques futurs et dans la demande des consommateurs de biens domestiques de grande technicité. On peut expliquer le phénomène comme suit : si les fabricants des produits à base de logiciels ne transmettent pas au consommateur toutes les potentialités des produits, celui-ci se sent trahi et il a tout à fait raison. C'est alors qu'un fabriquant trouve les moyens de faire mieux que ses concurrents et que le marché s'améliore. Que se passerait-il si ce scénario se produisait dans le petit monde de la fabrication des dictionnaires ? Certains éditeurs de dictionnaires utilisent actuellement des corpus très sophistiqués et des systèmes de bases de données qui produisent des dictionnaires de très haute qualité. Beaucoup d'acheteurs seront satisfaits de la qualité des dictionnaires produits de cette façon, mais d'autres pourraient réagir différemment, surtout s'ils constatent que l'utilisation de ces dictionnaires en version CD-ROM ne présente pas toutes les options possibles ou souhaitables. Les acheteurs de dictionnaires pourraient formuler ainsi leurs revendications : «Pourquoi nous empêcher de faire ce que vous savez faire ? Vous nous faites acheter un produit limité. Nous voulons les mêmes possibilités que vous !».

Continuons sur cette voie. Une base de données lexicographique bien organisée et bien gérée représente une ressource puissante qui peut-être configurée de différentes façons pour différents affichages à l'écran. C'est ce même principe qui justifie la présence du même dictionnaire sur le marché en version papier et sous forme de produit électronique. Fondamentalement, on peut considérer que les deux versions du même dictionnaire, la version papier et la version électronique, sont identiques. Cela peut avoir deux origines, toutes deux simples à comprendre mais néanmoins regrettables. Dans le pire des cas, le CD-ROM se contente d'être un support de données, et le texte qui apparaît sur l'écran est le même que celui du dictionnaire papier. Il se peut que la seule différence soit un système de recherche plus simple. L'investissement du fabricant est minimal, tout comme le sont les bénéfices qu'en retire l'utilisateur. C'est ce qui arrive lorsque la (pseudo)-informatisation du dictionnaire a été faite pendant la phase post-lexicographique, alors qu'il n'était plus possible d'améliorer sensiblement,

sur le plan qualitatif ou même ergonomique, un produit construit dans un moule traditionnel et prisonnier de ce moule. Mais lorsque l'informatisation du produit a lieu pendant la phase lexicographique à proprement parler, qu'obtient-on, ou que peut-on obtenir, en termes d'amélioration de la qualité ? Il est clair qu'on a alors bien plus de chances de créer des chemins d'accès «globaux» et de permettre une navigation ainsi que des possibilités de tri et de regroupement beaucoup plus perfectionnées. Le produit devient immédiatement un dictionnaire en hypertexte. Les fabricants peuvent même introduire le son sans trop de difficultés. Il est non seulement possible mais raisonnable d'introduire des concordances de chaque élément de la langue source avec les rubriques appropriées dans la partie droite du dictionnaire. Ceci constitue une indiscutable amélioration du produit CD-ROM par rapport à la version papier, tant sur le plan qualitatif que sur le plan ergonomique. Les possibilités d'informatisation du dictionnaire sont-elles maintenant épuisées, ou du moins sont-elles devenues moins intéressantes par la loi des rendements décroissants en matière de qualité ?

### **3 La fabrication d'un dictionnaire bilingue d'encodage**

On se doit de répondre par un «Non» catégorique à ces deux questions. Répond-on mal aux demandes des consommateurs en matière de qualité lorsque l'informatisation d'un dictionnaire bilingue ou unilingue se réduit à adapter un travail lexicographique traditionnel sur version papier au cours de sa phase principale d'élaboration ? Les éditeurs de dictionnaires qui se contenteraient de cette solution passeraient à côté de l'occasion d'atteindre des objectifs importants, dont l'un est de satisfaire une exigence bien connue et intellectuellement reconnue : celle de fournir des dictionnaires offrant des orientations différentes selon les utilisations : a-t-on besoin d'un dictionnaire pour le décodage ou pour l'encodage ? Sera-t-il utilisé par des lecteurs de langue maternelle ou de langue étrangère ? La satisfaction de ces demandes est plus avancée dans le domaine des dictionnaires unilingues, mais elle est loin d'être totale, que le dictionnaire soit conçu pour le décodage ou l'encodage, qu'il vise des utilisateurs qui parlent la même langue que celle décrite dans le dictionnaire ou non. On continue à produire – selon des rythmes de production beaucoup trop lents – des dictionnaires bilingues traditionnels en version papier, et (bien entendu !) en deux volumes, pour les deux types de clients qui ne font pas forcément bon ménage : ceux qui veulent décoder à partir de la langue source, et ceux qui veulent encoder en langue cible. Il est inutile de souligner les différences de conception et de fabrication. Dans ce domaine, l'informatisation permet sans aucun doute d'envisager une bonne rentabilité des efforts et des investissements. Dans les lignes qui suivent, nous nous concentrerons sur la tâche la plus complexe du lexicographe qui travaille sur un dictionnaire bilingue, c'est-à-dire la conception et la fabrication d'un dictionnaire destiné à des utilisateurs qui veulent encoder en langue cible.

Commençons par des choses toute simples : pour envisager la production d'un nouveau dictionnaire, il est judicieux de poser certaines questions.

- A Quelle est l'origine des mots-vedettes et des autres données lexicographiques retenues, et sur quoi repose leur choix ?
- B Le dictionnaire est-il tiré d'une base de données lexicales existante et constamment renouvelée ?
- C1 Le dictionnaire résulte-t-il de l'analyse d'un corpus réputé à la fois pour sa qualité et pour sa taille ? Pour un dictionnaire bilingue, la question C1 doit être formulée comme suit :
- C2 Le dictionnaire résulte-t-il d'analyses coordonnées de corpus en langue source et en langue cible réputés à la fois pour leur qualité et pour leur taille ? Si tel n'est pas le cas, les utilisateurs ne pourront avoir aucune garantie raisonnable quant à certains aspects importants, classés ci-dessous par importance décroissante.
  - 1 Les mots-vedettes ont-ils été retenus sur la base de données empiriques, c'est-à-dire statistiques et distributionnelles ?
  - 2 Comment ont-ils été collectés et systématisés ? (la présentation des mots-vedettes devrait être assez claire, mais il se peut que certains mystères restent à élucider).
  - 3 Le traitement lexicographique standard des mots-vedettes convient-il aux spécifications du dictionnaire bilingue ?
  - 4 Le traitement du vocabulaire cité ou illustré en langue source/langue cible rend-il compte de manière exhaustive de son potentiel combinatoire ?
  - 5 Quel est le pourcentage d'exemples et d'illustrations qui ne sont pas tirés directement du corpus ?
  - 6 Le dictionnaire reflète-t-il de manière satisfaisante, par le choix des illustrations et des exemples authentiques, c'est-à-dire tirés directement du corpus, les différences lexicales, grammaticales et structurales entre les deux langues concernées ?
  - 7 Comment les «intrusions» encyclopédiques sont-elles traitées ?
  - 8 Quelle est la qualité, la nature et l'étendue des feuilles liminaires insérées au début et à la fin du dictionnaire ?
  - 9 Y a-t'il des prises de position implicites dans le texte du dictionnaire lui-même et quel est le ton par lequel l'«auteur» fait entendre sa voix ?

Plusieurs positions et/ou hypothèses de départ découlent tout naturellement des interrogations ci-dessus :

- Un corpus est indispensable, parce que lui seul peut garantir la qualité d'un dictionnaire; s'il n'y a pas de corpus pour générer la nomenclature, le dictionnaire ne pourra qu'être le résultat d'une distillation de vocabulaire authentique et écotypique, à cause de la part d'arbitraire dans la sélection.
- Un corpus doit être informatisé et doit être accessible en-ligne pour un traitement à l'aide de logiciels d'analyse et de gestion de données textuelles.

- Quel que soit le degré de sophistication informatique mobilisé au cours de la phase lexicographique principale, le succès de l'informatisation de la fabrication d'un dictionnaire dépend de la mobilisation de corpus informatisés dès la phase «pré-lexicographique».
- Lorsqu'il s'agit de la fabrication d'un dictionnaire bilingue, l'acquisition du corpus, si elle est menée jusqu'à sa fin logique, est un double problème, puisque les citations issues du corpus doivent former une proportion très élevée des illustrations et des exemples, surtout dans un dictionnaire d'encodage en langue cible.
  - a Il faut utiliser deux corpus : un pour la langue source et un pour la langue cible.
  - b Ces deux corpus doivent être «harmonisés» et rendus compatibles, d'une façon ou d'une autre.
- Lorsqu'il s'agit de la fabrication d'un dictionnaire bilingue, il faut parvenir à un degré maximum de compatibilité et d'harmonie en ce qui concerne la nomenclature métalinguistique générale et les taxonomies utilisées pour chaque langue. Les textes qui se trouvent au début et à la fin du dictionnaire doivent être parfaitement équivalents dans leurs deux versions, L1 et L2.

#### **4 L'utilisation de deux corpus**

Venons-en au problème tout à fait particulier de l'association de deux corpus. Nous partons de l'hypothèse que dans la fabrication d'un dictionnaire bilingue, le corpus n'est généralement pas destiné à être utilisé pour la rédaction de la partie langue cible (mais voir Grundy, dans ce volume). On ne voit pas très bien comment il pourrait l'être, étant donné que dans cette situation des critères et des jugements de valeurs sont nécessaires pour la mise au point des procédures. Une conséquence de cela est que le corpus en langue cible ne doit en aucun cas être conçu, ou perçu, comme accessoire par rapport au corpus source, ou dépendant de lui. Si c'était le cas, la confiance dans le corpus cible s'en trouverait affectée. Le corpus en langue cible doit être autonome et placé sur un pied d'égalité avec le corpus source. Même si l'on envisageait de faire traduire massivement le corpus source, cela ne garantirait pas l'autonomie du texte, pour la simple raison que les conditions de la traduction seraient totalement différentes de celles dans lesquelles le texte source aurait été produit. La communauté émettrice du discours – selon la terminologie de Swales<sup>3</sup> – serait composée de traducteurs qui malgré tout leur professionnalisme ne pourraient produire que des substituts imparfaits des textes authentiques créés par leurs véritables «acteurs».

---

3 Swales, J. M., *Genre Analysis* (1990).

Il subsiste une incohérence non résolue dans tout cela. Les mots-vedettes des dictionnaires unilingues représentent des significations potentielles, souvent concrétisées par des segments de textes dans lesquels ils sont mis en œuvre à l'intérieur de micro-contextes. Ces segments de textes ont pour objectif de rassurer l'utilisateur du dictionnaire en fournissant des significations concrétisées du lexème, en l'absence quasi-totale des contextes qui les génèrent et les justifient ! Le même phénomène se produit, *mutatis mutandis*, dans les dictionnaires bilingues. Les équivalents, qu'il y en ait un seul ou plusieurs, sont présentés de la même manière acontextuelle, mais ils sont souvent ensuite désambiguïsés par des micro-contextes centrés sur différentes suites syntagmatiques proposées comme des équivalents réels dans ces contextes, ces équivalences étant justifiées par les attestations. Il est pratique, quoique peu scientifique, de se contenter de présenter l'information sous une forme compacte et par « blocs » discrets. Le travail lexicographique est chargé de résoudre ces problèmes, mais il laisse souvent l'utilisateur se débrouiller seul pour la traduction d'un mot. La seule solution est une option qui n'a pas été totalement comprise par les détenteurs de corpus et de bases de données destinés à la fabrication de dictionnaires bilingues informatisés : il s'agit d'offrir à l'utilisateur la possibilité d'appeler sur l'écran un certain nombre – déterminé par l'utilisateur lui-même – d'extraits du corpus, pour l'aider à résoudre son problème, et éventuellement apprendre d'autres choses. Il est tout à fait étrange que cette option ne soit pas encore disponible sur le marché, qu'il s'agisse de lexicographie monolingue ou bilingue. Il est évident que personne, parmi les lexicographes et les éditeurs de dictionnaires, n'a compris (ou alors faut-il imaginer qu'ils refusent de comprendre ?) deux choses. Tout d'abord, que les citations des concordances sont en général très éloquentes en elles-mêmes et peuvent être utilisées avec ou sans intermédiaires. Ensuite, que le matériau lexical, lorsqu'il est présenté directement, c'est-à-dire dans un (con)texte, peut très souvent se passer de toute explication fournie par un tiers.

Cette discussion n'a pas pour but de montrer qu'il est impossible de mettre au point un corpus satisfaisant en langue cible, mais vise plutôt à faire apparaître les difficultés et les risques dont il est important de réduire les conséquences. S'agissant des dictionnaires bilingues, il est permis d'espérer encore que l'examen attentif des unités de traduction potentielles dans les textes sources anglais et des équivalents choisis, pour chaque cas, dans les textes cibles en français du *Hansard* (annales du Parlement canadien) fournira une base à la fois spécifique et généralisable pour l'étude contrastive des deux langues. Ces étapes sont indispensables d'un point de vue pragmatique, et il peut même être nécessaire de changer un corpus conçu pour fonctionner en langue cible en un corpus source pour la réalisation de la deuxième partie du dictionnaire bilingue.

Une autre question qui n'a cessé de causer des difficultés pour les spécialistes de l'analyse des corpus est celle de l'homogénéité. Tout d'abord, qu'appelle-t-on homogénéité, et ensuite comment la mesure-t-on ? On peut voir l'homogénéité comme le résultat de deux facteurs : d'abord, une certaine constance dans les

références encyclopédiques et les programmes thématiques, et ensuite une constance dans la sélection des différentes ressources linguistiques et dans les méthodes utilisées pour en tirer des énoncés caractéristiques. Cela n'a rien à voir ni avec la «langue», prototype universel qui occupe une position inaccessible à une extrémité du continuum, ni avec les caractéristiques linguistiques individuelles, même s'il s'agit d'individus célèbres, mais avec les habitudes et les tendances écotypiques de communautés linguistiques que l'on peut identifier, dénombrer et décrire, et qui fonctionnent sur la base d'intérêts et de préoccupations communes. La conséquence est que la constance dans l'usage lexical généré par des références encyclopédiques s'observe plus facilement, apparaît plus nettement et se révèle plus stable dans ce que l'on appelle les technoclectes, ou les sociolectes professionnels. La collecte des lexiques utilisés dans les discours des communautés dont nous venons de parler ne permet pas d'aboutir à des groupements lexicaux plus conséquents, plus clairs et plus stables, mais à un estompement des frontières et à un flou typique de communautés linguistiques dont la population est à la fois plus nombreuse et plus diversifiée. Il devient possible de discerner des tranches lexicales en fonction de la fréquence, et ces tranches font preuve d'une certaine stabilité.

## **5 Le dictionnaire Hachette - Oxford**

Tous ceux qui s'intéressent à ces questions doivent prendre leur mal en patience et attendre les premiers résultats détaillés de l'expérience tirée du travail sur les deux corpus de 10 millions de mots chacun mis au point pour le *Hachette-Oxford*<sup>4</sup>. Cet ouvrage impressionnant, qui revendique plus de 350 000 mots et expressions et plus de 520 000 équivalences de traduction, s'impose à notre attention.

Les renseignements les plus fiables dont nous disposons actuellement sont tirées de la préface du dictionnaire et de son introduction, et de documents photocopiés distribués par Oxford University Press au moment du lancement du dictionnaire. Ces documents sont clairs : le dictionnaire a été conçu pour répondre aux deux types de clientèle évoqués ci-dessus, et pour satisfaire leurs besoins dans les deux «directions».

«Les rédacteurs ont, pour la première fois dans l'histoire de la lexicographie bilingue, utilisé un corpus électronique composé de deux bases de données textuelles de français et d'anglais. Les bases de données française et anglaise comportent chacune plus de 10 millions de mots de la langue usuelle. L'utilisation de ces bases de données a permis un repérage des mots et des traductions, toujours authentique et pertinent. Les utilisateurs peuvent en toute confiance se

---

4 Corréard, M.-H. et Grundy, V., *Hachette-Oxford French Dictionary French-English English-French* (1994).

servir des traductions proposées : elles sont issues de la langue réelle et employées en contexte. Le dictionnaire recense et traite de nombreux mots nouveaux, des expressions familières, des expressions idiomatiques et des tournures actuelles du français et de l'anglais... Les éditeurs se sont fixé un seul but : allier recherche linguistique et effort éditorial afin d'offrir aux traducteurs, étudiants de tous niveaux, enseignants et au monde des affaires, un ouvrage de référence inégalé à ce jour en matière d'autorité, d'actualité et de fiabilité.»

Une première réaction consiste à se demander de quand date le corpus, pour pouvoir juger si notre conception du terme «usuel» correspond à celle des lexicographes. Il convient également de s'attarder un peu plus longuement sur le passage suivant : «l'utilisation de ces bases de données a permis un repérage des mots et des traductions, toujours authentiques et pertinents». Cette formulation trompeuse est heureusement désambiguïsée dans la phrase suivante par la remarque : «Les traductions proposées sont issues de la langue réelle et employées en contexte.» Personne, assurément, ne contesterait que les équivalences françaises sont authentiques. Elles sont peut-être bien, d'une certaine façon, des traductions, mais elles ne sont pas des équivalences stables, inaltérables et garanties : elles sont des indicateurs de sens en langue cible et par conséquent, toute concordance les réunissant ne pourra jamais être qu'une concordance réunissant des indicateurs de sens présentés dans les contextes divers. Avant leur introduction dans un texte, il ne s'agit que de traductions potentielles. L'utilisateur du dictionnaire doit donc rester prudent, malgré les efforts louables consentis par les lexicographes.

«... Le choix des traductions a fait l'objet d'un soin particulier. En règle générale, à chaque acception d'un terme correspond une seule traduction. Cette traduction a été vérifiée dans les très nombreux exemples de langue réelle tirés de nos corpus électroniques. On ne trouvera deux ou plusieurs traductions équivalentes que dans les rares cas où elles sont vraiment interchangeables. Nous avons délibérément écarté les synonymes douteux et avons illustré par des exemples les traductions qui ne peuvent convenir que dans certains contextes».

L'investissement énorme en capital financier (3,5 millions de livres) et humain (environ 200 personnes/an) a indiscutablement permis la production d'un dictionnaire de grande qualité, dont la publication n'aurait pas été possible sans l'intervention massive de l'informatique. Les détails de cette intervention sont décrits dans les documents photocopiés mentionnés ci-dessus :

« L'ordinateur a joué trois rôles principaux dans ce projet :

- \* Il a fourni le matériel lexicographique (issu des bases de données) dont le texte du dictionnaire s'inspire.
- \* Il a permis de stocker la version finale et étiquetée du dictionnaire.
- \* Il a permis une planification et un contrôle efficaces du projet malgré toute sa complexité, grâce à un système de «suivi» construit par OUP Reference Computing. Les données dans ce système comprenaient approximativement

40 000 mots-vedettes dans chaque langue, rassemblés en groupes morphologiques et classés selon leur complexité lexicographique. Le système a mis au point des «tâches individualisées» pour ceux qui travaillent dans le cadre des structures pré-définies<sup>5</sup>, les traducteurs et les éditeurs, en fonction des spécifications sur la taille, la quantité et la complexité données par les gestionnaires du projet et il a permis de suivre chaque entrée lexicale pendant toute la durée du projet. Cela a permis aux gestionnaires de rester souples dans leur contrôle, de distribuer le travail aux compilateurs selon leurs capacités et leurs méthodes de travail et de s'assurer que la longueur du texte compilé correspondait bien aux objectifs qu'ils s'étaient fixés. Ce système de suivi a également permis de produire des rapports sous diverses formes sur la progression du travail.»

Par conséquent, l'outil informatique a participé sous trois formes différentes à la création de ce dictionnaire, et cela montre bien l'éventail des possibilités offerte aux créateurs de dictionnaires :

- 1 la production rapide d'exemples contextualisés d'acceptions faisant partie du groupe du mot-vedette;
- 2 la diminution des problèmes logistiques d'exactitude et d'intégrité du texte du dictionnaire au fur et à mesure qu'il grandit;
- 3 un système souple d'aide à la gestion qui permet de répartir les tâches, d'assurer le maximum de synergie et de réduire le temps perdu.

## **6 Les collocations dans un dictionnaire bilingue**

Les collocations constituent également un chapitre important de toute discussion sur l'informatisation de la lexicographie bilingue (voir Roberts, dans ce volume). Le phénomène de la collocation se manifeste par la juxtaposition d'éléments lexicaux régulièrement associés. Les collocations sont parfois basées sur l'assonance, parfois sur l'effet prosodique et parfois sur des critères moins aisément identifiables. En général, elles n'ont ni rime, ni raison ! Quelle que soit leur origine, on peut montrer qu'elles sont statistiquement significatives en termes de régularité et de distribution. Le besoin est constant, et urgent, d'ouvrages de références et d'outils pédagogiques pour aider ceux qui n'ont pas eu la chance, depuis leur jeune âge, de rencontrer en nombre suffisant ces suites lexicales qui, même si elles sont souvent idiosyncratiques, sont néanmoins fréquentes. Le problème est que la «collocabilité» des éléments lexicaux est le fruit des structures cachées et subliminales ainsi que des rythmes d'une langue, et qu'il faut donc procéder à des descriptions non pas fondées sur des critères, quels qu'ils soient, mais plutôt sur des observations statistiques sur les occurrences et les co-occurrences.

Seuls les dictionnaires informatisés, qu'ils soient monolingues ou bilingues, et les corpus dont ils sont tirés, peuvent générer ce saut quantique dont le domaine

de la lexicographie des collocations a tant besoin. Ces deux types de dictionnaires traitent pour l'essentiel le même problème : quelqu'un essaie de produire un texte en langue cible et rencontre des difficultés soit parce qu'il ne trouve pas dans son expérience ce qui lui permettrait de faire le bon choix, soit parce que la logistique de la recherche dans le dictionnaire se révèle trop difficile (on ne lui propose pas de point de départ adéquat, et les chemins d'accès aux mots n'ont pas été convenablement préparés, signalés et mis en place. La destination est topographiquement inaccessible).

Où faut-il donc chercher si l'on a besoin d'aide en matière de collocations et comment cette aide est-elle fournie ? Tant que des progrès n'auront pas été accomplis dans le domaine de la lexicographie des collocations, nous n'aurons pas de meilleur point de départ que celui qui a bien été décrit par Hausmann<sup>6</sup> : la pratique du placement asystématique à l'intérieur d'un même dictionnaire et dans divers dictionnaires. Nous utiliserons la terminologie de Hausmann. L'un des exemples qu'il cite est *knotty problem*. Il affirme avec toute la vigueur nécessaire pour faire passer son message, qui est important, qu'il est vital à la fois sur le plan conceptuel et sur le plan lexicographique de distinguer entre la **base** et le *collocatif*, autrement dit, dans ce cas, respectivement **problem** et *knotty*. La base est la partie de la collocation qui fait partie des mots qui occupent une place privilégiée dans la hiérarchie traditionnellement utilisée pour les parties du discours : les verbes, les noms, les adjectifs, les adverbes. La base peut également être un concept exprimé du point de vue linguistique par son «nom», auquel cas il a évidemment plus d'importance que l'autre membre de la collocation dans le sens où il le précède ou le domine dans le processus de raisonnement qui mène à une formulation linguistique et à une énonciation<sup>7</sup>. Cette hiérarchie aboutit à trois combinaisons fréquentes et indiscutables : [adjectif + nom, *a tenuous grasp*], [adverbe + verbe, *he instinctively felt*], [adverbe + adjectif, *implacably hostile*]. Les autres combinaisons sont un peu plus complexes : [nom (comme sujet) + verbe, *the wind abated*], [verbe + nom (comme objet direct), *he quashed the verdict*] [nom «of» nom a *torrent of abuse*]. Hausmann cite l'expression *to meet a situation*<sup>8</sup> tirée du dictionnaire Wolf/Canavan et fait remarquer que c'est bien *situation* qui est la base, même si l'on a tendance implicitement à penser que c'est *to meet* qui joue ce rôle.

5 Dans ce cas précis, toutes les analyses sous-jacentes des systèmes linguistiques des deux langues intéressant le lexicographe ont été préparées par ceux qui travaillent sur les grilles «sous la forme d'entrées lexicales structurées pour qu'elles soient cohérentes et compatibles». Cette procédure au cours de laquelle de nombreux contrôles de qualité ont eu lieu a permis une certaine harmonisation des opérations lexicographiques et a donc évité tous risques de mésententes qui pourraient être issus de pratiques professionnelles et d'intuitions différentes.

6 Hausmann, F. J., «Collocations in Monolingual and Bilingual English Dictionaries» (1991). On y trouve une longue référence à Wolf et Canavan, J., *Dictionary of English Words in Context* (1978).

7 Ceci inclut les sujets traditionnels et les prédicats, le contraste logique entre le thème et le rhème, et même la notion de «sujet psychologique».

8 On dit plus naturellement et plus couramment *to encounter a situation* !

Kozłowska<sup>9</sup> parle de «collocations à sens unique» :

«[...] Nous avons ce que nous pourrions appeler des collocations «à sens unique», lorsque le mot d'entrée [= la base, FK] appartient à une classe supérieure (ici les verbes ou les adjectifs) et le cooccurrent appartient à une classe inférieure (ici les adverbes). *English Adverbial Collocations* ne donne que des collocations «à sens unique», puisque les autres, s'il y en a, sont vues comme n'étant pas d'une grande utilité, sauf peut-être dans le domaine de l'apprentissage de la langue, mais certainement pas pour la traduction. Un traducteur part toujours d'un mot qui illustre le contexte particulier qu'il a en tête. Par exemple, en ce qui concerne les collocations adverbiales, il commencera sans doute par rechercher dans un dictionnaire de collocations un mot d'une classe supérieure (par exemple un verbe) qui convient à son contexte, comme «*resent sth*». Il trouvera alors une série de mots de classe inférieure (des adverbes) comme *bitterly, deeply, keenly* etc. parmi lesquels il pourra faire son choix. La recherche serait infructueuse s'il commençait ses recherches à partir de mots de classes inférieures (adverbes) pour trouver des mots de classe supérieure (verbes). Ainsi, les collocations basées sur un adjectif, un verbe ou un adverbe comme

Calm	<i>sea/expression/approach</i> (adjectif nom)
Organize	<i>education/a meeting/a tour</i> etc. (verbe nom)
Rapidly	<i>deteriorate/disappear/increase</i> etc (adverbe verbe)

ne sont pratiquement d'aucune utilité, au moins pour un traducteur.»

La seule façon réaliste de représenter la voie d'accès empruntée par quelqu'un qui cherche à formuler une expression comme *a tenuous grasp* est de postuler que le mot *grasp* a été choisi en premier et qu'ensuite on a cherché à caractériser le substantif par un adjectif. Cela revient à se demander quels sont les adjectifs qui peuvent qualifier *grasp*. Il n'est pas raisonnable de se demander quels noms pourraient qualifier l'adjectif *tenuous*. La question pourrait intéresser un lexicologue mais pas quelqu'un qui est engagé dans un processus de production d'un texte. On peut appliquer le même raisonnement à tous les autres exemples ci-dessus. Il s'agit bien de qualification : ce n'est que lorsque le *qualificandum* a été identifié que l'on peut choisir le *qualifians* approprié.

Le postulat lexicographique qui en résulte est que c'est la base qui est indispensable dans tout dictionnaire qui aspire à proposer un traitement adéquat des collocations. Le rôle du collocatif est relativement secondaire : il n'est pas essentiel du point de vue du fonctionnement de la langue de savoir quels autres noms l'adjectif *tenuous* pourrait qualifier.

Les collocations constituent un problème considérable pour les lexicographes, mais l'envergure du problème peut être notablement réduite dans les dictionnaires informatisés. La relation entre la base et le collocatif étant de nature

9 Kozłowska, C.D., *English Adverbial Collocations* (1991).

syntagmatique, les dictionnaires doivent adopter le type de technique décrit ci-dessus, et permettre l'accès aux informations sur la nature et le fonctionnement des collocations au moyen de citations tirées directement du corpus de base. Ces citations peuvent être générées à la demande, et le problème de leur gestion et de leur stockage informatiques n'est donc que transitoire, quel que soit leur volume. Il est également possible, bien sûr, de s'en tenir à des comptages par fréquence, mais le contexte est alors complètement absent. L'exploration du fonctionnement des collocations est une tâche que seuls les ordinateurs peuvent mener à bien au niveau de qualité nécessaire. Les lexicographes et les lexicologues ont mis au point, pour leurs besoins professionnels propres, des méthodes de mise en évidence des collocations potentielles et des logiciels permettant de mettre en œuvre ces méthodes et de doter chaque collocation d'un coefficient d'utilité. Les collocations doivent également être envisagées du point de vue de leur fréquence et donc de leur probabilité d'apparition dans la langue générale ou dans les langues spécialisées. Ces moyens d'investigation doivent être plus facilement accessibles au public. Le problème qui sera sans doute le plus difficile à régler sera celui de la disponibilité d'un corpus de base adéquat<sup>10</sup>, mais il est raisonnable de penser que les dictionnaires informatisés ne doivent pas être séparés de leurs modules de corpus. Si ce cordon ombilical était coupé, la possibilité d'établir une interaction dynamique avec le dictionnaire informatique serait fortement compromise.

D'autre part, il ne faut pas perdre de vue certains détails pratiques liés au corpus. Il se peut qu'une phrase simple extraite d'un corpus ne soit pas une unité adéquate en tant qu'exemple lexicographique. Les informations supra-segmentales doivent également être consignées. Les concordances du mot français *ou*, qui sert à structurer l'argumentation, montrent bien ce phénomène. En outre, il faudrait noter les «positions typiques» en même temps que les «collocatifs typiques». La position (dans une phrase, dans un texte, par rapport à d'autres mots) est parfois le seul facteur, ou un des facteurs, permettant de distinguer des équivalents, comme par exemple *certainly* et *certes*. Le type de discours, le contexte ou les conventions culturelles (qu'il ne faut pas confondre avec les mentions de «registre» présentes dans beaucoup de dictionnaires) peuvent aider à la compréhension de certains mots et peuvent également expliquer pourquoi le dictionnaire propose parfois une expression qui a l'air longue et maladroite en guise de traduction. La notion de «corpus» rend ces éléments difficiles à distinguer. D'autres problèmes peuvent surgir lorsque la traduction en langue cible d'un mot de la langue source provient d'une tranche lexicale différente du point de vue de la fréquence, à cause par exemple d'une disparité (encyclopédique ?) des usages. En outre, les traducteurs optent souvent pour la reformulation syntaxique, pour de multiples raisons : de la simple maladresse d'expression à la volonté d'améliorer la structure du message, en passant par l'existence de différences dans les

10 Une initiative de ce type a conduit à la création du British National Corpus.

rôles fonctionnels des membres de la famille d'un lemme. Un autre problème est celui de la difficulté d'expression, ce qui devrait nous faire prendre conscience qu'il existe toujours des degrés d'équivalence dans les traductions réelles.

## 7 Le dictionnaire français – russe de Gak et Triomphe

Tournons-nous maintenant vers d'autres caractéristiques que l'on peut exiger de dictionnaires bilingues informatisés, tirées de l'observation de dictionnaires bilingues qui ne le sont pas encore. Le *Dictionnaire Français-Russe* de F. Gak et J. Triomphe<sup>11</sup>, ouvrage impressionnant et très complet (plus de 1000 pages d'une typographie serrée, 37 000 mots-vedettes et plus de 150 000 équivalences), est un dictionnaire unidirectionnel pour apprenants destiné aux francophones qui veulent apprendre à parler le russe. Les auteurs accordent une attention toute particulière à la nécessité d'inclure tout ce qui peut servir spécifiquement à résoudre les difficultés propres au russe en tant que langue cible : les accords, la déclinaison des noms propres, le concept animé/non animé, etc. Le dictionnaire regroupe également les mots selon des critères sémantiques généraux et selon leurs ressemblances du point de vue des structures collocationnelles, et s'attache en outre à décrire les façons de «transférer» en russe les spécificités de certaines catégories grammaticales du français. Pour aider les utilisateurs à franchir le pas qui sépare la langue de la parole, il utilise quatre méthodes.

«La transition entre *langue* et *parole* s'opère en quatre étapes : la première étape consiste à fournir une ou des équivalences de traduction qui tiennent compte, bien sûr, de toute polysémie non liée au contexte; l'étape suivante consiste à donner des exemples dans lesquels figurent une ou des traductions du mot-vedette russe; ensuite viennent des équivalences dans lesquelles ne figurent pas les équivalents principaux du mot-vedette russe; enfin l'article fournit des expressions courtes dans lesquelles le choix de l'équivalence est déterminé avant tout par des considérations contextuelles».

Il est intéressant de rappeler ici l'analyse pénétrante de John Sinclair, rédacteur en chef du *Cobuild Dictionary*, lors d'un congrès<sup>12</sup>. Sinclair affirmait avec force et conviction qu'il n'y a plus aucune raison de faire la différence entre langue et parole. Saussure avait fait la même remarque, mais seulement en se fondant sur l'impossibilité de décrire ce qu'était la langue à partir de données concrètes. Cette impossibilité est maintenant en passe de disparaître, affirme-t-on, grâce à la linguistique basée sur les corpus : la parole est représentée par les lignes de concordances et la langue apparaît dans la colonne centrale. Nous avons

11 Gak, V.G. et Triomphe, J., (1991). Une critique de ce dictionnaire a été publiée dans *International Journal of Lexicography* (1994). Un autre dictionnaire d'encodage mérite d'être mentionné : Neubert, A. et Gröger, E., (1988). Il est maintenant distribué par Langenscheidt.

12 Lors du symposium international sur la phraséologie à Leeds, en 1994.

déjà vu, surtout à propos de la lexicographie unilingue, quelles améliorations énormes l'utilisation de concordances pouvait apporter en ce qui concerne la qualité des résultats et des hypothèses sur le fonctionnement linguistique. Ceci est particulièrement évident dans le domaine du lexique, mais il faut admettre que les bases de données lexicales fondées sur des corpus et qui servent à la fabrication des meilleurs dictionnaires modernes représentent des ressources à la fois gigantesques et puissantes. Il ne faut pas oublier, en particulier, que les dictionnaires et les corpus qui leur donnent naissance contiennent une grande quantité d'informations grammaticales, qui méritent d'être étudiées pour elles-mêmes. La distinction quelque peu artificielle entre langue et parole peut également être réduite dans le domaine grammatical. De plus, de la même façon que les citations tirées du corpus font clairement apparaître les significations lexicales parce qu'elles sont une démonstration directe et immédiate, il devrait être possible d'atteindre un degré comparable de transparence et de liberté de navigation en ce qui concerne les phénomènes grammaticaux. Tout du moins, ce type d'approche devrait permettre de se débarrasser dans un avenir proche des formules maladroitement et incompréhensibles utilisées pour indiquer la valence des mots.

## 8 Conclusion

L'informatisation des dictionnaires bilingues était dans les limbes avant la publication du *Hachette-Oxford Dictionary*. Ce dictionnaire constitue désormais un modèle, qui sera difficile à égaler, et encore plus à surpasser. Nous l'avons dit : nous attendons maintenant avec impatience non seulement des rapports détaillés sur la fabrication de ce dictionnaire mais également l'apparition sur le marché de dictionnaires informatisés bilingues publiés par d'autres maisons d'édition. Surtout, beaucoup d'utilisateurs espèrent désormais voir paraître des dictionnaires informatisés bilingues ou monolingues distribués sous forme de CD-ROM qui se distingueront par une plus grande sophistication et une plus grande fonctionnalité. Tous ceux dont la profession bénéficiera de ces perfectionnements les attendent avec impatience, et le grand public lui-même mérite qu'on lui propose ces dictionnaires nouveaux.

## Chapitre 9

# LES DICTIONNAIRES BILINGUES EN-LIGNE ET LE POSTE DE TRAVAIL DU TRADUCTEUR

Elliott MACKLOVITCH

Le concept de «poste de travail du traducteur» que nous présentons dans ce chapitre est relativement récent. Il a fait l'objet d'une présentation générale dans le cadre des Iles Journées Scientifiques du Réseau Lexicologie, Terminologie et Traduction tenues à Mons (Belgique) en 1991 (Clas et Safar, 1992). Nous nous concentrerons ici sur un point spécifique : le rôle et la place des dictionnaires informatisés accessibles directement depuis un traitement de texte, maintenant appelés «dictionnaires en-ligne». Nous illustrerons nos propos par l'exemple de notre expérience au Centre d'Innovation en Technologies de l'Information (Laval, Québec).

### 1 Qu'est-ce que le poste de travail du traducteur ?

À notre sens, le poste de travail du traducteur, ou PTT tel qu'il est maintenant communément appelé, est beaucoup plus un concept qu'un produit fixe. L'idée de base est très simple; il s'agit de regrouper, sur la plate-forme relativement peu coûteuse qu'offre l'ordinateur personnel, un ensemble de logiciels qui peuvent aider le traducteur à exécuter de façon plus efficace un certain nombre de ses tâches de routine périphériques.

Essayons de préciser certains éléments de cette définition. D'abord, par tâches périphériques, nous entendons tout ce qui s'écarte de la tâche centrale du traducteur, laquelle consiste à reformuler le contenu du texte de départ en un texte correct et équivalent dans la langue d'arrivée. De ce point de vue, le PTT s'inscrit dans la catégorie des aides automatisées à la traduction humaine et non pas dans celle des véritables systèmes de traduction automatique, où la machine propose une traduction du texte de la langue de départ. Ainsi, parmi les tâches périphériques que le PTT viserait à automatiser, on pourrait inclure la recherche dans les archives de traductions passées, la comparaison de deux versions d'un même texte dans la langue de départ, pour repérer les mises à jour, par exemple, la consultation de glossaires, de dictionnaires, de bases de données terminologiques,

ou d'autres ouvrages de références, la transcription de la traduction dans la langue d'arrivée, la relecture finale pour relever les fautes de frappe et autres coquilles – bref, tout ce que les traducteurs peuvent faire entre la réception du texte à traduire et l'expédition de la traduction finale, à part l'acte traductionnel lui-même.

Autre point à souligner, il n'y pas un PTT, mais des PTT. Puisqu'il s'agit d'un concept, le PTT peut admettre plusieurs réalisations, regroupant divers produits logiciels sur différentes plate-formes matérielles. D'ailleurs, la flexibilité est un des avantages de cette approche : la composition du PTT peut varier selon la situation, les besoins, voire même les goûts des traducteurs.

Le Centre d'Innovation en Technologies de l'Information (CITI) a élaboré, pour le compte du Bureau de la traduction de l'État canadien, trois versions successives du PTT, adaptées à la situation particulière des traducteurs de cet organisme. Les spécifications fonctionnelles de ces postes ont été décrites ailleurs (Macklovitch, 1989, 1992, 1993). Nous insisterons seulement ici sur une de leurs caractéristiques importantes : l'utilisation presque exclusive de logiciels commerciaux. Même si, en principe, on peut envisager l'élaboration d'un PTT à partir de zéro, la démarche adoptée au CITI est de voir jusqu'où on peut aller avec des logiciels disponibles sur le marché. L'avantage évident de cette approche est qu'elle nous permet de mettre au point un système opérationnel beaucoup plus rapidement, et à un coût bien inférieur à celui d'un développement intégral<sup>1</sup>. Notre travail dans le projet PTT a donc d'abord consisté à trouver divers logiciels d'applications, dont la plupart n'avaient pas été conçus spécifiquement pour la traduction. Il a fallu aussi les évaluer afin de vérifier leur bonne intégration avec les autres composantes du poste. Chaque version du PTT fut ensuite transférée au Bureau de la traduction, où elle a fait l'objet d'un essai en bonne et due forme mené par des traducteurs en milieu opérationnel, la réaction des utilisateurs devant bien sûr influencer la forme de la version suivante.

## 2 Et les dictionnaires en ligne ?

On pourrait, pour résumer, dire qu'un poste de travail conçu spécialement pour les traducteurs devrait normalement aider ces derniers à accomplir les tâches courantes qui leur sont propres. Or, une des caractéristiques du travail des traducteurs est la fréquence avec laquelle ils consultent des ouvrages de référence, particulièrement les dictionnaires. Donc, un des objectifs du projet PTT est de faciliter l'accès à ces sources d'information lexicologique en intégrant autant de dictionnaires en-ligne que possible. Par l'expression «en-ligne», nous dési-

<sup>1</sup> Cette approche n'est pas sans inconvénients. Voir Macklovitch (1992) pour une discussion du problème de l'hétérogénéité qui peut résulter du regroupement de logiciels disparates.

gnons ici la version électronique d'un dictionnaire que le traducteur peut consulter directement sur son ordinateur, sans avoir à se déplacer ni à fermer les autres applications qu'il est en train d'utiliser. Ceci s'applique évidemment prioritairement aux logiciels de traitement de texte. Du reste, on voudrait idéalement que ce dictionnaire soit complètement intégré avec les autres applications du PTT; nous reviendrons sur ce point plus loin.

Et qu'entend-on ici par «dictionnaire»? Tout simplement, les mêmes ouvrages que ceux que les traducteurs ont l'habitude de consulter sous forme papier, à savoir, les meilleurs dictionnaires unilingues et bilingues disponibles pour une paire de langues donnée. Dans ce chapitre nous établirons une distinction entre, d'une part, les dictionnaires classiques et, d'autre part, deux autres ressources lexicographiques offertes sur le PTT : un programme de gestion de glossaires et une banque de terminologie. Le traducteur se sert d'un glossaire pour consigner sur des fiches électroniques les résultats de ses propres recherches terminologiques. Contrairement à ce qui se passe avec un dictionnaire standard, ce sont les utilisateurs eux-mêmes qui définissent et modifient le contenu de ces glossaires en fonction de leurs propres normes de traduction ou des préférences de leurs clients. Quant à la banque de terminologie, il est rare que ce soit un outil directement modifiable par les traducteurs. Mais, par rapport à un dictionnaire standard, son contenu est plus spécialisé et généralement limité au vocabulaire technique. Ainsi, une banque comme **TERMIUM** ne contient pas d'entrées pour les mots les plus généraux et les plus fréquents de la langue.

Dans un premier temps, nous nous limiterons aux dictionnaires classiques, c'est-à-dire à ces ouvrages de référence bien connus qui offrent un répertoire ordonné des unités lexicales d'une ou de plusieurs langues, proposant pour chaque mot une définition de chacun de ses sens ou un équivalent dans la langue cible, ainsi que d'autres informations sur sa prononciation, son étymologie, sa flexion, ses variantes orthographiques, sa catégorie syntaxique, les constructions typiques dans lesquelles il est employé, des exemples d'usage, etc.

Dans les parties 3 et 4 nous décrirons les différents dictionnaires qui ont été intégrés par le CITI dans les trois versions du PTT. Mais avant d'entrer dans les détails, il convient de s'interroger sur les avantages qu'offrent les dictionnaires en-ligne par rapport à leurs versions papier standard. Nous en voyons au moins trois.

Premièrement, les dictionnaires en-ligne sont plus faciles et plus rapides à consulter que les dictionnaires traditionnels. À partir de son traitement de texte, le traducteur peut passer au dictionnaire électronique et l'interroger en beaucoup moins de temps qu'il ne lui en faut pour trouver le même article dans son dictionnaire papier. Ce qui est plus important encore que les quelques secondes gagnées à la fin de la journée, c'est que cette facilité de consultation encourage le traducteur à vérifier son dictionnaire plus souvent, avec pour conséquence une amélioration de la qualité du texte final. Deuxièmement, les dictionnaires

électroniques installés sur le PTT permettent au traducteur d'importer le contenu des articles directement dans son traitement de texte, via l'intégrateur (soit *Windows*, soit *DesqView*) qui encadre les applications du poste – ce qui évite au traducteur d'avoir à recopier des citations ou des exemples d'usage attesté. Enfin on peut effectuer certains types de recherche dans un dictionnaire électronique qui ne sont pas possibles dans un dictionnaire imprimé traditionnel. En effet, certains dictionnaires électroniques indexent non seulement la vedette des articles, mais aussi certains (ou tous les) mots contenus dans le corps de l'article. L'*Oxford English Dictionary on CD-ROM*, par exemple, offre un index séparé pour les expressions ou les mots combinés qui sont inclus dans d'autres articles, ce qui est fort pratique quand l'utilisateur ne sait pas sous quelle vedette les chercher. Et ce n'est qu'un exemple d'un type de recherche simple. Dans ce même *OED*, on peut également chercher toutes les entrées qui contiennent une citation d'un auteur donné ou d'une des ses œuvres. Dans le *Robert électronique*, on peut, lorsqu'on n'est pas sûr de l'orthographe d'un mot, le taper tel qu'il est prononcé, et voir apparaître à l'écran le mot correctement épilé et tous ses homonymes. Tous ces types de recherche seraient impensables dans un dictionnaire papier. Donc, les dictionnaires en-ligne ne sont pas seulement une étape de plus vers l'avènement du bureau sans papier; ils offrent aux traducteurs de nouvelles façons de mieux tirer profit de toute l'information que recèlent leurs dictionnaires et qui reste trop souvent inexploitée.

### 3 Les dictionnaires monolingues

Étant donné que la traduction de l'anglais vers le français occupe la plus grande part du marché de la traduction au Canada, les efforts dans le projet PTT du CITI se sont concentrés sur cette paire de langues. On peut se demander si le traducteur francophone a réellement besoin de consulter un dictionnaire de la langue source. Les participants aux essais du PTT répondent clairement par l'affirmative, puisqu'ils trouvent ainsi le moyen de mieux cerner le sens d'un mot ou d'une expression dans le texte à traduire. Le PTT est également conçu pour des traducteurs anglophones, et nous n'avons pas hésité à inclure dans le poste plusieurs bons dictionnaires électroniques de l'anglais.

Nous n'en mentionnerons ici que deux, à titre indicatif, aux deux extrémités de l'échelle de prix. La société WordPerfect offre le *Random House Webster's* en version *Windows* pour un prix très raisonnable<sup>2</sup>. Dans ses 180 000 entrées, on retrouve les mêmes informations que dans la version papier du *Random House Webster's College Dictionary*, édition 1991; de plus, le logiciel comprend une version électronique du *Random House College Thesaurus*, un dictionnaire de

2 moins de 100 \$ US.

synonymes de 275 000 mots. Malgré le nombre impressionnant d'articles, il n'est pas évident que le contenu de ces éditions soit suffisamment étoffé pour les traducteurs professionnels, surtout pour ceux qui n'utilisent pas l'anglais américain. Ceci dit, le programme lui-même est très convivial et assez agréable à utiliser, grâce à son interface *Windows*. Nous avons mentionné ci-dessus qu'on souhaitait une intégration complète du dictionnaire électronique avec les autres applications du PTT. Le *Random House* est un excellent exemple d'une intégration réussie des applications. À partir de son traitement de texte, le traducteur peut sélectionner un mot avec la souris dans le document anglais et, par une simple touche d'appel, obtenir automatiquement sa définition dans une boîte de dialogue («pop-up window») du dictionnaire, sans même avoir à retaper le mot<sup>3</sup>. Et dans la direction inverse, si le traducteur veut copier une partie de la définition dans le document de son traitement de texte, il lui suffit de la sélectionner, puis de frapper deux petites touches d'appel, et le tour est joué.

L'*Oxford English Dictionary* (ou OED) est un monument de la lexicographie. Il est sans aucun doute le dictionnaire le plus complet de la langue anglaise, comprenant environ un demi-million de définitions et plus de 2,4 millions de citations. C'est avant tout un dictionnaire diachronique, qui trace au fil des citations l'évolution du vocabulaire de l'anglais à travers les siècles. L'apparition sur CD-ROM de ce dictionnaire, longtemps réservé aux philologues, soulève une question intéressante : un dictionnaire peut-il offrir trop d'informations pour un traducteur ? Il faut bien dire en effet que, la plupart du temps, le traducteur, et en particulier le traducteur en langue technoscientifique, trouvera superflue toute cette information historique. Cependant, l'interface de l'*OED sur CD-ROM* est tellement bien faite qu'elle rend la question caduque, puisque l'utilisateur peut supprimer l'affichage des citations et de l'étymologie pour ne retenir que les définitions. Donc, le seul véritable obstacle demeure le prix, non seulement de l'*OED* lui-même<sup>4</sup>, mais aussi d'un lecteur de CD-ROM<sup>5</sup>.

Si le traducteur francophone hésite à déboursier de telles sommes, l'existence du *Robert électronique*, également disponible sur CD-ROM, lui fournira une raison supplémentaire de se procurer un lecteur de disque compact. À notre connaissance, ce *Robert* est le plus complet des dictionnaires du français disponibles sous forme électronique. Il contient toute l'information du *Grand Robert de la langue française* – c'est-à-dire 80.000 entrées, comprenant la prononciation, l'étymologie, des citations et des synonymes – et, comme nous l'avons déjà mentionné, il permet des recherches qui sont impraticables dans la version reliée du *Grand Robert*. On peut, par exemple, lui soumettre une recherche avec troncature au début d'un mot, et obtenir une liste de toutes les formes qui riment avec ce mot. Plus utile pour le traducteur, peut-être, il offre la

3 À condition que le traitement de texte soit également conçu pour *Windows*.

4 1 250 \$ CDN

5 entre 200 \$ et 1 200 \$ CDN, selon la vitesse, l'interface, etc.

conjugaison complète de tous les verbes contenus dans le dictionnaire. Le *Robert électronique* n'offre pas autant de citations ni d'informations historiques que l'*OED*; mais pour le traducteur ceci n'est pas nécessairement un inconvénient. Son interface DOS n'est peut-être pas des plus conviviales, mais ceci est largement une question de goût personnel. De toute manière, son éditeur a déjà annoncé une nouvelle version avec interface *Windows*, ce qui devrait rendre le couper-coller entre le dictionnaire et le traitement de texte plus aisé qu'il ne l'est actuellement. Quoiqu'il en soit, les participants aux essais du PTT ont été unanimes : même avec son interface DOS, le *Robert électronique* constitue une ressource indispensable pour le traducteur francophone et il vaut son pesant d'or<sup>6</sup>. Pour les traducteurs autonomes qui trouvent que le prix du *Robert* demeure prohibitif, il existe quelques autres dictionnaires électroniques du français – dans la série *Larousse Référence Électronique*, par exemple –, mais aucun qui puisse rivaliser avec ce grand CD-ROM.

#### 4 Les dictionnaires bilingues

Ce qui distingue le travail du traducteur de celui du rédacteur, c'est que le premier concerne non pas un mais deux textes : un texte source pré-existant et un texte cible à produire, qui doit respecter une certaine relation d'équivalence avec le premier. Les dictionnaires bilingues sont ceux qui se rapprochent le plus de ce travail contrastif, puisqu'ils mettent en regard les unités lexicales de la langue source et celles de la langue cible. Pour cette raison, un poste de travail conçu spécialement pour les traducteurs devrait intégrer tous les meilleurs dictionnaires bilingues qui existent pour une paire de langues donnée. Malheureusement, ceci n'est pas possible actuellement, du moins en ce qui concerne la traduction anglais > français. D'abord, le nombre de dictionnaires bilingues disponibles sur le marché sous forme électronique est inférieur au nombre de dictionnaires monolingues. Qui plus est, les quelques dictionnaires bilingues informatisés sont loin d'être les meilleurs. En fait, nous n'en connaissons que trois ou quatre. Dans la série *Larousse Référence Électronique*, on citera le dictionnaire bilingue *Business Larousse* de Peter Collin, qui, selon les distributeurs, contient plus de 40 000 mots et expressions dans chaque langue. Nous trouvons ce dictionnaire assez limité, même pour le domaine des affaires, et son interface, ou plus exactement la disposition des données dans les fenêtres, laisse beaucoup à désirer.

---

6 Combien d'or exactement ? Au Canada, le CD se vend à plus de 1 400 \$ l'unité. Voilà une autre bonne raison, s'il en fallait, de relier en réseau les PTT des traducteurs qui travaillent dans un même service : il coûte moins cher de partager des ressources comme le *Robert* dans leur version réseau que d'acheter des disques compacts pour un nombre équivalent de postes individuels, même en comptant le coût du réseautage.

Pour ce qui est des dictionnaires bilingues généraux, il en existe au moins deux pour l'anglais et le français qui proviennent de la maison Harrap. D'abord, les développeurs du logiciel *Mercury/Termex*, un des programmes de gestion de glossaires que nous avons testé dans le PTT, offrent une version du *Harrap's Concise*. Plus précisément, ils ont repris le contenu de ce dictionnaire et l'ont reformaté afin qu'on puisse l'interroger comme un glossaire *Termex* non-modifiable. Il comprend 43 500 fiches *Termex*, ou plus de 70 000 équivalents, mais uniquement dans le sens anglais > français<sup>7</sup>. En outre, l'utilisateur doit absolument disposer du programme *Termex* pour pouvoir consulter ce dictionnaire, ce que certains pourraient voir comme un inconvénient. Ceci n'est pas le cas pour la série de dictionnaires bilingues *Collins*, qu'on peut installer et consulter en tant que programmes autonomes. Mais, ne renfermant qu'environ dix mille entrées chacun, ces dictionnaires de poche ne sont pas assez détaillés pour les besoins des traducteurs professionnels.

L'autre dictionnaire Harraps, le *Harrap's Shorter électronique*, qui est publié par la société Chambers Harrap d'Edimbourg, est beaucoup plus intéressant. D'abord, il est offert sur CD-ROM ainsi que sur disquettes, et il contient un plus grand nombre d'entrées que le *Harrap's Concise* – plus de 235 000. Ensuite, c'est un vrai dictionnaire bi-directionnel, exactement comme son pendant sur support papier. En fait, le *Harrap's Shorter* comprend deux fichiers de données distincts auxquels on accède par le même programme de recherche. Lorsque le traducteur veut chercher un terme dans l'autre direction, il doit changer de volume en cliquant sur une des icônes de livre rangées dans la bibliothèque. Autre avantage du *Harrap's Shorter* : son interface *Windows*, qui est assez bien conçue, contrairement à celle du *Business Larousse*<sup>8</sup>. Encore une fois, la version en-ligne de ce dictionnaire permet des recherches qui seraient difficiles à effectuer avec l'édition papier. On peut demander, par exemple, l'affichage de tous les articles qui contiennent, dans leurs définitions ou leurs exemples, un mot donné. Autre option offerte non seulement par le *Harrap's* mais par plusieurs dictionnaires électroniques : le programme conserve en mémoire la liste de tous les articles consultés pendant la séance de travail, ce qui facilite le retour aux articles précédents.

Mais, dans le fond, ce n'est pas tellement sa facilité d'interrogation qui amène les traducteurs à préférer un dictionnaire bilingue à un autre, mais plutôt son contenu. Et pour la traduction anglais > français, force est de constater que le dictionnaire bilingue le plus utilisé par les traducteurs, du moins au Canada, n'est pas disponible dans un format qui permettrait son intégration dans le PTT. Pendant longtemps, les grandes maisons d'édition gardaient jalousement leurs

7 Même si un utilitaire *Termix* permet de renverser la direction des entrées, c'est-à-dire de créer la vedette d'une nouvelle fiche à partir de chaque équivalent français, l'opération demeure compliquée et les résultats ne sont pas entièrement satisfaisants.

8 En fait, c'est la même société néerlandaise, AND Software B.V., qui a conçu l'interface du *Harrap's Shorter électronique* et du *OED sur CD-ROM*.

dictionnaires chez elles et empêchaient leur parution sous forme électronique. Ceci n'est pas le cas du *Robert-Collins*, qui est sorti il y a quelques années dans un format de mini-disque compact que Sony essayait de promouvoir. Nous ignorons si Sony a réussi à vendre beaucoup d'exemplaires de cette version du *Robert-Collins* en Europe, mais au Canada, elle n'a pas percé du tout, et l'absence de ce dictionnaire sur le marché grandissant des aides informatiques à la traduction est certainement regrettable. On peut espérer que la mise sur le marché, annoncée pour le début 1995 par les éditeurs, d'une version CD-ROM du tout récent *Oxford-Hachette French Dictionary* permettra de mieux occuper le terrain.

## 5 Une mémoire de traduction

Comme on l'aura vu, notre approche jusqu'ici dans le projet PTT a été résolument axée sur l'utilisation de ressources disponibles dans le commerce. En conformité avec cette approche, tous les dictionnaires que nous avons intégrés dans le PTT sont des versions électroniques de dictionnaires traditionnels. D'ailleurs, il est à prévoir que, de plus en plus, les nouveaux dictionnaires paraîtront en version électronique en même temps qu'en version papier (voir Knowles, dans ce volume). Dans le cadre du projet PTT, nous nous hâterons de les évaluer et, le cas échéant, de les intégrer dans notre poste. Cependant, même dans les meilleurs dictionnaires, le traducteur ne trouve pas toujours ce qu'il recherche. Le vocabulaire d'une langue naturelle est si vaste que les lexicographes qui préparent les dictionnaires n'ont guère le choix : ils doivent faire une sélection des unités lexicales qu'ils vont retenir. Souvent, cette sélection est basée sur la fréquence des mots ou des expressions (ou de leurs sens) dans un corpus représentatif de la langue courante. Il arrive parfois que les auteurs choisissent d'inclure dans leur dictionnaire des mots peu courants ou des expressions rares, estimant que ces éléments de lexique peuvent être problématiques pour un lecteur qui serait ainsi conduit à consulter un dictionnaire<sup>9</sup>. Mais quels que soient les critères, les lexicographes sont obligés d'exclure certaines unités lexicales. Car un dictionnaire papier, même en plusieurs volumes, ne peut pas contenir tous les sens de tous les mots d'une langue vivante, et toutes ses expressions figurées, ainsi qu'un nombre suffisant de citations pour illustrer tous leurs emplois et toutes leurs nuances; sans mentionner toute la nouvelle terminologie de tous les domaines techniques. Il n'est donc pas surprenant que les utilisateurs ne trouvent pas toujours ce qu'ils recherchent, puisque les dictionnaires, même les plus récents, sont nécessairement incomplets.

---

9 Pour une discussion édifiante des critères du lexicographe, voir Sinclair (1987).

Lorsque le traducteur est confronté à ce problème il a la ressource de consulter ses collègues. Les résultats ne sont pas toujours fiables puisque le temps et la mémoire peuvent faire défaut. Même si l'expression recherchée a été traitée dans un document déjà traduit par le service rien ne permet d'assurer qu'elle sera facilement retrouvée.

Une autre solution peut heureusement être imaginée. On peut supposer que le PTT inclue un système d'interrogation de bases de données permettant au traducteur d'obtenir rapidement tous les équivalents traductionnels d'un mot ou d'une expression problématique, tel qu'ils figurent dans tous les textes déjà traduits par son service – même si cela représente des dizaines de millions de mots de production antérieure. Ainsi, sans quitter son poste, il lui suffirait d'ouvrir la fenêtre du PTT qui donnerait accès à cette base de données et, par le truchement d'une interface graphique simple, de soumettre l'expression problématique au système. Le système afficherait de façon quasi instantanée toutes les occurrences de l'expression trouvées dans la base de données, chacune dans son contexte phrastique, et présenterait en regard de chaque exemple la phrase en langue cible contenant sa traduction. Le traducteur pourrait alors parcourir ces exemples et choisir l'équivalent qui lui semblerait le plus approprié.

En fait, un tel système ne relève pas de l'imaginaire; il existe déjà depuis quelque temps, quoiqu'il ne soit pas encore intégré au PTT. Techniquement, il s'agit d'un concordancier bilingue, et celui qui a été développé au CITI porte le nom de *TransSearch*<sup>10</sup>. Pour le moment, *TransSearch* tourne uniquement sur des stations SUN et ne permet d'interroger que quelques bases de données, dont la plus importante compte des dizaines de millions de mots provenant des débats de la Chambre des communes du Canada. Le CITI est actuellement en pourparlers avec des partenaires du secteur privé en vue de porter *TransSearch* sur une plate-forme d'ordinateur personnel et de permettre aux utilisateurs éventuels du système de constituer des bases de données à partir de leur propre production antérieure.

*TransSearch* représente une façon d'exploiter ce qui est parfois appelé une mémoire de traduction; il en existe d'autres. Tous les systèmes incorporant une mémoire de traduction cherchent à recycler les traductions passées au moyen de l'appariement de segments correspondants dans une série de textes sources et cibles : que ce soit pour éviter la retraduction de phrases répétées, comme le font les systèmes *TM/2* d'IBM ou *Optimizer* d'Eurolang, ou pour composer une nouvelle traduction par l'assemblage des traductions de syntagmes similaires à ceux du nouveau texte à traduire, comme essaient de le faire certains projets de recherche japonais<sup>11</sup>. *TransSearch* est une application plutôt modeste de la

---

10 Voir Isabelle (1992) et Isabelle et al. (1993) pour plus de détails sur *TransSearch*, ainsi que sur le concept de bi-texte sur lequel il repose.

11 Voir, entre autres, Sato & Nagao (1990), l'article à l'origine de tous les travaux sur EBMT (c.-à-d. «example-based machine translation») au Japon.

mémoire de traduction, qui se démarque des autres applications par son interactivité. L'utilisateur, et non le système, décide quand et comment interroger la base de données de traductions mémorisées pour trouver la réponse à un problème particulier – exactement comme il décide de consulter un dictionnaire bilingue. Mais, en soi, le concordancier bilingue n'est pas un dictionnaire bilingue, et il est intéressant d'essayer de préciser en quoi les deux diffèrent.

Un dictionnaire bilingue présente la synthèse de certaines informations traductionnelles. Normalement, chaque article du dictionnaire inclut une indication de la prononciation de la vedette, suivie de sa catégorie grammaticale. Vient ensuite toute l'information proprement traductionnelle, et c'est ici qu'on voit l'effort de synthèse effectué par les lexicographes. Dans le cas de mots ayant plusieurs équivalents dans la langue cible – et ceci représente la grande majorité des entrées – on essaie de regrouper les équivalents par domaine. En outre, quand les traductions sont conditionnées par des facteurs grammaticaux ou sémantiques, on essaie d'illustrer ceux-ci par des exemples. Ainsi, en consultant un article typique, l'utilisateur apprendra que le mot source *x* peut se traduire par T1 dans le domaine de la pharmacie, par exemple, ou par T2 dans le domaine commercial; il verra que s'il a tel sens, *x* peut se traduire par T3, T4 ou T5, alors qu'il se traduit par T5 ou T6 s'il est employé dans un autre sens. Enfin, quand son sujet est du même type sémantique que le nom *y*, *x* peut se traduire par T7, alors qu'avec un objet de type *z*, il se traduit par T8. Toutes ces informations métalinguistiques sont codées de façon presque télégraphique dans l'article et exigent de la part d'un utilisateur intelligent une bonne dose d'interprétation. Qui plus est, l'information est offerte de façon indicative : elle n'est pas présentée sous forme de règles de traduction qui seraient directement interprétables par un système de traduction automatique, par exemple. Et même si la plupart des distinctions dans un article sont illustrées d'exemples accompagnés de bribes de contexte, il arrive souvent que l'utilisateur ne trouve pas dans ces exemples le contexte exact du mot qu'il recherche tel qu'il figure dans son document.

Le concordancier bilingue n'a pas les ambitions du dictionnaire bilingue. Au lieu de chercher à synthétiser toute l'information traductionnelle, le concordancier bilingue se contente de donner accès à ces mêmes informations, mais en vrac. En fait, une des tendances de la lexicographie moderne est de substituer aux intuitions du lexicographe le recours systématique à d'énormes corpus de textes comme source principale d'information lexicale, et il se trouve que le concordancier monolingue est justement un des outils informatisés que les lexicographes utilisent depuis un certain temps déjà pour faire du dépouillement dans ces corpus<sup>12</sup>. Le concordancier bilingue, lui, a dû attendre l'avènement récent de techniques statistiques permettant l'appariement automatique d'énormes quan-

<sup>12</sup> Church & Hanks (1990) discutent de l'utilisation du concordancier par les lexicographes et proposent de nouveaux instruments informatisés très intéressants.

tités de textes traduits. Maintenant que ces corpus bi-textuels sont disponibles, le traducteur peut les interroger au même titre que le lexicographe, grâce à un système comme *TransSearch*. Mais là où le lexicographe utiliserait ce genre d'outil pour l'aider à sélectionner ce qu'il va retenir pour les articles de son dictionnaire, le traducteur peut l'utiliser pour compléter l'information qui manque au dictionnaire bilingue.

Car il faut bien le dire, un concordancier comme *TransSearch* ne remplace pas un bon dictionnaire bilingue. Dans son état actuel, *TransSearch* n'effectue aucun tri dans les résultats qu'il affiche : il répète les mêmes résultats autant de fois qu'il les retrouve dans la base de données et il les affiche dans un ordre qui est strictement chronologique. Donc, pour les mots les plus fréquents, l'utilisateur risque d'être noyé dans un surplus d'information, et souvent il ne prendra pas le temps de tout compiler. Autre différence importante : l'information repérée par le concordancier bilingue n'a pas toujours été validée, contrairement à celle contenue dans un dictionnaire bilingue. Ainsi, il est possible que le traducteur retrouve dans la base de données certaines traductions qui sont en fait à éviter; puisque le concordancier ne fournit aucune indication sur l'usage, le traducteur pourrait être porté à réutiliser ces solutions douteuses. C'est pour ces raisons que nous considérons le concordancier bilingue plutôt comme un complément au dictionnaire bilingue traditionnel. Lorsque le traducteur ne trouve pas ce qu'il recherche dans ses ouvrages de référence standard – ce qui arrive assez souvent dans le cas d'expressions figurées, par exemple – *TransSearch* lui permet de puiser dans cet énorme dictionnaire virtuel constitué de toutes les traductions antérieures dont dispose son service<sup>13</sup>.

Nos propres travaux sur un énorme corpus de ce type nous ont vite convaincus que la production passée des traducteurs constitue une source presque inépuisable pour la traduction. Utilisé avec discernement, un concordancier bilingue comme *TransSearch* permet de libérer toute la richesse de cette mémoire. Après tout, qu'est-ce qu'un dictionnaire bilingue, sinon une mémoire de traduction d'orientation lexicale ?

---

13 À titre d'exemple, nous avons inclus en annexe un extrait des résultats obtenus à la suite d'une recherche par *TransSearch* sur le mot anglais «hardball», qui n'est jamais employé au sens littéral dans les débats de la Chambre des communes.

## Annexe

TransSearch - Recherche

Soumettre Arrêter Langue source Indéterminée  
Majuscules pertinentes

anglais →-| hardball ←-|

---

TransSearch 2.2 -- [TransBase: Hansard]

Fichier Options Chercher... Arrêter

Liste Document Requête: 1/1 Résultat 1/7  
Document: hans:34.03132.250 rev

In May 1992 after the government had capitulated and given in to the <b>hardball</b> tactics of the tobacco industry, the exports jumped from 39 million to 366 million kilograms of fine cut tobacco.	En mai 1992, après que le gouvernement a capitulé et cédé aux tactiques brutales de l'industrie du tabac, les exportations de tabac à coupe fine sont passées de 39 millions à 366 millions de kilogrammes.
The government acts as if it was dealing with a technical problem when what we are watching is <b>hardball</b> politics in which Canada is being made the victim.	Le gouvernement se comporte comme s'il s'agissait là d'un problème technique, alors que nous assistons, en fait, à des manoeuvres politiques pures et simples dont le Canada est la victime.
When it comes to negotiations, the U.S. and Europe play <b>hardball</b> .	Lorsqu'il s'agit de négocier, les Américains et les Européens se montrent intransigeants.
My caller says his dissatisfaction with PSAC started last year, when the support staff at PSAC headquarters went on strike — and Daryl Bean played <b>hardball</b> .	Mon correspondant disait que son mécontentement à l'égard de l'AFPC avait commencé l'an dernier, lors de la grève du personnel de soutien au siège social de l'AFPC, lorsque Daryl Bean a pris une position extrêmement rigide à leur égard.

Recherche : complétée (7résultat(s))

## Chapitre 10

# LE TRAITEMENT DES COLLOCATIONS ET DES EXPRESSIONS IDIOMATIQUES DANS LES DICTIONNAIRES BILINGUES

Roda P. ROBERTS

### 1 Importance des unités idiomatiques dans les dictionnaires bilingues

La langue, nous rappelle Pergnier (1980:307), est non seulement une structure, mais aussi un idiome. Plus précisément, la langue consiste en deux types d'unités : des unités structurales, telles que les morphèmes, les mots et leurs combinaisons libres, et des unités idiomatiques. Les unités structurales sont faciles à comprendre ou à composer pour qui connaît une langue, même si on ne la maîtrise pas parfaitement. Cependant, les unités idiomatiques, qui ne s'expliquent pas entièrement par la structure de la langue, posent d'autres difficultés. Si la structure de la langue explique le rapport entre les éléments de la phrase *It is raining cats and dogs*, rien n'explique pourquoi on dit *cats and dogs* plutôt que *pigs and giraffes*. Et si la connaissance de la structure de la langue et de son vocabulaire permet de comprendre le groupe de mots *un hiver rigoureux*, il n'y a rien qui explique pourquoi on dit en anglais *a severe winter* et non *a rigorous winter*. Ces unités soi-disant idiomatiques, on les assimile, plus ou moins par osmose, dans sa propre langue (sa langue dominante), mais ce n'est généralement pas le cas dans une langue seconde ou étrangère, où il faut les apprendre consciemment. Ainsi, les dictionnaires bilingues, dont le but principal est de permettre à ceux qui travaillent dans une langue seconde ou étrangère de bien décoder ou encoder un texte dans cette langue, doivent attribuer un rôle important à ce type d'unités dans leur microstructure.

---

1 Je remercie le Conseil de Recherches en Sciences Humaines du Canada, qui a subventionné la recherche pour ce travail. Je remercie également Johanne Blais, associée de recherche, qui a révisé ce travail.

## 2 Définition et classification des unités idiomatiques

J'entends, par «unités idiomatiques», des groupes de mots qui sont non seulement reliés par des règles grammaticales, mais par d'autres restrictions, surtout sémantiques, et aussi, dans certains cas, syntaxiques. Ces unités se divisent en deux grandes catégories : celle des expressions idiomatiques et celle des collocations.

L'expression idiomatique, que Pergnier (1980:310) définit comme un «ensemble de mots dont le désigné est global, et qui ne peut par conséquent trouver son équivalent dans une autre langue que globalement» et que Dagneaud (1965:95) décrit comme un ensemble de mots faisant bloc et qu'on emploie dans un sens fixe, est régie à la fois par des restrictions sémantiques et morphosyntaxiques. Du point de vue sémantique, chacun des éléments linguistiques qui la compose perd son sens individuel à la faveur d'un sens unique pour l'expression, qui est souvent, mais pas toujours, figurée : ainsi, dans le groupe de mots *c'est une autre paire de manches*, le sens individuel de *paire* et de *manches* disparaît, mais ce groupe entier de mots prend le sens «c'est une chose entièrement différente». Du point de vue morphosyntaxique, une telle unité ne permet pas de modifications ou peu : on ne peut dire, par exemple, «ce n'est pas une autre paire de manches», ni «c'est deux autres paires de manches», ni «ce sont une autre paire de manches»; et bien qu'il semble être possible d'insérer dans l'expression un nombre limité de mots (le Grand Robert donne *c'est toujours une autre paire de manches*), l'étude d'un corpus français présentant 14 occurrences de cette expression sans aucun ajout démontre que l'expression généralement est perçue comme entièrement fixe.

D'autre part, la collocation, que Mounin (1974:71) définit comme «l'association habituelle d'une unité lexicale avec d'autres unités», n'est pas fixe du point de vue morphosyntaxique. Pour reprendre l'exemple *un hiver rigoureux*, on peut dire sans problème *nous avons subi deux hivers rigoureux* ou *c'était un hiver très rigoureux* ou même *l'hiver est très rigoureux*. D'ailleurs, du point de vue sémantique, chacun des éléments de la collocation garde un sens individuel, même si le sens d'un élément (*rigoureux*) est limité par celui de l'autre. Ce qui fait que ce groupement est une collocation et non une combinaison libre, c'est que le nombre d'adjectifs qu'on peut utiliser pour décrire un hiver qui est dur à supporter du point de vue du climat est limité : à part *un hiver rigoureux*, on peut dire *un hiver rude*, mais c'est à peu près tout. Ainsi *rigoureux* est habituellement associé avec *hiver*, ce qui se voit par la récurrence de cette unité idiomatique dans les textes (3 fois dans notre corpus). C'est cette récurrence qui nous permet de dire que *un hiver rigoureux* est une collocation.

Si j'ai pris soin de décrire les types d'unités dont le traitement dictionnaire sera abordé dans ce chapitre, c'est qu'il n'y a consensus ni sur leur définition ni sur leurs caractéristiques. En effet, comme le signale Cowie (1981:133), la ligne de démarcation entre les combinaisons libres, les collocations et les expressions

idiomatiques n'est pas nette, car il s'agit d'un continuum. Et bien que les linguistes et les lexicographes soient d'accord sur le fait qu'il existe différents types d'expressions idiomatiques et de collocations, ils les classent différemment suivant soit des critères grammaticaux, soit des critères sémantiques.

Ainsi, l'*Oxford Dictionary of Current Idiomatic English* (1975) et le *Dictionnaire des expressions et locutions figurées* (1979) utilisent des critères grammaticaux pour catégoriser les expressions idiomatiques. L'*ODCIE* les divise en deux grandes catégories : celles qui contiennent des verbes avec particules ou prépositions (présentés dans le volume 1) et celles qui sont constituées de locutions, de propositions ou de phrases (présentées dans le volume 2). Dans le *Dictionnaire des expressions et locutions figurées*, les différentes expressions contenant un même mot-clé sont divisées en (a) locutions nominales, (b) locutions adjectivales, adverbiales, prépositives et conjonctives; (c) locutions verbales; et (d) locutions-phrases. Par contre, Cowie (1981) établit des distinctions entre les expressions idiomatiques en s'appuyant essentiellement sur des critères sémantiques. Il les divise en expressions idiomatiques véritables, c'est-à-dire les expressions dont le sens ne peut plus être analysé (par exemple, *une autre paire de manches*) et les expressions idiomatiques figurées, soit les expressions qui ont un sens figuré pour ce qui est de l'ensemble, mais qui conservent également une interprétation littérale (par exemple, *make a U-turn*).

Quant aux collocations, elles sont classées elles aussi soit grammaticalement, soit sémantiquement. Hausmann (1989:1010) va jusqu'à intégrer un classement grammatical dans sa définition même de la collocation. Selon lui, on appellera «collocation» la combinaison caractéristique de deux mots dans une des structures suivantes : (a) substantif + adjectif (épithète); (b) substantif + verbe; (c) verbe + substantif (objet); (d) verbe + adverbe; (e) adjectif + adverbe; et (f) substantif + (préposition) + substantif. Cette catégorisation de Hausmann se limite à ce que Benson, Benson et Ilson (1986) appellent des «collocations lexicales», qui consistent typiquement en des substantifs, des adjectifs, des verbes et des adverbes. Ces derniers utilisent des critères grammaticaux pour distinguer également entre la classe des collocations lexicales et celles des collocations grammaticales, «consisting of a dominant word (noun, adjective, verb) and a preposition or grammatical structure such as an infinitive or a clause». D'autre part, Mel'čuk (1984) classe les collocations lexicales en fonction de leur sens : il présente ces collocations d'après la fonction lexicale, qui est un type de sens dont l'actualisation est strictement déterminée par le lexème auquel ce sens est associé. Par exemple, la fonction lexicale *Magn* correspond en gros au sens de «très, intensément ou à un degré élevé» et peut être actualisée par les moyens suivants : (*mémoire*) : *prodigieuse*; (*bruit*) : *infernal*; (*désir*) : *ardent*; (*remercier*) : *vivement*. Mel'čuk a repéré une cinquantaine de fonctions lexicales habituelles, qui sont décrites par des symboles.

Le fait que les expressions idiomatiques et les collocations sont mal définies et aussi classifiées différemment par certains linguistes et lexicographes les rendent difficiles à repérer et aussi à traiter dans les dictionnaires.

### 3 Repérage des collocations et des expressions idiomatiques à des fins d'inclusion dans les dictionnaires bilingues

Le lexicographe bilingue dispose de quatre sources principales pour repérer les collocations et les expressions idiomatiques : la compétence linguistique du lexicographe, la compétence linguistique des informateurs, d'autres dictionnaires, et les corpus. Aucune de ces sources n'est suffisante en soi, ce qui implique qu'il faut les utiliser toutes. Reprenons les exemples *un hiver rigoureux* et *c'est une autre paire de manches*.

La compétence linguistique du lexicographe joue un rôle critique dans le repérage de ces unités, car même s'il les trouve par d'autres sources, il faut qu'il puisse les identifier en tant que telles. Il y a en effet peu de sources qui lui disent «voilà une collocation, voilà une expression idiomatique». Cependant la compétence linguistique du lexicographe n'est jamais exhaustive, surtout quand il est appelé à travailler dans deux langues comme c'est le cas dans la lexicographie bilingue. Il fait donc souvent appel à d'autres informateurs.

Le problème ici est que les informateurs hésitent souvent quand on sollicite leur avis. Par exemple, quand j'ai demandé à une informatrice française si *un hiver rigoureux* était pour elle une collocation, elle a commencé à parler plutôt d'*un hiver dur*, sans se prononcer sur la première. Et quand j'ai demandé à une autre si, selon elle, l'expression *c'est une autre paire de manches* était complètement invariable, elle a dit qu'elle pensait que non, mais a voulu vérifier sa réponse dans un dictionnaire !

Quelles réponses trouve-t-on dans les dictionnaires ? Même si on trouve les collocations et les expressions idiomatiques qu'on cherche dans les dictionnaires généraux unilingues (et cela peut parfois être long, car on ne sait jamais sous quel mot de l'unité on va les trouver), elles sont difficiles à identifier comme collocations ou expressions idiomatiques, car elles sont généralement insérées parmi d'autres exemples qui sont des combinaisons libres, souvent sans étiquette pour signaler leur statut différent. Ainsi, *c'est une autre paire de manches* figure, parmi d'autres exemples, à l'intérieur d'une division sémantique, dans le *Grand Robert* et le *Petit Robert* (sous *manche* et *paire*), ainsi que dans le *Lexis* (sous *manche*), mais elle n'est marquée comme locution que dans l'entrée «manche» du *Grand Robert* et dans le *Lexis*. Et *un hiver rigoureux*, qui se trouve dans le *Grand Robert* et le *Petit Robert* (sous *hiver* et *rigoureux*), ainsi que dans le *Lexis* (sous *rigoureux*), n'est marquée nulle part comme une collocation. Par ailleurs, les dictionnaires, généralement unilingues, d'expressions ou de collocations sont des sources plus sûres, car le fait même de trouver ces unités dans ces types de dictionnaires les identifie comme collocations ou comme expressions idiomatiques. L'expression idiomatique *c'est une autre paire de manches* se trouve dans le *Dictionnaire des expressions et locutions figurées* de Rey et Chantreau, bien qu'elle n'y soit que sous une forme tronquée (*une autre paire de manches*). Toutefois, le seul dictionnaire de collocations françaises (Lacroix, 1956) est trop

restreint pour même contenir les collocations avec le mot *hiver*. Enfin, les dictionnaires généraux bilingues ne sont pas meilleurs que les dictionnaires généraux unilingues pour l'identification des collocations : *hiver rigoureux* se trouve dans le *Robert Collins Senior* (sous *rigoureux*), bien que sans indication qu'il s'agit d'une collocation, mais non dans le *Oxford Hachette*<sup>2</sup> ni dans le *Larousse*. Quand les expressions idiomatiques sont incluses dans les dictionnaires généraux bilingues (*c'est une autre paire de manches* l'est), elles sont généralement bien isolées des autres exemples et donc faciles à repérer (voir section 5 ci-dessous).

Enfin, comme source de repérage, peut-être la meilleure, on a les corpus, qui, s'ils sont suffisamment vastes et variés, présentent les collocations et les expressions idiomatiques courantes. Cependant les concordanciers standard laissent au lexicographe le soin de les repérer dans les lignes de contexte (les KWIC). Il est certes assez facile d'y noter les unités répétées, qui sont le signe des unités idiomatiques (*un hiver rigoureux* se trouve 3 fois dans notre corpus<sup>3</sup>, *c'est une autre paire de manches*, 14 fois). Cependant, toutes les unités répétées ne sont pas forcément des collocations ou des expressions idiomatiques et le repérage final de ces unités dépend finalement de la compétence du lexicographe. Il y a maintenant des logiciels qui commencent à se développer pour le repérage de ce genre d'unités (par exemple, Xtract de Frank Smadja<sup>4</sup>), mais ils ne sont pas encore très perfectionnés (Xtract atteint un taux de «précision» d'environ 80 %), et, de plus, ils ne distinguent pas les collocations des expressions idiomatiques.

En résumé, il ressort que le repérage de ces unités, qui est l'étape qui précède celle de leur traitement à proprement parler, n'est pas une opération simple. Le lexicographe doit avoir recours à toutes les sources – sa propre compétence linguistique, la compétence linguistique des informateurs, d'autres dictionnaires, et les corpus – pour en établir une liste plus ou moins complète.

#### **4 Choix des collocations et des expressions idiomatiques pour inclusion dans les dictionnaires bilingues**

Bien que, pour les raisons indiquées dans la section 1, un dictionnaire général bilingue destiné au grand public doive, idéalement, contenir autant de collocations et d'expressions figées que possible, en réalité, le lexicographe est obligé de faire

2 Dans ce dictionnaire, cependant, cette collocation est présentée indirectement (sous le mot *rigoureux*) par l'utilisation des mots contextuels [*climat, saison*].

3 Il s'agit du corpus informatisé TEXTUM, du projet de *Dictionnaire bilingue canadien*, qui contient plus de 100 millions de mots, dont près de la moitié en français. La partie française comprend des journaux canadiens et français, ainsi que quelques livres.

4 Voir F. Smadja (1994) : «Retrieving Collocations from Text : Xtract», in S. Armstrong, *Using Large Corpora*, Cambridge, Mass., MIT Press, pp. 143-177.

un choix, car un dictionnaire général ne présente pas que ces types d'unités. Comment alors doit-il faire cette sélection ?

D'abord, il doit tenir compte de leur fréquence d'emploi. Si, par exemple, il repère les collocations *hiver rigoureux* et *hiver rude* dans les dictionnaires, mais que pour des raisons d'espace ou autres il ne veut pas les inclure toutes les deux, le nombre d'occurrences de chacune dans un corpus peut lui servir de guide. Ainsi, selon notre corpus, *hiver rigoureux* (3 occurrences) est plus fréquent que *hiver rude* et serait donc la collocation à retenir. Mais puisqu'un corpus n'est qu'une représentation possible de la langue, il faut avoir recours à d'autres critères.

Un autre critère très important pour la lexicographie bilingue sera ce que Ivir (1988:49) appelle «the contrastive pattern of collocation in the two languages». C'est un fait bien connu que les collocations et les expressions idiomatiques sont particulières à une langue donnée et qu'il n'y a pas de règles qui indiquent à l'utilisateur si une collocation ou expression figée dans une langue est traitée pareillement ou différemment du point de vue sémantique dans une autre langue. Il est évident que le lexicographe doit donner la priorité à des unités qui ne coïncident pas d'une langue à l'autre. Selon ce critère, il retiendrait non seulement *hiver rigoureux* mais aussi *hiver rude* dans un dictionnaire français > anglais, car les deux collocations sont rendues par «severe winter». Par ailleurs, selon ce même critère, il peut éliminer *hiver dur*, car cette collocation peut être traduite par «hard winter». Ce critère, si logique soit-il, n'est cependant pas évident pour l'utilisateur, qui pourrait se demander comment rendre *hiver dur*, par exemple; s'il est adopté, il faut que le lexicographe l'explique dans l'introduction. En outre, l'adoption de ce critère exige que le lexicographe commence d'abord par traduire une liste relativement longue de collocations et d'expressions idiomatiques pour éliminer celles qui se traduisent «directement».

Le lexicographe peut encore décider de ne retenir qu'une collocation ou expression idiomatique pour une idée donnée. Selon ce critère, il éliminerait soit *hiver rigoureux*, soit *hiver rude*, car les deux collocations dénotent la même réalité; par contre, il retiendrait *hiver dur*, car il ne s'agit pas précisément de la même réalité. Ce critère a, cependant, le désavantage de laisser perplexe l'utilisateur qui cherche la collocation éliminée et qui ne sait pas qu'elle est synonyme de celle qui est retenue.

La question des critères de choix a encore besoin d'être approfondie. Ce qui est clair, c'est qu'ils ne seront pas forcément les mêmes que ceux qui sont utilisés dans un dictionnaire unilingue.

## 5 Statut et placement des collocations et des expressions figées dans les dictionnaires bilingues

Le placement des collocations et des expressions figées dans les dictionnaires bilingues dépend du statut qu'on leur accorde. Les considère-t-on comme des types spéciaux d'exemples ou plutôt comme des sous-entrées ? Cette question est soulevée par Jacobsen *et al.* (1989:2784) qui semblent suggérer que ces unités sont plutôt des sous-entrées regroupées avec un mot-vedette pour plus de commodité :

it is true in a sense to say that everything in a bilingual dictionary that is not the headword or explicit metalinguistic information exemplifies the use of the headword or its translation equivalents. Therefore lexicographers tend to use *example* vaguely when referring to sub-entries (idioms, collocations and other syntagms with their translation equivalents). This situation is largely due to the historical dependence of bilingual dictionaries on the content and categories of monolingual dictionaries, in which there is considerably more justification for seeing sub-entries as exemplifications of the headword. In bilingual dictionaries, however, the majority of sub-entries provide one-to-one translation equivalents. They *exemplify* nothing.

En regardant la microstructure des dictionnaires bilingues publiés avant les années 80 et depuis, on remarque une certaine évolution dans le statut accordé aux expressions figées. Là où le *Harrap's New Standard French and English Dictionary* (1972) présentait encore ensemble pêle-mêle les combinaisons libres, les collocations, et les expressions figées, comme «exemples» d'un sens donné du mot-vedette, le *Larousse* (1992), le *Robert Collins Senior* (1993), et le *Oxford Hachette* séparent nettement les expressions idiomatiques des exemples à proprement parler et les placent dans une sous-division indépendante, qui est clairement distinguée des sous-divisions sémantiques par une abréviation comme *loc* (dans le *Robert Collins Senior*), par un signe graphique (dans le *Larousse*), ou par un mot comme *idioms* ou *idiomes* (dans le *Oxford Hachette*).

Examinons maintenant comment le *Harrap's* et le *Oxford Hachette* traitent le mot *manche* (nom féminin), et plus particulièrement le statut qu'ils accordent aux expressions idiomatiques et aux collocations<sup>5</sup>.

*Harrap's*

**manche**<sup>1</sup>... *s.f.* **1.** (a) sleeve; être en manches de chemise, to be in one's shirt sleeves; ... robe à manches, sans manches, sleeved, sleeveless, dress; fausses manches... oversleeves; relever, retrousser, ses manches, to roll up one's sleeves; retroussons nos manches, let's get down to work; mettre qch. dans sa m., (i) to put sth. up one's sleeve; (ii) to get hold of, lay hands on, sth.; garder (qn, qch.) dans sa m., to keep (s.o., sth.) in reserve; avoir qn dans sa m., to have s.o. at one's disposal; F : in one's pocket; ... tirer qn par la m., (i) to pull

5 Les articles ont été abrégés. En outre, je n'ai pas pu reproduire certains signes graphiques.

s.o. by the sleeve...; (ii) to attract s.o.'s attention; (iii) to prevent s.o. from leaving; (iv) to coax, entice, s.o.; **se faire tirer par la m.**, to do sth. with reluctance; to be persuaded into doing sth.; *F* : **ça, c'est une autre paire de manches**, that's another matter, another story, another pair of shoes, another cup of tea; **avoir la m. large**, to be easy-going, broadminded; (*b*) *Tchn* : **m. à eau, m. en toile**, (canvas) hose pipe;... **2. Sp** : *etc* : (i) heat; (ii) round; (iii) *Ten* : set; (iv) *Cards* : hand (played); single game; **nous sommes m. à m.**, we're even; we're game all; we're neck and neck;... **3. F** : (*a*) **faire la m.** to beg; to go around with the hat; (*b*) *coll.* **la m.**, beggars;... **4. Geog** : **la M.**, the English Channel.

*Oxford Hachette*

**manche**... *I nm*...

**II nf** 1 *Cout* sleeve; ~ **courte/trois-quarts** short/three-quarter sleeve; **robe à ~s courtes/longues** short-sleeved/long-sleeved dress; **sans ~** sleeveless; **2 Jeux**, Sport round; (aux cartes) hand; (au bridge) game; (au tennis)... set; **3...** (quête) **faire la ~** [*baladin*] to pass the hat round...; [*mendiant*] to beg.

- ~ **à air** Naut air shaft; *Météo* wind sock; ~ **à balai**...

**IDIOMES**... **avoir qn dans la ~** to have sb in one's pocket; **se faire tirer par la ~** to need coaxing; **c'est une autre paire de ~s**... it's a different ball game...

La comparaison de ces deux entrées fait ressortir deux points quant aux expressions idiomatiques : (1) le *Harrap's* donne beaucoup d'importance aux expressions idiomatiques, si l'on juge le nombre de ces unités par rapport au nombre dans le *Oxford Hachette*<sup>6</sup>; (2) par contre, malgré leur nombre, elles attirent moins l'attention dans le *Harrap's*, car elles y sont noyées dans une division sémantique, avec d'autres «exemples» qui comprennent aussi des combinaisons libres, des collocations et des composés. En d'autres termes, aucun statut spécial ne leur est accordé dans ce dictionnaire. Par contre, les trois petites expressions idiomatiques que donne le *Oxford Hachette* sont très «visibles», car on les a séparées des autres exemples, et, en fait, de toutes les divisions sémantiques pour les faire figurer séparément. Cette séparation, qui leur donne un statut différent des autres exemples, celui de vraies sous-entrées, se retrouvera aussi dans le *Dictionnaire bilingue canadien*<sup>7</sup>, où les expressions sont placées après les divisions sémantiques dans une division spéciale indiquée par (*exp*).

Cependant, ce statut n'est pas accordé aux collocations même dans les dictionnaires bilingues récents, comme le démontre l'entrée suivante où figure *un hiver rigoureux* :

6 Bien entendu, la taille du *Harrap's* explique aussi un plus grand nombre d'expressions idiomatiques.

7 Le Dictionnaire bilingue canadien sera un des produits principaux d'un vaste projet de recherche intitulé «Lexicographie comparative du français et de l'anglais au Canada». Ce projet, qui s'échelonne sur environ 10 ans, est présentement subventionné par le Conseil de Recherches en Sciences Humaines du Canada.

*Robert Collins Senior*

**rigoureux**... **adja** (*sévère*) *punition, discipline* rigorous, harsh, severe; *mesures* rigorous, stringent, harsh; (*fig*) *climat* rigorous, harsh; *maître, moraliste* rigorous, strict, rigid. **hiver** ~ hard *ou* harsh winter. **b** (*exact*)... **c** (*absolu*) *interdiction, sens d'un mot* strict. **observation** ~**euse** **du règlement** strict observation of the rule; **ce n'est pas une règle** ~**euse** it's not a hard-and-fast *ou* absolute and unbreakable rule.

Ici, le placement d'*hiver rigoureux* (après les équivalents dans une division sémantique) n'est pas différent de celui d'*observation rigoureuse du règlement*, qui est une combinaison plutôt libre. Ainsi, il n'y a aucune indication tangible quant à son statut spécial.

Il est certes vrai que le statut des collocations n'est pas aussi distinct que celui des expressions figées. Tandis que pour ces dernières il est parfois difficile de déterminer à quel sens d'un mot elles se rattachent, pour les premières ce problème ne se pose pas. Il semble donc logique de les placer à l'intérieur de la division sémantique appropriée. Cependant, je crois qu'il est important de les distinguer nettement des combinaisons libres à cause de leur rôle différent. Les combinaisons libres présentent l'usage typique d'un mot dans une phrase et servent de modèles selon lesquels l'utilisateur peut faire d'autres phrases, tandis que les collocations présentent l'association conventionnelle d'un mot avec un autre, une association que l'utilisateur ne peut pas changer ou faire varier à son gré. C'est pour cette raison que, tout en plaçant les collocations dans les divisions sémantiques, le *Dictionnaire bilingue canadien* sépare les collocations des combinaisons libres par des moyens graphiques. La présentation ci-dessous d'une division sémantique de l'entrée pour *cœur* préparée pour ce dictionnaire montre clairement comment la distinction est faite entre les combinaisons libres, qui sont précédées d'un seul astérisque, et les collocations, qui sont précédées de deux astérisques.

*Dictionnaire bilingue canadien*

**cœur** *nm* **1** (*organe*) heart. \* **le** ~ **fait circuler le sang** the heart pumps blood. \*\* **avoir le** ~ **malade** to have a weak heart/ a heart condition/ heart trouble; **le** ~ **me manqua** I fainted. **qv** **affection, arrêter, battement, battre, défaillir, faible, greffe, lâcher, maladie, opération, palpiter, pulsation, percer, sain, solide.**

Cette entrée soulève une autre question qui n'a pas été abordée jusqu'ici. Puisque les collocations et les expressions idiomatiques comportent plus d'un mot, sous quel mot faut-il les inclure ? Idéalement, du point de vue de l'utilisateur, les collocations et les expressions idiomatiques devraient pouvoir se retrouver sous tous les mots qui les composent; ainsi, il ne serait pas obligé de les chercher dans deux ou trois entrées différentes avant de tomber sur la bonne. Cependant, aussi longtemps qu'on aura des dictionnaires publiés en format livre, le facteur espace limitera le dédoublement d'informations. Le lexicographe aura donc à choisir l'entrée dans laquelle l'unité sera placée, tout en facilitant la recherche de

l'usager. Dans le *Dictionnaire bilingue canadien*, nous avons adopté une politique très claire à ce sujet. La collocation ou l'expression idiomatique sera placée dans l'entrée du premier mot important qu'elle contient (nom, verbe, adjectif, adverbe). La seule exception à cette règle est quand ce mot est un mot passe-partout comme *faire* ou *avoir*, qui fait partie de tant de collocations et d'expressions idiomatiques que celles-ci risquent d'être «perdues» dans l'entrée pour ce mot. Cette décision éditoriale sera clairement énoncée dans l'introduction du dictionnaire pour guider l'usager. En outre, pour faciliter sa consultation du dictionnaire, nous indiquons dans les entrées pour chaque mot important d'une collocation ou d'une expression idiomatique sous quel mot se trouve la collocation ou l'expression recherchée : cette indication est donnée par des «qv» à la fin de la section des collocations ou de la division des expressions idiomatiques.

## 6 Forme de présentation des collocations et des expressions idiomatiques

L'emplacement des unités idiomatiques dans l'entrée est une façon de permettre à l'usager de les reconnaître. Une autre façon serait la forme sous laquelle elles sont présentées.

Il y a en effet deux formes sous lesquelles ces unités se retrouvent généralement dans les dictionnaires : soit isolées dans leur forme «lemmatisée» (par ex. *livrer un combat*), soit intégrées dans une phrase (par ex. dans le *Lexis* : «Cette ethnie [...] livra aux Belges le dernier combat»).

Cependant, toutes ces unités, et surtout les expressions idiomatiques, ne se laissent pas facilement lemmatiser. Ainsi, la présentation de *c'est une autre paire de manches* comme *être une autre paire de manches* donnerait la fausse impression qu'elle peut être précédée de n'importe quel sujet, ce qui n'est pas le cas. Et comment présenter les collocations nom + verbe (par ex. *le cœur me manqua*) autrement que par une phrase ?

Selon Burger (1989), un autre inconvénient de la présentation des unités idiomatiques sous forme lemmatisée est que cela oblige les usagers (surtout ceux qui apprennent la langue) à posséder assez de connaissances morphologiques pour lemmatiser correctement l'unité qu'ils désirent chercher. Cependant, cet argument ne me semble guère valable, car les dictionnaires procèdent par lemmatisation des mots-vedettes et les usagers en prennent vite l'habitude.

Toutefois, la façon de présenter les unités idiomatiques uniquement par des exemples a aussi un inconvénient : elle ne permet pas à l'usager de les reconnaître clairement. Ainsi, il n'y a rien dans la phrase du *Lexis* : «Cette ethnie [...] livra aux Belges le dernier combat» qui indique qu'elle intègre une collocation. En effet, il n'y a rien qui distingue cette phrase-exemple de celle qui suit («Un soldat avait eu les deux bras emportés dans un combat»), qui ne contient pas de collocation.

Les dictionnaires généraux bilingues semblent favoriser la forme lemmatisée quand celle-ci est possible, et une phrase courte ou tronquée quand elle ne l'est pas. Ainsi, le *Oxford Hachette* présente sous forme lemmatisée la collocation *avoir mal au cœur* et l'expression idiomatique *avoir le cœur sur la main*. Toutefois, la collocation *le cœur m'a manqué* est donnée dans une phrase courte et l'expression idiomatique *tant que mon cœur battra* est présentée telle quelle, comme élément de phrase. On remarque la même forme de présentation pour les trois premières unités dans le *Robert Collins Senior* et le *Larousse* (la quatrième ne s'y trouvant pas).

Mais il y a encore une autre forme de présentation que les dictionnaires bilingues utilisent parfois pour des collocations : il s'agit des mots contextuels que le *Oxford Hachette* appelle des indicateurs de collocations, et qui sont utilisés non seulement dans ce dictionnaire mais aussi dans le *Robert Collins Senior* et le *Larousse*. Ainsi, au lieu de trouver la collocation *distinct preference* comme sous-entrée dans ces dictionnaires (soit sous *distinct*, soit sous *preference*), l'utilisateur doit la reconstituer selon les informations suivantes :

*Oxford Hachette*

**distinct**... *adj*... 2 (definite) [*resemblance, preference, progress, impression, memory*] net/nette...

*Robert Collins Senior*

**distinct**... **adj a** (*clear*)... *preference, likeness* marqué, net; ...

*Larousse*

**distinct**... *adj*... 3. [*decided, evident - accent*] prononcé; [- *preference*] marqué; ...

Cette façon de procéder a un avantage évident : elle permet de regrouper plusieurs collocations sans perte d'espace. Mais le problème est que seule la collocation de la langue de départ est donnée en entier, alors qu'on n'a qu'un élément de la collocation de la langue d'arrivée. Ainsi, cette forme de présentation, bien qu'elle soit adéquate pour le décodage, ne l'est pas pour l'encodage, surtout pour les usagers qui commencent à apprendre la langue d'arrivée.

Un autre inconvénient de cette méthode de présentation est que, étant donné le nombre relativement grand des mots contextuels qu'on peut présenter de cette façon, l'utilisateur tend à croire que toute la gamme est couverte, ce qui n'est souvent pas le cas. Ainsi, le *Robert Collins Senior* présente neuf mots contextuels avec *distinct* (*landmark, voice, memory, promise, offer, preference, likeness, increase, progress*) mais «oublie» *possibility* et *advantage*. Bien qu'on puisse dire pour sa défense que les lexicographes ont le droit de choisir les collocations présentées, la forme de présentation que ce dictionnaire a adopté pour les collocations de certains mots se prête à une liste quasi-exhaustive de «collocates» et donc à des fausses attentes de la part des usagers.

Pour éviter d'allonger inutilement la liste de «collocates», les lexicographes remplacent parfois les mots contextuels spécifiques par un générique. Ainsi, pour le mot *abolish*, le *Robert Collins Senior* donne parmi les compléments d'objet «typiques» *practice* et *custom*, qui sont des génériques. Mais l'utilisateur risque de chercher l'équivalent de *abolish* suivi d'un mot spécifique comme *slavery* ou *head tax*. Dans un tel cas, il doit d'abord se demander s'il s'agit bien d'une pratique ou d'une coutume. Il arrive aussi que le mot générique qui est donné mette l'utilisateur sur une fausse piste : par exemple, un des sujets typiques donnés pour le verbe *abate* dans le *Robert Collins Senior* est *emotions* : mais l'usage démontre que c'est surtout avec les émotions négatives (comme *fear* et *rage*) que ce verbe s'associe.

Puisque chaque forme de présentation des collocations et des expressions idiomatiques a à la fois des avantages et des inconvénients, il est impossible de prôner l'une ou l'autre. Il n'y a en fait rien qui empêche le lexicographe bilingue d'utiliser l'une ou l'autre forme, comme font beaucoup de dictionnaires bilingues existants. Mais il faut qu'il y ait une certaine logique dans le choix et que cette logique soit apparente ou indiquée aux usagers. Dans le *Dictionnaire bilingue canadien*, par exemple, nous préférons mettre bien en évidence les collocations et les expressions idiomatiques en les présentant sous forme lemmatisée chaque fois qu'il est possible de le faire; quand ce n'est pas possible, nous donnons une phrase tronquée ou courte; enfin, nous utilisons la méthode des mots contextuels pour présenter les collocations quand le mot-vedette entre en collocation avec un grand nombre de mots.

## 7 Équivalents des collocations et des expressions idiomatiques dans les dictionnaires bilingues

En général, on ne trouve dans les dictionnaires bilingues que la collocation, ou l'expression idiomatique, dans la langue de départ et son équivalent dans la langue d'arrivée<sup>8</sup>. Cet équivalent n'est pas en principe une traduction directe ou littérale, car peu de ces unités se prêtent à ce type de présentation.

Il faut d'abord que le lexicographe essaie de trouver une collocation ou une expression idiomatique équivalente dans la langue d'arrivée pour retenir l'idiomaticité de la langue de départ. Ainsi *c'est une autre paire de manches* sera rendu par *that's another story* (comme dans le *Robert Collins Senior*) ou *that's a different kettle of fish* (comme dans le *Larousse*), mais certainement pas par *\*that's another pair of sleeves*.

8 Il y a aussi les marques d'usage, dont je parlerai plus loin.

Je n'aborderai pas la question du nombre d'équivalents synonymes à donner pour les collocations et surtout pour les expressions idiomatiques. La réponse à cette question est liée à la décision éditoriale sur le nombre d'équivalents synonymes à donner pour le mot-vedette, pour les exemples, et ainsi de suite. Cependant, quelle que soit la politique lexicographique sur les équivalents synonymes, le lexicographe bilingue ne doit jamais oublier qu'une collocation ou une expression idiomatique, comme le mot-vedette, peut avoir des équivalents différents selon le contexte. Ainsi, les collocations *découvrir/dévoiler/ouvrir son cœur* et *dire ce qu'on a sur le cœur* se rendraient par *to pour one's heart out* ou *to bare one's soul* quand il s'agit d'exprimer n'importe quel sentiment intime, par *to get something off one's chest* quand il est question de sentiments négatifs, et par *to make a clean breast of things* quand il s'agit de culpabilité. Il faut, dans ce cas, qu'au moins trois équivalents soient donnés et qu'on fasse la discrimination sémantique entre eux, comme suit :

*Dictionnaire bilingue canadien*

**cœur** nm...

8 (*siège des pensées intimes*) heart, feelings. \*\* *découvrir/dévoiler/ouvrir son ~ = dire ce qu'on a sur le ~ to unburden/pour out (sep) one's heart = to bare one's soul; (sentiments négatifs) to get something off one's chest; (sentiments de culpabilité) to make a clean breast of things...*

Il peut cependant arriver qu'une expression idiomatique n'ait pas d'équivalent «idiomatique» dans la langue d'arrivée. C'est le cas, par exemple, d'*avoir du cœur au ventre* pour lequel il n'existe pas d'expression idiomatique en anglais. Dans ces cas, le lexicographe n'a d'autre choix que de donner un équivalent explicatif, qui en clarifie le sens (par exemple, *to be brave, to have guts*). De même, certaines collocations dans la langue de départ sont rendues non par des collocations mais par des combinaisons libres dans la langue d'arrivée. Ainsi, la collocation française *avoir mal au cœur* se traduit par la combinaison libre *to feel sick / nauseous / nauseated / queasy*. Il faut dans ces cas que le lexicographe bilingue trouve la façon la plus naturelle d'exprimer, dans la langue d'arrivée, le sens de l'unité de la langue de départ.

Faut-il indiquer que tel équivalent est lui aussi une collocation ou une expression idiomatique et tel autre équivalent ne l'est pas ? C'e serait certes une indication utile pour l'utilisateur qui se sert du dictionnaire bilingue pour l'encodage : il saurait ainsi de quelle liberté de manœuvre il peut disposer. Et il serait facile de présenter une telle information par des moyens graphiques (par exemple, en mettant les équivalents qui ne sont pas des collocations ou des expressions idiomatiques, comme pour les gloses). Cependant, aucun dictionnaire bilingue ne l'a fait jusqu'ici.

## 8 Autres informations pertinentes relatives aux collocations et aux expressions idiomatiques qu'on peut inclure dans les dictionnaires bilingues

Bien que les dictionnaires de collocations ne donnent pas les définitions de ces unités<sup>9</sup>, puisqu'elles sont en général faciles à comprendre, certains dictionnaires d'expressions idiomatiques les incluent (par exemple, Rey et Chantreau, 1979; Lafleur, 1966). Faut-il les ajouter dans les dictionnaires bilingues ? L'argument contre l'inclusion des définitions, même pour le mot-vedette, a toujours été que les équivalents les remplacent. Quoique les dictionnaires anglais-français récents tendent de plus en plus à présenter de brèves indications sémantiques sous forme de synonymes ou de paraphrases pour le mot-vedette, ils ne le font toujours pas pour les expressions idiomatiques. Je considère que les définitions des expressions idiomatiques en mots simples seraient une addition éventuelle désirable dans les dictionnaires bilingues, mais qu'il est plus important d'y inclure un grand nombre d'expressions idiomatiques et de collocations (avec leurs équivalents) et surtout d'en donner des exemples d'usage.

Bien que j'aie indiqué dans la section 6 ci-dessus que le fait de présenter des collocations et des expressions idiomatiques dans une phrase-exemple cache leur statut spécial et qu'il vaut mieux les introduire sous leur forme lemmatisée ou invariable, il reste que leur usage mérite d'être illustré par un exemple, surtout quand l'unité a une certaine flexibilité ou quand elle peut être rendue différemment selon les contextes. La possibilité d'ajouter ou non des mots à l'expression *c'est une autre paire de manches*, qui a déjà été discutée dans la section 2, peut être illustrée au moyen d'exemples. Examinons maintenant l'expression *sur le cœur* qui veut toujours dire «on one's mind», mais qui ne se traduit pas toujours ainsi, comme le démontrent les exemples suivants :

**dis-moi ce que tu as sur le cœur** tell me what's bothering you / what's eating you / what's the matter.

**sa remarque, je l'ai gardée longtemps sur le cœur** his remark rankled in my mind / with me for a long time.

Il serait important de présenter dans un tel cas non seulement l'unité idiomatique avec sa traduction mais aussi des phrases-exemples qui montrent son usage avec leurs traductions.

Des exemples seraient aussi utiles pour indiquer les variations syntaxiques possibles dans l'usage de certaines expressions, surtout les verbes à particule en anglais. Ainsi, *to pour out* peut être utilisé avec un objet avant ou après la particule : *to pour out his heart* ou *to pour his heart out*. Cette variation, qui est

9 Le *BBJ Combinatory Dictionary of English* donne souvent des définitions du mot-vedette, pour indiquer le sens du mot-clé dans une collocation.

clairement démontrée par ces exemples, peut cependant aussi être indiquée par un commentaire métalinguistique comme *to pour out (sep)* (où *sep* veut dire séparable).

Des commentaires métalinguistiques, généralement sous forme d'étiquettes, sont de rigueur pour marquer la situation de discours dans laquelle les collocations et surtout les expressions idiomatiques s'emploient. Si tous les types de marques d'usage peuvent être utilisés dans ce contexte, les marques de registre ont une importance particulière pour les expressions idiomatiques, qui sont souvent familières. C'est le cas, par exemple, de l'expression *c'est une autre paire de manches*, qui est, d'ailleurs, marquée comme familière dans tous les dictionnaires bilingues consultés.

Il y a ainsi plusieurs types d'informations lexicographiques qui peuvent accompagner une collocation ou expression idiomatique et son équivalent dans un dictionnaire bilingue. Certaines informations (par ex. les marques) semblent plus importantes que d'autres (par ex. les définitions). Le choix de celles qui figureront dans une entrée de dictionnaire bilingue dépendra d'une part de l'espace disponible<sup>10</sup>, et, d'autre part de l'importance attribuée à ces unités par la direction du projet dictionnaire.

## **9 Arrangement des collocations et des expressions idiomatiques dans un dictionnaire bilingue**

Une fois qu'on a décidé quelles collocations et expressions idiomatiques doivent être incluses dans un dictionnaire, où les placer et comment les traiter, il reste encore une question à régler : celle de l'arrangement de ces unités, l'une par rapport à l'autre. En d'autres mots, si l'on a repéré un certain nombre de collocations et d'expressions idiomatiques à inclure dans une entrée, comment faut-il les ordonner les unes par rapport aux autres ? Selon ce que j'ai dit ci-dessus (section 5), les collocations sont placées dans les divisions sémantiques appropriées. Il faut donc les trier d'après le sens du mot-vedette. Mais après ce premier tri, il faut ordonner les collocations qui vont entrer dans une même division sémantique, tout comme on doit ordonner les expressions idiomatiques à l'intérieur d'une division spéciale.

Prenons à titre d'exemple certaines collocations de *cœur* (dans le sens de «disposition à la bonté, la générosité» - *avoir un / du cœur, en appeler au (bon) cœur de qn, avoir (un) bon cœur, être plein de cœur, avoir le cœur au bon endroit, avoir le cœur sensible, avoir le cœur dur, avoir le cœur sec, être sans cœur, ne pas avoir de cœur, avoir un cœur de marbre, avoir un cœur de pierre, homme de cœur,*

---

<sup>10</sup> Avec la production des dictionnaires sur disque optique, la question de l'espace disparaît.

*femme de cœur*. Comment les ordonner pour rendre la consultation facile pour l'utilisateur ? Il y a plusieurs façons de le faire. La plus évidente est l'utilisation de l'ordre alphabétique selon le premier mot de la collocation (suivi du second, etc.). Selon ce critère, les collocations ci-dessus se présenteraient ainsi :

*avoir (un) bon cœur*  
*avoir le cœur au bon endroit*  
*avoir le cœur dur*  
*avoir le cœur sec*  
*avoir le cœur sensible*  
*avoir un/du cœur*  
*avoir un cœur de marbre*  
*avoir un cœur de pierre*  
*en appeler au (bon) cœur de qn*  
*être plein de cœur*  
*être sans cœur*  
*homme de cœur*  
*femme de cœur*  
*ne pas avoir de cœur*

Un tel arrangement, quoique facile pour le lexicographe et l'utilisateur, a certains désavantages. Le premier est l'interruption des collocations verbe + nom par deux collocations nom + préposition + nom. Ce problème peut être réglé en groupant d'abord les collocations selon les catégories grammaticales des constituants, et ensuite en établissant un ordre de présentation des sous-divisions ainsi établies. Ainsi, les collocations de *cœur* traitées ici peuvent être réorganisées pour mettre les deux collocations nom + préposition + nom (*homme de cœur*, *femme de cœur*) avant ou après les collocations verbe + nom. Deuxièmement, *en appeler au (bon) cœur de qn* est une collocation verbe + substantif, même si son premier mot est un pronom : donc, il vaut mieux la classer selon l'ordre alphabétique du verbe, comme toutes les autres collocations de cette catégorie. Après avoir fait ces changements, les collocations seraient ordonnées de la façon suivante :

*en appeler au (bon) cœur de qn*  
*avoir (un) bon cœur*  
*avoir le cœur au bon endroit*  
*avoir le cœur dur*  
*avoir le cœur sec*  
*avoir le cœur sensible*  
*avoir un/du cœur*  
*avoir un cœur de marbre*  
*avoir un cœur de pierre*  
*être plein de cœur*  
*être sans cœur*  
*ne pas avoir de cœur*  
*homme de cœur*  
*femme de cœur*

Mais même cet ordre présente certains inconvénients : les collocations synonymes comme *avoir (un) bon cœur*, *avoir un/du cœur*, et *être plein de cœur*, qui ont toutes les mêmes équivalents (*to have a good/kind/warm heart* or *to be good-/kind-/warm-hearted*) sont séparées et le lien de la synonymie ne se voit que par la répétition des équivalents. Il vaut peut-être mieux regrouper les collocations synonymes. L'ordre final adopté pour ces collocations dans le *Dictionnaire bilingue canadien* est le suivant :

**cœur** nm... 4 (*disposition à la bonté, la générosité*) heart, generosity. \*\* **en appeler au (bon) ~ de qn** to appeal to sb's generosity; **avoir un/du ~ = avoir (un) bon ~ = être plein de ~** to have a good/kind/warm heart = to be good-/kind-/warm-hearted; **avoir le ~ au bon endroit** to have one's heart in the right place; **avoir le ~ sensible** to be sensitive;... **avoir le ~ dur/sec = être sans/manquer de/ne pas avoir de ~** to be heartless/unfeeling/cold-hearted/hardhearted = to have no heart; **avoir un ~ de marbre/de pierre** to have a heart of stone = to be stony-hearted; **homme\ femme de ~** good-hearted man\ woman.

Cet arrangement, qui a le mérite de regrouper les collocations synonymes, est cependant moins facile à lire que l'arrangement alphabétique. Et les mêmes problèmes se posent avec plus ou moins les mêmes solutions, pour l'arrangement des expressions idiomatiques.

C'est peut-être à cause de la complexité de la question d'arrangement que les dictionnaires bilingues font peu d'efforts pour la régler. En effet, même dans les dictionnaires bilingues qui regroupent les expressions idiomatiques (le *Robert Collins Senior*, le *Larousse*, et le *Oxford Hachette*), l'ordre de ces unités n'est pas évident.

## 10 Conclusion

En résumé, il est clair que le traitement des collocations et des expressions idiomatiques dans les dictionnaires généraux bilingues est assez délicat pour plusieurs raisons : (a) ces unités sont relativement mal délimitées; (b) leur statut lexicographique reste vague; (c) il n'y a pas encore de méthodes sûres pour les repérer; (d) à cause de leur nombre, il faut les sélectionner; (e) leurs équivalents sont souvent peu évidents; et (f) leur placement et leur arrangement posent des problèmes. Cependant, leur importance dans les dictionnaires bilingues ne fait aucun doute. Il faut donc que les lexicographes redoublent leurs efforts pour bien les traiter.



# PROBLÈMES DE PRÉPARATION RÉDACTIONNELLE DE DICTIONNAIRES BILINGUES SPÉCIALISÉS : QUELQUES RÉFLEXIONS

André CLAS

## 1 Introduction

Pour mieux cerner leur objet d'étude, les linguistes ont cherché depuis fort longtemps à élaborer une typologie des ouvrages de référence que l'on appelle dictionnaires, parfois répertoires, lexiques ou glossaires (voir Marelllo, dans ce volume). De nombreuses catégorisations ont été proposées pour les divers types de dictionnaires. Cependant la classification n'est pas aisée puisque les dictionnaires sont évidemment fort variés, tant par leur taille que par leur contenu. En effet, il existe des dictionnaires de poche et des dictionnaires en plusieurs volumes, des dictionnaires généraux et des dictionnaires spécialisés, des dictionnaires monolingues et des dictionnaires bilingues et multilingues, des dictionnaires fonctionnels (dictionnaires de synonymes, d'orthographe, de difficultés...) et des dictionnaires encyclopédiques. Sans oublier qu'il y a encore des dictionnaires pour des usagers dont la langue du dictionnaire est la langue maternelle et des dictionnaires pour des usagers dont la langue cible est langue étrangère. En plus, on connaît encore des dictionnaires d'encodage (dictionnaires pour écrire dans la langue cible) et de décodage (dictionnaires pour déchiffrer un texte de la langue cible), sans d'ailleurs que, sauf dans de rares cas, la distinction soit toujours très claire.

Bref, on peut affirmer sans risque de se tromper que toutes les variantes imaginables de dictionnaires sont possibles; certains ouvrages usurpent même, d'une certaine façon, leur étiquette de dictionnaire parce qu'on sait que le titre sera bien perçu par la clientèle potentielle qui y attache grand mérite.

Quoi qu'il en soit, et sans vouloir chercher dans ces essais de classification plus de nécessité qu'il n'en faut, on peut répartir globalement les dictionnaires en deux grandes catégories, soit les **dictionnaires intralinguistiques** (les diction-

naires monolingues) et les **dictionnaires interlinguistiques** (les dictionnaires bilingues et multilingues, mais dont la langue cible n'est pas une métalangue). Mais les dictionnaires interlinguistiques, tout comme d'ailleurs les dictionnaires intralinguistiques, peuvent être de types fort variés, en fonction des objectifs, des usagers visés, des domaines, des critères de détermination de la nomenclature (macrostructure) ou de l'orientation des entrées (microstructure).

## 2 Difficultés d'élaboration des dictionnaires

Il est, faut-il le rappeler, évidemment très difficile de traiter tout le lexique d'une langue, puisqu'il s'agirait de recenser et d'expliquer plusieurs centaines de milliers de mots et de locutions, à condition bien entendu de s'entendre au préalable sur la définition de mot et de locution, tout comme il est quasi impossible de fournir tout le lexique d'une langue et tous les équivalents correspondants dans l'autre. En clair, il faut admettre dès le départ qu'un dictionnaire n'est au fond jamais complet, mais simplement représentatif d'une langue ou d'un secteur spécifique d'une langue. Un dictionnaire ne peut fournir toutes les réponses à toutes les questions.

## 3 Nomenclature

Comme la nomenclature d'un dictionnaire, c'est-à-dire la globalité des mots qui en forment les entrées ou les adresses (la macrostructure) ne peut être identique au lexique d'une langue dans sa totalité, la nomenclature d'un dictionnaire est toujours un choix délibéré. Un dictionnaire doit nécessairement prendre en charge, dans sa macrostructure, une part importante et représentative du lexique d'une langue, et sa macrostructure doit obligatoirement être structurée pour apporter les informations sur les entrées données. En fait, il s'agit dans tous les cas de déterminer la nomenclature idéale du type ou du format de dictionnaire choisi, choix fait en fonction, le plus souvent, de la clientèle possible. Le public visé joue un rôle capital dans la rédaction d'un dictionnaire. Rappelons qu'un dictionnaire est d'une part une liste de mots avec des explications – reflet culturel puisque les mots ont tous une réalité sociale – et d'autre part un outil pédagogique qui vise à fournir certains renseignements à un groupe d'usagers. Le format, la nomenclature, l'information fournie dans les articles de dictionnaire, tout comme d'ailleurs la présentation matérielle, la typographie, les illustrations, par exemple, ciblent un certain public. Il n'y a pas, il faut le dire, de nomenclature idéale pour tous les domaines ou tous les usagers, il n'y a que des choix qui se veulent nécessairement représentatifs des lexèmes retenus dans le lexique d'une langue et qui par répercussion sont significatifs.

Faut-il même rappeler, et on a voulu y voir une typologie possible des dictionnaires, que le dictionnaire est aussi un objet commercial, et que souvent à tort, ou malgré les rédacteurs qui prennent toutes les précautions voulues dans leurs préfaces, le public assimile dictionnaire et norme ou modèle à suivre, quelle que soit la visée des auteurs ? Mais cela est encore un autre problème.

#### **4 Langue et réalité**

Lorsqu'on parle de dictionnaire, en fait, on met l'accent sur la langue. Il est alors indispensable de se souvenir qu'une langue, et le dictionnaire n'en est que le reflet plus ou moins prototypique, a une existence sociale et qu'elle renvoie à une réalité. Si le signe linguistique est arbitraire avant l'acte de dénomination, il devient obligatoire et nécessaire une fois la dénomination attribuée. La réalité a une prise sur la langue et elle devient source de créativité linguistique lorsqu'il y a prise de conscience d'existence. Il y a encore peu de temps, pour ne citer que quelques exemples banals, personne n'eût osé parler de «mémoire d'ordinateur» ou de «souris», ou encore de «cliquer avec une souris» ! Les nouvelles réalités imposent de nouveaux lexèmes et/ou transforment de nombreux aspects de ceux qui existent déjà. Une langue qui ne s'enrichit pas ou ne se transforme pas est, comme chacun le sait, une langue morte.

Comme il est difficile, pour ne pas dire impossible, d'être complet, le problème fondamental auquel se heurte tout lexicographe est de déterminer ce qu'il faut inclure dans un dictionnaire et ce qu'il faut exclure, ce qui correspond à une réalité durable et ce qui est mode passagère, c'est-à-dire comment en fait bâtir une macrostructure représentative.

Un dictionnaire est donc toujours un choix, plus ou moins calculé. Bien entendu, la question la plus délicate est de fixer des critères raisonnables pour arrêter ce choix. Sous-jacent à toute entreprise lexicographique, il y a donc toute une série de principes qui doivent guider le rédacteur : principes économiques, principes linguistiques, principes sociaux, principes historiques, principes géographiques, etc. On peut donc reprendre, sans risque d'être démenti, l'affirmation qu'il n'existe aucun dictionnaire neutre, en ce sens que tout dictionnaire est nécessairement conditionné par des théories linguistiques, sociales, économiques ou même parfois politiques, même s'il tente de rester représentatif.

#### **5 Les dictionnaires intralinguistiques**

Les dictionnaires intralinguistiques (dictionnaires généraux et dictionnaires spécialisés : synonymes, locutions, verbes, rimes, etc.) sont ceux qui décrivent une seule langue : les entrées et les explications ou définitions sont fournies dans

la même langue. De façon un peu simpliste, on peut dire que l'article de dictionnaire explicite l'entrée, ou mieux, l'article de dictionnaire, au moins une partie, est synonyme de l'entrée ou vedette (l'adresse). La définition commence souvent par une catégorisation très générale, d'une certaine façon par la recherche d'un hyperonyme qui délimite le champ d'appartenance et de spécificateurs qui réduisent ou précisent ce dernier. Bien entendu, toutes ces informations sont accompagnées d'indications métalinguistiques plus ou moins déterminées.

## 6 Les dictionnaires interlinguistiques

Les dictionnaires interlinguistiques ont des entrées dans une langue source et des équivalents dans une langue cible (dictionnaires bilingues) ou dans plusieurs (dictionnaires multilingues). En fait, on peut affirmer que les équivalents des entrées dans la langue cible (l'autre langue) correspondent à ce que représente la définition ou le synonyme dans les dictionnaires intralinguistiques. Mais, il faut encore ajouter, contrairement à ce que l'on peut croire, que les dictionnaires interlinguistiques font, eux aussi, souvent usage d'une métalangue pour fournir les explications permettant d'éclairer tout particulièrement tel ou tel aspect ou de fournir des renseignements indispensables pour telle ou telle raison explicite. On sait, en effet, que l'équivalent de la langue cible ne mérite pas toujours sa fonction d'équivalent et qu'il faut corriger ou préciser quelque peu. D'ailleurs, là encore, la précision peut être variable et être déterminée par le public visé, par exemple. Souvent l'équivalent exact n'existe pas, et il faut traduire, c'est-à-dire fournir une explicitation (glose) dans la langue cible, qui peut être considérée comme le synonyme de l'adresse de la langue source. Ainsi, si *white blue-collar workers* peut correspondre à une entrée de dictionnaire, l'équivalent français apparaît moins lexicalisé : *cols bleus de race blanche*, c'est-à-dire que le syntagmatique n'est pas encore tout à fait du paradigmatique (voir aussi Szende, dans ce volume).

On peut à nouveau répartir les dictionnaires interlinguistiques en deux grandes catégories : les dictionnaires généraux et les dictionnaires spécialisés. La différence que nous introduisons ici concerne l'opposition entre la langue commune et ce qu'on appelle les langues de spécialité.

### 6.1 Les dictionnaires interlinguistiques généraux

Dans le dictionnaire interlinguistique, c'est l'équivalent ou la traduction dans la langue cible qui sert de synonyme à l'entrée. Cela signifie en clair qu'à toute entrée dans la langue source doit correspondre une réponse, un équivalent dans la langue cible. On admet donc obligatoirement comme corollaire que tout est

traduisible. Bien plus, on prend pour acquis que ce qui se dit dans une langue peut être dit dans l'autre et qu'il y a un isomorphisme entre les langues. Dans la réalité cependant, si tout ce qu'on peut dire dans une langue est explicable dans une autre, il faut admettre que l'entreprise n'est pas toujours commode et qu'il peut y avoir certains risques de dire plus ou moins, ou même quelque chose qui n'est que ressemblant ou suggérant. Les possibilités de réponse du lexicographe, quand la langue cible n'a pas d'équivalent du lexème de la langue source, peuvent être, en théorie, l'emprunt, le calque ou l'explicitation dans la langue cible. Le lexicographe doit absolument faire face aux défaillances du lexique de la langue cible. Quoi qu'il en soit, le dilemme du lexicographe se pose sur deux plans. Tout d'abord sur le plan du réel, puis sur le plan linguistique. Le réel, ou mieux ce même réel, existe-t-il dans la langue cible ? Quelle forme linguistique faut-il privilégier ? Fabriquer un emprunt ? Mais cela n'explique rien dans un dictionnaire bilingue. Faire un calque ? Il y a là bien entendu, risque d'influence de la langue prêteuse. Jusqu'à quel point peut-on accepter ce mode de formation ? Expliciter simplement, c'est-à-dire donner une certaine définition dans la langue cible. Le choix du lexicographe n'est pas fait au hasard, ni en toute liberté. Sa base documentaire lui permet de trouver des attestations ou, dans le cas d'une lacune totale, lui donne un modèle de formation statistiquement fréquent dans le domaine étudié pour la création d'un néologisme intégrable au domaine. En fait, la base documentaire (banque de données informatisées, par exemple, et personnes ressources) sert de guide pour choisir une forme linguistique adéquate. Mais là encore, la solution est liée au public visé et au domaine traité. C'est ici qu'il faut introduire une séparation rigoureuse dans le type de dictionnaire en fonction du domaine lexicographique exploré et du public visé.

## 6.2 Les dictionnaires interlinguistiques spécialisés

On sait que les exigences de la communication, l'influence des médias, – presse audio-visuelle, presse écrite –, la nécessité de l'enseignement, les facteurs politiques, économiques et sociaux créent toute une série de conditions qui favorisent l'homogénéité et tendent à gommer les écarts régionaux et variations linguistiques des diverses strates sociales d'une langue. Pourtant, il faut constater qu'aucune communauté linguistique n'est parfaitement homogène puisqu'un ensemble de facteurs, – âge, sexe, occupation, emploi, géographie, statut social, etc.–, introduit des variétés dans l'usage. Ainsi, un facteur important de diversification linguistique est la différenciation des activités socio-économiques. On peut admettre en effet que les développements de la science et des techniques ont largement influencé, – et même accéléré le changement, surtout à l'heure actuelle –, l'évolution de la langue. Ainsi, on peut montrer que le *Petit Robert* de 1982 contient plus de 4500 mots dont la datation est postérieure à 1960 et que le *Nouveau Petit Robert* de 1993 contient, par exemple pour la lettre A, environ

1100 entrées étiquetées comme relevant d'un domaine de spécialité, pour la lettre B, environ 600, pour la lettre C, environ 1900 et pour la lettre D, environ 900. A titre indicatif, ces mots relèvent des domaines de la zoologie (*acanthoptérygiens*), de la chimie (*aldostérone*), de la géologie (*cône adventif*), des mathématiques (*fonction abélienne*), de la médecine (*albuminurique*), etc. On y trouve encore, pour la lettre D, plus de 80 verbes marqués techniques. Une étude sur le *Petit Larousse* entre 1945 et 1960 avait fait ressortir qu'en une dizaine d'années plus d'un quart du vocabulaire recensé par un dictionnaire avait changé.

S'il y a donc d'une part, tendance à une certaine unification, il y a d'autre part mouvement de renouvellement linguistique puisque de nouveaux mots apparaissent et pénètrent dans le vocabulaire usuel, et mouvement de dénomination scientifique et technique. Les scientifiques et les techniciens pour parler de leurs activités utilisent une langue dont le lexique est nettement différent de l'usage quotidien, en plus de privilégier un certain type de syntaxe. Il semble donc que l'on puisse dégager les traits permettant de distinguer ce que l'on peut appeler *langues de spécialité* ou, mieux, *langues spécialisées*, en regard de ce que l'on nomme *langue commune*.

D'un point de vue assez superficiel, il semble possible d'opposer facilement langue commune à langues spécialisées par le simple fait que l'une est la langue de tous et de tous les jours, l'autre, ou plus exactement les autres, ne s'emploient que dans des circonstances de communication liées à des activités particulières, à des activités faisant appel à des concepts spécifiques partagés. C'est la différenciation de ces activités, et leurs transformations, qui créent les diversifications linguistiques.

Si, il y a quelques années encore, les préoccupations technologiques et tout le vocabulaire qui y était relié étaient considérés comme «peu dignes de l'honnête homme» et de la «culture» et si, dans les dictionnaires, la limite entre langue commune, c'est-à-dire surtout langue littéraire, et langue spécialisée était nettement tracée, de nos jours la banalisation scientifique et technique verse dans le vocabulaire de la langue de tous les jours un nombre de plus en plus considérable de termes ayant perdu quelque peu leur valeur technique ou scientifique. C'est ainsi que des termes comme *bisulfure*, *cryogénique*, *dyslexique*, *électronique*, *transistor*, *nucléaire*, *asymptote*, *enzyme*, pour n'en citer que quelques-uns, sont devenus des mots «ordinaires». On peut donc raisonnablement se demander lesquels de ces mots doivent apparaître dans un dictionnaire de spécialité.

La distinction entre mot scientifique ou technique et mot courant n'est pas toujours évidente. Un mot scientifique ou technique sera considéré comme tel s'il est utilisé dans une communication scientifique ou technique portant sur un sujet scientifique ou technique. En fait, les langues de spécialités, comme on l'a déjà dit, doivent leur existence à la nécessité de dénommer des objets ou des notions qui n'ont pas de dénomination dans la langue commune. Les langues de spécialité se veulent donc un moyen de communication traitant d'un domaine spécialisé par des spécialistes du domaine. On pourrait ici être tenté de conclure que les langues

de spécialité sont tout ce qui est linguistiquement ignoré par la langue commune. Si cela est vrai de façon très générale, il faut cependant faire intervenir toute la problématique de la banalisation des vocabulaires techniques et scientifiques, des niveaux de communication et des liens entre langue commune et langues de spécialité.

Il est évident que les langues de spécialité ne sont pas des entités autonomes, même si la plupart des utilisateurs de ces langues ont subi une certaine initiation : «toute langue technique conserve le caractère initiatique inclus dans l'obligation de n'accéder au secret des mots qu'après avoir été instruit du secret des choses» écrit Bernard Quemada (1978 : 1148). En effet, les langues de spécialité utilisent les mêmes moyens lexicologiques, morphologiques, syntaxiques que la langue commune et conservent des rapports étroits avec celle-ci. Il existe même des échanges constants entre les deux. Nous devrions toujours utiliser l'étiquette *langues de spécialité* au pluriel pour bien montrer qu'il y a des langues différentes selon les domaines de spécialité, mais aussi selon les niveaux de spécialité et toute la zone de transition entre elles et la langue commune. On peut donc, comme le font Robert Galisson et Daniel Coste (1976 : 307-308) distinguer :

- *les langues scientifiques* qui s'utilisent pour rendre compte d'un domaine de «champs d'expérience caractérisés par un objet ou une méthode d'investigation déterminés et de connaissances fondées sur les relations objectives vérifiables»;
- *les langues techniques* qui renvoient aux domaines des applications pour décrire les produits et les productions, les opérations, les procédés, les moyens de production, les agents de production, les producteurs, etc.;
- *les langues professionnelles ou de métiers* qui décrivent les champs d'expérience que les individus exercent soit de façon manuelle ou mécanique.

En plus de cette classification, il faudrait distinguer les «champs d'expérience non professionnalisés» et de passe-temps. Galisson et Coste réservent l'appellation «langues spécialisées» aux champs d'expérience non professionnalisés (par exemple, le football) et désignent par «langues vulgarisées» ou «langues banalisées» les langues qui permettent l'acquisition d'informations aux non-spécialistes.

Pour être honnête, il faut rappeler que cette classification n'a pas pour but de montrer que les domaines d'activité sont très clairement tranchés. En fait, il n'en est rien, et de moins en moins puisque les spécialités s'interpénètrent et se chevauchent souvent. Où s'arrête une spécialité et où commence une autre ? La chimie ne fait-elle pas appel à la physique, à la biologie ? Est-ce que la biochimie est totalement différente de la biologie et de la chimie ou est-ce un amalgame, et dans quelles proportions des deux ? Et que dire de la médecine, de l'anatomie, de la biologie, de la physiologie, de la chimie, du génie chimique ? Il en est de même pour la distinction entre langue scientifique et langue technique. La frontière reste floue et ce qui appartient à la langue scientifique se retrouve aussi dans la langue technique. Comme l'ont montré J.-L. Descamps et A. Phal (1978), si *coefficient*

est un mot du vocabulaire scientifique, *coefficient d'absorption totale linéaire* est un mot du vocabulaire technique.

Que faut-il alors répondre à la question que posait Louis Guilbert : «Peut-on transcender ces spécificités pour conclure à une homogénéité de tous en vocabulaires techniques et scientifiques, aboutissant ainsi à une langue particulière dans l'ensemble de la langue ?»

Il est vrai que si l'appellation *langue de spécialité* fait problème, comme le signale Bernard Quemada lorsqu'il écrit que la linguistique descriptive «condamne les désignations» de langues technique et scientifique qui sont également impropres, il convient plutôt de parler de vocabulaires, s'agissant d'emplois particuliers du français et de ses variétés, qui font appel, pour la prononciation, la morphologie et la syntaxe, au fond de la langue commune, caractère à partir duquel notre analyse établit des rapports fonctionnels mais non hiérarchisés. L'appellation *langue commune* n'est pas non plus tout à fait sans ambiguïté. On trouve en effet encore les adjectifs courant, général, usuel pour désigner la langue de tous, mais il est très difficile de circonscrire exactement la notion de langue commune. On pourrait dire qu'il s'agit de la langue utilisée par tous, c'est-à-dire le noyau commun, mais cela ne règle rien puisqu'en effet, on peut théoriquement mesurer la performance des locuteurs, mais non leur compétence linguistique. Peut-être faut-il simplement admettre l'existence de la langue commune et reprendre la formulation de R. Lagane lorsqu'il écrit : «...c'est la somme de (...) larges aptitudes à l'intercompréhension. (...) Sa réalité est prouvée par les faits, mais il est vain d'espérer en définir les contours avec précision.»

Cette même problématique se retrouve dans les dictionnaires intralinguistiques de langue ou spécialisés. On voit en effet que la multiplicité des domaines de la réalité humaine crée une variété et un enchevêtrement de domaines pour le lexicographe. Tous les dictionnaires intralinguistiques deviennent alors des «dictionnaires restrictifs» ou sectoriels, et leurs limites sont en fait basées sur la définition d'un «usager théorique» et d'un domaine d'étude regroupé.

L'usager théorique est souvent défini de la façon la plus large possible, ce qui permet commercialement de «ratisser» large et justifie ainsi temps et investissement. Il n'en reste pas moins qu'il y a une variété d'usagers potentiels et que l'auteur de dictionnaire à tout intérêt à bien juger les caractéristiques de sa «clientèle» pour déterminer sa nomenclature. Il est clair que les dictionnaires pour les utilisateurs généraux, les étudiants du domaine de spécialité, les traducteurs ou les spécialistes du domaine seront nettement différents tant par la nomenclature que par l'information fournie et même par la métalangue utilisée.

### 6.2.1 Le choix de la nomenclature

La détermination de la nomenclature est une étape importante dans la rédaction d'un dictionnaire. Il faut rappeler que nous ne traiterons pas ici du dictionnaire

interlinguistique de langue commune, où le choix de la nomenclature se pose de façon différente : il faut déterminer quels mots scientifiques et techniques sont banalisés pour les inclure dans la nomenclature. Nos observations ne concernent que les dictionnaires spécialisés. Comme on l'a déjà dit, le choix est fixé par l'usager théorique, mais aussi par le secteur étudié. La première décision du lexicographe sera de réunir tous les mots techniques pour son champ d'études. Mais dès ce moment, il est confronté à une première difficulté : quels mots retenir, quels critères établir pour faire ce choix ?

Les mots techniques comprennent les mots spécialisés qui relèvent du domaine d'étude mais qui appartiennent aussi, du moins certains, à la langue générale et peuvent en plus participer à la formation de terminologismes, c'est-à-dire ces mots techniques, très spécialisés, hiérarchisés et généralement étoffés par une série de spécificateurs. Ces derniers constituent ce que l'on appelle parfois des «composés lourds» du type *charrue pour labours à plat à traction animale sans avant-train*, pour reprendre un exemple souvent cité. Des exemples de ce type ne sont pas rares. On pourrait citer : *ten year level premium term insurance = assurance temporaire de dix ans à prime nivelée; reusable 10-micron corrugated stainless-steel wire cloth element = élément filtrant réutilisable en acier inoxydable ondulé de 10 microns; identical three-position four way solenoid-operated selector valve = robinet sélecteur identique à solénoïde, à quatre voies, à trois positions*. Le lexicologue devra donc se demander si sa nomenclature doit inclure des composés lourds comme ceux cités ci-dessus ou même des termes du type *marteau*, avec les divers types de marteaux : *marteau à palissier, marteau de géologue, marteau à piquer, marteau à emboutir, marteau à gorge, marteau à rentrer, marteau à dresser, marteau à main, marteau à enfoncer les pointes, marteau à ébarber, marteau à crochet pour pompier, marteau de couvreur, marteau à frapper devant, marteau de maçon, marteau à charbon, marteau de paveur...* sans oublier les 15 autres variétés et les synonymes *asseau, besaiguë, boucharde, chasse, ferretier, mailloche, masse, massette, matoir, merlin, picot, rustique, smile...* en fournissant bien entendu les domaines de spécialisation auxquels ils correspondent. En plus, faut-il donner les mots de la même «famille», soit *martelage, martèlement, marteler, martelet, marteleur, marteline* et *martellerie*, par exemple ? Là encore la connaissance des besoins du «destinataire» du dictionnaire est primordiale. En plus, la limite entre des domaines différents n'est pas étanche : un même terme peut appartenir à plus d'un secteur. Un exemple est celui de l'informatique qui touche, à l'heure actuelle, sans exception à tous les domaines. D'un autre côté, il reste à régler la question de la délimitation des domaines. En effet, comment les circonscrire ? Combien de domaines recenser ? Quels domaines retenir ? Faut-il être très spécifique ou plus convivial ? Les emplois spécialisés sont de plus en plus nombreux, mais faut-il dans un dictionnaire marquer tous les usages ? N'y a-t-il pas risque de semer la confusion dans l'esprit de l'usager ? Un dictionnaire polytechnique, par exemple, ne peut, à notre avis, recenser ni tous les domaines ni même un trop grand nombre de domaines, sous peine d'être obligé de

multiplier les références et de fournir ainsi une «surspécialisation» et de gêner plus que d'aider un certain usager non spécialiste de tous les secteurs. Le lexicographe se doit donc de regrouper les indications sous de grandes rubriques. A notre avis, on peut difficilement dépasser environ quatre-vingts domaines sans créer une certaine redondance et un risque de confusion. Il sera alors plus prudent, de tout marquer *CHIMIE* et *GENIE CHIMIQUE*, par exemple, même s'il s'agit de termes qui relèvent d'un domaine plus spécialisé, de la *chimie organique* de la *chimie des polymères* ou de *chimie des solides*.

### 6.2.2 L'équivalence

Comme on l'a déjà dit, le problème de l'équivalence se pose très souvent. En reprenant les exemples de Roger Goffin (1985 : 18-19) tirés de la terminologie des aciers, on peut montrer que, même si les découpages de l'expérience suivant le critère linguistique ne sont pas superposables d'une langue à l'autre, toutes les langues arrivent, lorsque la nécessité les oblige à le faire, à la même précision quant à la description de la réalité. Bien sûr, les moyens linguistiques sont souvent différents, mais peut-il en être autrement puisque les langues utilisent leurs «potentialités» morphologiques et syntaxiques pour suppléer aux lacunes lexicales. Ainsi, si l'allemand a *Erwärmdauer*, le français note *durée de mise en température* ou *durée de chauffage* qui est le «temps écoulé entre le début du chauffage et le moment où l'on estime que la température recherchée est atteinte, à cœur ou dans la partie de la pièce qui doit être traitée». Cette opération est découpée par l'allemand en : 1) *Anwärmdauer* : «durée au bout de laquelle la température visée est atteinte en surface» (= *durée de chauffage*) et 2) *Durchwärmdauer* : «laps de temps compris entre le moment où la température est atteinte par la surface et celui où elle est atteinte par le cœur». De même l'allemand distingue entre *Überhitzen*, si la surchauffe résulte d'une température trop élevée, et *Überzeiten*, si le chauffage est d'une durée trop longue. Le français n'a que *surchauffe* qui est un «chauffage effectué dans des conditions de température et de durée telles qu'il se produit un grossissement exagéré du grain». L'allemand distingue encore pour *Entkohlung*, qui est la *décarburation*, c'est-à-dire la réduction de la teneur en carbone sous l'action du milieu extérieur, *Auskohlung*, qui est la *décarburation totale* ou quasi totale, et *Abkohlung* ou *décarburation partielle*. La perception de la réalité n'est pas imposée par la langue, mais la langue facilite la prise en charge linguistique de cette différenciation. Il en est de même pour *Ausethern*, et toute la série de même formation, que le français doit rendre par *extraction à l'éther* ou *extraction à...* Dans toutes les paires de langues, on peut sans difficulté trouver quantités d'exemples semblables mais encore une fois, si le lexique est insuffisant, la syntaxe y supplée. Si l'allemand parle de *Augenpulver* pour désigner quelque chose qui est écrit en «fine print», on peut rendre cette même réalité en français, par exemple, par *texte en très petits caractères* ou *texte en microcaractères*. La correspondance permet

d'évoquer la même réalité, la communication fonctionne et le message est relayé. Les langues ont des moyens différents, il est banal de le dire, l'une exploite davantage tel ou tel procédé morphologique, l'autre favorise le moyen syntaxique, l'une explicite, l'autre garde une partie implicite, mais le comportement final du locuteur est identique. Nous avons déjà fait remarquer que le syntagmatique supplée aux défaillances du paradigmatique. Le panneau allemand (qui devra obligatoirement choisir entre *Einfahrt* ou *Ausfahrt*) aura, par exemple, *Ausfahrt freihalten*, c'est-à-dire «Garder la sortie libre», ce que le français notera simplement par *Sortie de véhicules*. Manque de précision, surcaractérisation ? Qu'importe, le message est le même ! De même, mettre en relief la différence entre *Lebensgefahr* et *Danger de mort* n'apporte pas grand chose, puisqu'il y a équivalence de situation même si pour les uns l'accent est mis sur «le danger menaçant la vie» et pour les autres sur «le risque de perdre la vie». L'allemand dira *Ausbaggerung*, c'est-à-dire marquera clairement l'instrument de l'action, alors que le français notera simplement *excavation*. Si la précision est indispensable, le français peut l'indiquer syntaxiquement tout comme les autres langues. Doit-on en tirer des conclusions ? Le dictionnaire n'a pas à plonger dans la psychologie profonde pour mettre en relief une différence de «vision» du monde, mais à fournir des correspondants qui peuvent et doivent apporter la contribution sémantique dont le lexique a besoin pour permettre une communication utile. Si l'anglais fait une différence, parfois vitale, entre *president* et *chairman* d'une assemblée, le français n'ayant en principe que *président*, il y a cependant moyen de marquer la différence, si elle est pertinente, par des «ajouts». Ce que le lexicographe doit retenir, et l'usager savoir, c'est que les dictionnaires donnent des équivalences à différents niveaux.

Nous ne voulons cependant pas laisser croire que le lexicographe a toute liberté de création. Il n'en est rien. Bien au contraire, il y a un usage consacré, une terminologie parfois imposée par des normes qu'il appartient au lexicographe de noter et de recenser. Le lexicographe est d'abord le «recenseur de l'usage». Un *Ausfallast* est en français un *nœud non adhérent* ou un *nœud bouchon*, selon la norme NF B50-002, et un «bon» dictionnaire se doit de l'indiquer.

La liberté du lexicographe est, comme on vient de le voir, conditionnée et sa réaction face aux possibilités d'équivalences entre les signes de la langue source et de la langue cible sont celles d'une recherche d'équilibre où les «pertes» et les «ajouts» ne sont que minimales ou minimalisés dans toute la mesure du possible. Le lexicographe doit donc constamment «peser» et «mesurer». Bien entendu, dans le domaine des langues spécialisées son travail de comparaison est partiellement facilité par la prise en charge d'une «réalité plus technique ou plus scientifique» qui se transpose plus facilement que dans le domaine où la langue est avant tout un agrégat historique, en ce sens qu'il y a une accumulation culturelle fort ancienne et qui a laissé dans la langue de nombreuses connotations. Le domaine des langues de spécialité, là encore, il faudrait préciser certains domaines, car l'éventail des spécialités va du plus proche aux sciences humaines

au plus loin. Quoi qu'il en soit, les comparaisons du lexicographe lui permettent de dégager, de façon très schématique, les cas suivants :

- la réalité pour la langue source et la langue cible est la même;
- la réalité n'existe que dans la langue cible;
- la réalité permet des équivalences approximatives entre la langue source et la langue cible.

Dans le premier cas, on peut dire qu'il y a identité de perception de la réalité et la langue cible a un signe linguistique qui est l'équivalent parfait du signe linguistique de la langue source. C'est bien entendu le cas idéal où la dénotation est identique. Les exemples ne manquent pas : *Ausgleichspule* = *bobine de compensation* (domaine : électrotechnique); *Pfahlgründung* = *fondation sur pilotis* (domaine : construction); *parachuting* = *parachutage*. Mais déjà on peut percevoir des difficultés dans certains exemples. Ainsi, *fondation sur pilotis* a des synonymes *fondation sur pieux* ou *fondation sur radier général*. On doit donc se demander quel est le terme le plus courant (la fréquence des termes synonymes peut être différente et la fréquence de l'équivalent peut aussi varier). L'équivalence n'est donc que très théorique; elle relève de la langue et non de la parole. La même observation peut se faire avec *parachuting* qui, bien que l'équivalent de *parachutage*, est cependant nettement moins usité; les dictionnaires ne s'y trompent pas puisqu'ils ajoutent comme synonymes *dropping by parachute* et *landing by parachute*. Le lexicographe se tire donc de ce mauvais pas en accumulant les équivalents dans la langue cible (mais voir Grundy, dans ce volume). La traduction de *abgasentgiftetes Fahrzeug* devient ainsi *véhicule propre* ou *véhicule non polluant* où le premier terme laisse au contexte le soin de la désambiguïsation.

Dans le cas où la réalité n'existe pas dans la langue cible, le lexicographe doit trouver un équivalent approximatif en glosant l'information manquante. Il s'agit très souvent de termes liés à la vie culturelle. Bien entendu un emprunt est toujours possible, mais là encore il y a lieu de fournir quelques éclaircissements pour orienter convenablement l'interprétation de l'utilisateur et lui suggérer les associations sémantiques indispensables. Le lexicologue cherche à créer l'équation entre la dénotation (et la connotation) des deux langues. On peut citer ici *boxingday*, c'est-à-dire le «lendemain de Noël». On peut encore évoquer toute la question de la difficulté de rendre les termes juridiques entre les systèmes de pays différents. Comment rendre l'équivalent de *common law* ? Faut-il gloser par *droit civil* ou *droit du citoyen* ou tout simplement faire un emprunt ? Est-ce que *African satinwood* est l'équivalent d'*olovongo* ? N'y a-t-il pas plutôt équivalence «rapprochée» par le choix d'un terme «sémantiquement proche» du lexique (*satinwood*) et un ajout corrigeant la déviation (*African*) ou explicitant. Ainsi, *œufs en neige* correspond à *stiffly beaten eggwhites*. Dans des cas similaires, l'information linguistique est souvent redressée par un renseignement de nature encyclopédique.

Dans le cas où la réalité permet des équivalences approximatives, le lexicographe est lié aux potentialités de formation de mots de la langue cible. On sait, par exemple, que la morphologie de l'allemand permet une composition facile de «mots nouveaux»; et le français n'a à ce moment d'autres recours que l'explicitation, l'emprunt étant impossible à cause de la forme du mot créé (rejet de l'accumulation de consonnes). Ainsi le français ne peut emprunter *Dienstmarscheinweihungsschritt* et doit donc gloser un tel mot par *pas ni trop rapide ni trop long que l'on utilise pour inaugurer les monuments*.

## 7 Conclusions

On pourrait être amené à penser que les dictionnaires bilingues ne réalisent que difficilement la correspondance entre les lexies ou les termes de la langue source et de la langue cible, mais cela donnerait une vision fautive de la lexicographie spécialisée bilingue. En fait, dans la majorité des cas, la défaillance de l'équivalence n'est pas insurmontable. Elle est plus marquée dans les domaines qui relèvent des sciences humaines où la réalité a imposé une décision linguistique à importance culturelle marquée. La langue montre dans ces secteurs des activités humaines (droit, politique, structuration sociale, etc.) une empreinte historique importante qui a en quelque sorte figé son évolution. Dans le domaine technologique ou scientifique, les divergences ne sont pas si marquées et les équivalents, parfois «redressés» par un «complément d'information» plus ou moins encyclopédique ou un étiquetage métalinguistique bien déterminé apporte la solution sinon tout à fait exacte, du moins correcte.



## Chapitre 12

# LA LEXICOGRAPHIE BILINGUE CATALANE CONTEMPORAINE : ANALYSE ET ÉVALUATION

M.Teresa CABRÉ et Cristina GELPÍ

### 1 Quelques considérations autour des dictionnaires bilingues

La lexicographie catalane dispose d'un bon nombre d'œuvres représentatives qui permettent de la situer à un excellent niveau scientifique. C'est en tout cas l'opinion de la majorité des linguistes et philologues qui ont étudié ce sujet.<sup>1</sup> Cela est vrai si l'on considère la qualité des œuvres lexicographiques les plus représentatives et le grand nombre de dictionnaires édités, surtout si l'on tient compte du poids démographique de la zone catalanophone. Dans ce contexte, il faut souligner l'importance quantitative des dictionnaires bilingues et, spécialement, des dictionnaires bilingues qui traitent à la fois les langues catalane et espagnole.<sup>2</sup>

Dans cet article, nous utiliserons la notion de dictionnaire telle qu'elle est définie par Cabré et Lorente (1991) : «œuvre lexicographique qui contient une sélection des mots d'une langue, sur lesquels elle offre des informations variées et qui présente ces mots ordonnés à partir d'un système préétabli.» Ainsi, nous excluons explicitement de cette notion : les œuvres qui ne sont pas des recueils d'unités linguistiques; les œuvres d'extension très limitée, ou dont le format ne correspond pas à ce que l'on considère comme un dictionnaire; les encyclopédies; les glossaires de termes qui ne contiennent pas d'autres informations que le mot clé; les œuvres non éditées sur papier; les œuvres sans indépendance physique,

---

1 Ainsi Germà Colon affirmait en 1976, à l'occasion du IV Colloqui Internacional de Llengua i Literatura Catalanes, que la lexicographie catalane a atteint un très haut niveau de perfection, en faisant référence à cinq œuvres lexicographiques catalanes fondamentales : le *Diccionari català-valencià-valenc d'Alcover* et Moll, le *Diccionari Aguiló*, le *Diccionari Batari*, le *Glossarium mediae latinitatis Cataloniae* et le *Diccionari etimològic i complementari de la llengua catalana* de Coromines, en ce moment en phase d'élaboration (Cabré, 1995).

2 Nous utiliserons la dénomination 'espagnol', bien que les dictionnaires que nous analyserons aient généralement pour titre la dénomination 'castillan'.

tels que les recueils de mots qui apparaissent comme glossaire à la fin d'une autre œuvre; et les œuvres non éditées, à l'exception des œuvres publiées à l'usage d'entreprises ou de centres académiques.

En outre, nous croyons que les dictionnaires peuvent être inclus dans des typologies lexicographiques, établies en fonction de la déviation des divers types de dictionnaires par rapport à un patron lexicographique de base qui constituerait le dictionnaire de langue générale (voir aussi Marelllo, et Clas, dans ce volume). Ce patron de référence aurait les caractéristiques suivantes :

- les sources d'où sont tirées les informations sont généralement écrites
- les entrées sont sélectionnées suivant le critère des formes les plus usuelles
- la forme des entrées est lexématique
- les entrées sont ordonnées par ordre alphabétique
- les informations fournies dans l'entrée sont généralement grammaticales, sémantiques (définition, ou équivalence dans les dictionnaires bilingues), acceptions sémantiques et exemples
- la fonction principale du dictionnaire est descriptive
- les destinataires sont le grand public
- la fonction spécifique du dictionnaire est d'augmenter la compétence de l'usager et de lever les doutes ou de combler les lacunes linguistiques

Nous considérerons comme spécifique tout dictionnaire qui s'éloignera de ces caractéristiques. Bon nombre des dictionnaires que nous analyserons sont dans ce cas; il s'agit, en définitive, des recueils de terminologie, ou des dictionnaires spécialisés, fondamentalement, par le domaine. Nous résumerons ci-dessous quelques-unes des caractéristiques principales de ce type d'œuvres :

- les sources utilisées ne sont pas exclusivement écrites; ils font aussi usage des sources orales,
- la sélection de la nomenclature se fait sur la base du domaine et non pas nécessairement sur celle de la fréquence,
- la fonction principale de ces œuvres est la normalisation, au sens de standardisation, des termes propres d'un domaine spécialisé,
- le travail se fait généralement selon le processus onomasiologique.

Indépendamment des caractéristiques concrètes des dictionnaires, nous concevons le dictionnaire bilingue comme le type de dictionnaire qui contient deux langues historiques, dans lequel les mots de la langue source sont accompagnés de leurs équivalents dans la langue cible, et qui a comme objectif principal de coordonner, avec les unités lexicales d'une langue, les unités de deux langues entre lesquelles il existe une équivalence dans la signification. En plus de cet objectif, les dictionnaires bilingues en ont beaucoup d'autres, déterminés par le développement de la société : ils sont fondamentaux dans l'apprentissage des langues depuis l'enseignement primaire jusqu'à l'université; ils sont utiles comme appui pour la communication en langues étrangères, dans le monde commercial et dans l'administration publique; indispensables pour les traduc-

teurs et interprètes. Les dictionnaires de traduction spécialisés sont aussi un élément essentiel dans la communication spécialisée entre compagnies, institutions publiques et organisations internationales.

Il est évident que les caractéristiques des dictionnaires bilingues ne diffèrent pas beaucoup de certaines caractéristiques des dictionnaires monolingues, parce qu'il est bien connu que toute œuvre lexicographique sert de pont entre ce que l'utilisateur connaît et ce qu'il ne connaît pas. La connection entre ce qui est connu et ce qui est inconnu est représentée de façon différente, selon le type de dictionnaire. Ainsi, dans les dictionnaires monolingues, la définition a presque toujours la même structure : un hyperonyme qui indique la catégorie à laquelle appartient le mot et un identificateur de sens qui apporte les attributs spécifiques de l'entrée. Dans les encyclopédies, le corps de l'article offre les informations qui connectent l'entrée avec le monde réel et ces informations ont pour finalité d'apporter un niveau déterminé de culture à l'utilisateur. Dans les dictionnaires bilingues, en revanche, la connection entre le connu et l'inconnu est représentée par la traduction d'un mot d'une langue dans une autre langue.

Du point de vue de l'utilisateur, la traduction d'un mot est possible, parce qu'il ne doute pas que la traduction existe toujours; mais les problèmes de traduction sont réels et vastes, parce que souvent la réalité extra-linguistique à laquelle renvoie le signifié d'un mot a un sens dans la langue des locuteurs, mais le mot que cette réalité désigne n'existe pas dans l'autre langue, ou n'existe pas avec les mêmes connotations (voir Szende, et Clas, dans ce volume).

En dehors des problèmes connus de la traduction, il existe, aussi, dans les dictionnaires bilingues, une difficulté supplémentaire relative aux besoins des usagers possibles des dictionnaires. Selon Ščerba (1971), l'utilisateur d'un ouvrage lexicographique peut en avoir besoin, fondamentalement, pour comprendre ou pour produire un texte; et c'est pour cela qu'on peut utiliser les dictionnaires dans quatre situations différentes, selon la langue et l'utilisateur. Et vu que les informations dont l'utilisateur a besoin pour coder ou pour décoder des textes sont différentes, il paraît raisonnable de penser que la distinction entre dictionnaires de production et de compréhension, actifs ou passifs, est la meilleure pour refléter les besoins des usagers. Nous partageons l'idée qu'il n'est pas possible d'accorder la même attention aux usagers des deux langues d'un dictionnaire bilingue en même temps. Ainsi, les informations que contiennent les dictionnaires de production devraient être différentes de celles des dictionnaires de compréhension.<sup>3</sup>

3 Quelques exemples de ces différences : l'accent dans les dictionnaires de production est mis sur l'information interne de l'article, tandis que dans les dictionnaires de compréhension il est sur le mot-clé; la sélection des entrées des dictionnaires de production admettrait des dérivés et des composés, même si son signifié est transparent, alors que ce genre d'entrées n'est pas nécessaire dans un dictionnaire de compréhension; dans la présentation des mots-clés des dictionnaires de production les conditions d'usage sont très importantes, alors que dans les dictionnaires de compréhension on favorise l'accumulation d'équivalents, etc.

Le présent chapitre est centré sur l'analyse et sur l'évaluation des dictionnaires bilingues qui traitent le catalan et l'espagnol ensemble, indépendamment du fait que l'un ou l'autre soient langue source ou langue cible, parce que nous constatons que ce groupe de dictionnaires vient en deuxième position dans la hiérarchie des dictionnaires les plus fréquents de la langue catalane.<sup>4</sup>

Dans ce chapitre, nous procéderons à une analyse critique de la lexicographie bilingue catalan > espagnol ou espagnol > catalan, en montrant d'un côté les aspects positifs de son évolution et, de l'autre, les lacunes qu'elle présente, dont beaucoup sont explicables par l'évolution politique et sociale de la langue et de la société du domaine catalan dans sa globalité.<sup>5</sup>

## 2 Aspects historiques

Nous devons partir du constat que les dictionnaires bilingues catalans n'apparaissent pas dans notre panorama lexicographique au moment où nous en commençons l'étude ici. Cependant nous disposons d'échantillons représentatifs dès les X<sup>e</sup> et XI<sup>e</sup> siècles et, progressivement, et de façon continue, jusqu'au XX<sup>e</sup> siècle. Ces échantillons nous permettent d'affirmer que la tradition lexicographique bilingue catalane commence très tôt et avec des œuvres concrètes, bien que, dans la plupart des cas, on ne puisse pas parler de dictionnaires bilingues dans le sens que l'on donne à ce terme aujourd'hui. C'est pour cela que nous considérons qu'avant 1940, il existe des recueils lexicographiques bilingues que l'on conçoit comme les précurseurs des ouvrages lexicographiques bilingues actuels.

Aux origines de la lexicographie bilingue catalane, le type de dictionnaires que nous analysons ici ne prédomine pas. Les dictionnaires bilingues catalan et espagnol n'ont pas été les premiers à apparaître dans le panorama lexicographique catalan. En effet, les premiers échantillons lexicographiques que nous avons, les gloses (qui avaient pour objet d'expliquer les mots obscurs à partir de synonymes ou de périphrases), ont comme autre langue le latin. Dans le même groupe nous incluons les *rimari*, les *gramaticae proverbiandi* ou les exemples médiévaux d'explication de mots (voir Boisson, dans ce volume).

Les précurseurs de la lexicographie bilingue catalane, qui se sont manifestés par des œuvres concrètes et à partir du XIV<sup>e</sup> siècle, n'avaient pas non plus comme langue complémentaire prioritaire l'espagnol, mais la langue qu'ils décrivaient était aussi le latin. On pourrait citer comme exemples le *Comprehensorium* de

4 Le type le plus fréquent est constitué par les dictionnaires terminologiques monolingues ou multilingues, mais pas nécessairement bilingues dans ces deux langues.

5 Le corpus des données qui constituent la base de cet article est le catalogue de dictionnaires de Cabré et Lorente (1991), enrichi jusqu'en 1990. Ce catalogue prétend être un inventaire exhaustif des dictionnaires catalans publiés dans tout le domaine catalanophone. Il a été complété par deux appendices non publiés : 1989-1990 et 1991-1992.

Johannes, le *Liber elegantiarum* de Joan Esteve, le *Nebrija català*, etc. Toutes ces œuvres étaient destinées à clarifier pour les usagers catalans le signifié des mots de la langue latine. Déjà au XVI<sup>e</sup> siècle, plusieurs œuvres lexicographiques bilingues furent publiées avec le même objectif, celui de permettre l'apprentissage de la langue latine aux étudiants catalans. Quelques échantillons en seraient le *Thesaurus purerilis* d'Onofre Pou, le *Diccionari llatí-català* d'Antoni Font, le *Thesaurus* de Pere Torra ou le *Gazophylacium* de Joan Lacavalleria i Dulach.

Ce n'est qu'au XVIII<sup>e</sup> siècle, et bien que la langue catalane n'ait pas encore retrouvé son statut, qu'apparaîtront les premiers essais de la lexicographie bilingue de l'ensemble catalan et espagnol; on publie le *Vocabulari català-castellà de mots antics* de Fra Manuel Marià Ribera et les *Diccions espanyoles explicades en català* de Baldri i Reixach (œuvre qui prétend recueillir par ordre alphabétique les mots espagnols plus éloignés du catalan et plus compréhensibles pour les enfants qui étudiaient la langue). Pendant ce siècle, l'activité lexicographique la plus importante est celle de Carles Ros, notaire de Valence, qui publia plusieurs dictionnaires *valenciano-castellano*.

Finalement, dans la majorité des recueils lexicographiques édités au cours du XIX<sup>e</sup> siècle, l'élément commun est l'espagnol. On édite des dictionnaires valenciens, majorquins et minorquins, avec la correspondance en espagnol, et le latin, qui avait été présent dans la tradition lexicographique, disparut pratiquement. Quelques-unes des œuvres les plus représentatives sont le *Diccionari castellà-català* de fra Magí Ferrer (inspiré du dictionnaire de l'Academia Española) et aussi le *Diccionario castellano-catalán* du même auteur. Valence devint le centre des principales impressions de dictionnaires pendant ce siècle : *Ensayo de un Diccionario valenciano-castellano*, de Lluís Lamarca (liste d'un millier de mots qui part du catalan et donne l'équivalent espagnol); *Miscelánea* de Miquel Rosanes (œuvre d'environ six mille six cents mots catalan-espagnol); le *Vocabulari valencià-castellà* de Josep M. Cabrera (près de deux mille entrées, ordonnées alphabétiquement). Un des dictionnaires valenciens les plus prestigieux est le valencien-espagnol de Josep Escrig, dont la première édition apparut en 1851 (et contient 90 000 entrées, bien que la qualité du matériel soit un peu douteuse). Joaquim Martí i Gadea est l'auteur du *Novísimo diccionario general valenciano-castellano*, publié en 1891 (1976 pages, mais la qualité du matériel est aussi relative). Les îles Baléares constituent également un lieu d'activité lexicographique importante pendant le XIX<sup>e</sup> siècle : le *Diccionari mallorquí-castellà*, publié en 1840 par Pere Antoni Figuera (recueil de mots majorquins définis en catalan et suivis par l'équivalent espagnol); en 1869, Josep Hospitaler i Caballer publia le *Vocabulario castellano-menorquín y vice-versa* (qui n'est qu'un recueil de mots groupés thématiquement à partir de l'espagnol). L'objectif de ce dictionnaire est, comme dans la majorité des cas, de diffuser la connaissance de l'espagnol. Il faut aussi mentionner le *Vocabulario menorquín-castellano* de Joan Benejam, en 1885.

Pour finir cette revue des principaux dictionnaires bilingues catalan et espagnol publiés avant 1940, nous voulons mentionner l'apport, pendant la

deuxième partie du XIX<sup>e</sup> siècle, de trois érudits catalans : Aguiló, Balari et Alart, qui conférèrent à la lexicographie catalane un caractère scientifique inconnu jusque-là. Selon les méthodes historico-comparatives de la linguistique, ils analyseront la langue ancienne, et feront un gros travail de recueil de matériaux qu'ils essayeront de revitaliser pour revigorer la langue catalane.

Le premier, Marià Aguiló, recueillit les matériaux qui serviraient à la création du dictionnaire qui porte son nom, le *Diccionari Aguiló*, édité plus tard par Pompeu Fabra et Manuel de Montoliu, entre 1914 et 1934, en 8 volumes. Bien qu'il s'agisse d'un dictionnaire de langue générale, il fournit dans certains cas, au lieu de définitions, des équivalents en espagnol, et parfois aussi en français.

Le deuxième, Balari, nous a laissé le *Diccionari Balari* qui est, comme le précédent, un recueil posthume des matériaux recueillis par l'auteur. Il part d'une documentation ancienne ou contemporaine; dans l'avertissement aux lecteurs il fait remarquer que l'œuvre sera intéressante pour connaître l'équivalent exact des mots en catalan et ceci est une des caractéristiques fondamentales du dictionnaire, par rapport au recueil d'Aguiló. Le dernier, Julià-Bernat Alart, donna aussi l'*Essai de dictionnaire historique de la langue catalane*, où il traduit chaque mot en français.

### 3 De 1940 à 1990

Le nombre des dictionnaires bilingues catalan et espagnol publiés entre 1940 et 1990 est de 108, un chiffre important si nous tenons compte de ce que, dans cette période, 281 dictionnaires ont été publiés (et, par conséquent, les dictionnaires bilingues catalan et espagnol représentent plus d'un tiers de la totalité des dictionnaires catalans); il faut également penser au poids démographique de la société catalanophone et, aussi, aux vicissitudes politiques de 1940 à 1977.<sup>6</sup>

Cette production est répartie irrégulièrement tout au long de la période 1940-1990 :

ŒUVRES PUBLIÉES	CHIFFRE ABSOLU	%
1940-1950	0	0 %
1951-1960	3	2,2 %
1961-1970	15	11,4 %
1971-1980	35	26,7 %
1981-1990	78	59,5 %

<sup>6</sup> Nous incluons ici non seulement les dictionnaires strictement bilingues catalan et espagnol, mais aussi les terminologies qui traitent le catalan et l'espagnol ensemble.

Ce tableau montre comment la production lexicographique suit une progression continue : la décennie de 1940 à 1950 représente 0 % de la production; et celle de 1981 à 1990, 59,5 %. Les vingt dernières années (1971-1990) ont concentré près de 80 % de la production totale, alors que pendant 30 ans (1940-1970) la production n'a été que de 13,6 %.

Dans la période la plus productive (1971-1990), l'édition de dictionnaires est répartie régulièrement et marque une progression :

PÉRIODE	CHIFFRE ABSOLU	%
1971-1980	35	26,7 %
1981-1990	78	59,5 %

Une fragmentation des cinq périodes en segments de cinq ans montre qu'il y a un point à partir duquel la production de dictionnaires augmente spectaculairement :

PÉRIODE	CHIFFRE ABSOLU	%
1940-1945	0	0 %
1945-1950	0	0 %
1951-1955	0	0 %
1956-1960	3	2,2 %
1961-1965	5	3,8 %
1966-1970	10	7,6 %
1971-1975	14	10,6 %
1976-1980	21	16 %
1981-1985	40	30,5 %
1986-1990	38	29 %
TOTAL	131	100 %

En effet, sur les 15 dictionnaires de la décennie 1961-70, 10 ont été publiés entre 1966 et 1970; ces 15 dictionnaires représentent 11,4 % de cette décennie. Le point d'inflexion de la production de dictionnaires bilingues catalan et espagnol se situe donc en 1966 : c'est à partir de ce moment-là qu'une publication en masse de dictionnaires s'est produite.

En outre, cette date coïncide presque exactement avec la date qui concerne l'ensemble des dictionnaires catalans, non seulement les bilingues catalan-espagnol; c'est-à-dire, que la progression dans la publication des dictionnaires

bilingues suit la même ligne que la publication des autres dictionnaires. Nous voyons, dans le tableau suivant, un résumé du panorama lexicographique catalan :

PÉRIODE	CHIFFRE ABSOLU	%
1940-1950	1	0,3 %
1951-1960	4	1,4 %
1961-1970	24	8,5 %
1971-1980	69	24,5 %
1981-1990	183	65 %

De cette façon, on peut établir, à partir de ces données, selon Cabré (1995), trois grandes étapes lexicographiques :

- 1 1940-1965, le désert lexicographique
- 2 1966-1980, le réveil de la lexicographie
- 3 1981-1990, la productivité lexicographique

Ces trois segments de temps offrent le panorama quantitatif suivant :

PÉRIODE	CHIFFRE ABSOLU	%
1940-1965	8	6,1 %
1966-1980	45	34,3 %
1981-1990	78	59,5 %

Une autre observation : après avoir analysé les dictionnaires bilingues sur la base de la langue source et de la langue cible, on peut conclure que les auteurs des dictionnaires bilingues catalans choisissent l'une ou l'autre langue en fonction de la valeur de référence qu'en général on concède à l'espagnol. De la même façon, les dictionnaires bilingues catalans ne peuvent être distribués à partir de la distinction actif-passif, ou production-compréhension :

période / langue	1940-50	1951-60	1961-70	1971-80	1981-90	TOTAL	%
catalan > espagnol, espagnol > catalan	0	0	6	25	33	64	46,2 %
espagnol > catalan	0	0	4	8	23	35	29,6 %
catalan > espagnol	0	3	4	3	22	32	24 %

Selon les données du tableau, le nombre d'œuvres qui prennent le catalan et l'espagnol, l'un après l'autre, comme langue source, est considérablement supérieur au nombre des dictionnaires qui prennent le catalan seul ou l'espagnol seul comme langue source :

LANGUE SOURCE	CHIFFRE ABSOLU	%
catalan > espagnol, espagnol > catalan	64	48,8 %
espagnol > catalan	35	26,7 %
catalan > espagnol	32	24,4 %

En outre, il faut souligner que le nombre de dictionnaires bilingues qui traitent à la fois le catalan et l'espagnol est beaucoup plus important que celui des dictionnaires bilingues qui décrivent le catalan en combinaison avec une autre langue. Nous avons seulement 23 dictionnaires bilingues entre le catalan et une autre langue différente de l'espagnol. La distribution de cette production, par décennie, est la suivante :

DÉCENNIE / LANGUE DES DICTIONNAIRES	1940-50	1951-60	1961-70	1971-80	1981-90	TOTAL
catalan et français	0	0	1	3	1	5
catalan et anglais	0	0	0	1	5	6
catalan et allemand	0	0	0	0	2	2
catalan et italien	0	0	0	0	1	1
catalan et portugais	0	0	0	0	3	3
catalan et russe	0	0	0	0	1	1
catalan et japonais	0	0	0	0	1	1
catalan et hongrois	0	0	0	0	1	1
catalan et latin	0	0	0	0	1	1
catalan et grec	0	0	0	0	1	1

Les données que nous avons présentées autour de l'évolution de la production lexicographique permettent de suggérer quelques conclusions :

- 1 L'importance quantitative de la production de dictionnaires est considérable si l'on tient compte de la situation de la langue et des conditions politiques du contexte pendant une bonne partie de la période analysée.
- 2 On constate une concentration presque exclusive de la production de dictionnaires bilingues catalan et espagnol pendant la période 1966-1990. L'année

1966, presque quinze ans avant le rétablissement des autonomies et de la légalisation de l'usage de la langue catalane, est la ligne de partage des eaux pour la production lexicographique bilingue, et elle coïncide quantitativement avec la production lexicographique catalane générale; ceci indique que la revitalisation de la production des dictionnaires ne commence pas avec la récupération institutionnelle, mais qu'elle a une base fondamentalement sociale.

- 3 Le processus de revitalisation lexicographique étant déjà initié, la récupération des institutions est le facteur qui fait augmenter, régulièrement, la production de dictionnaires bilingues. Les dictionnaires publiés dans la période 1981-1990 représentent 59,5 % du total.
- 4 On observe une supériorité quantitative très remarquable d'œuvres lexicographiques qui traitent parallèlement le catalan et l'espagnol ensemble comme langues de départ des dictionnaires bilingues. Cette observation permet de supposer que les maisons d'édition ont l'intention de combler le manque de dictionnaires spécifiques de production et de compréhension par la publication de deux ouvrages en même temps, pour couvrir les besoins d'un maximum d'utilisateurs possibles.
- 5 Il faut noter le manque évident de dictionnaires bilingues en d'autres langues que l'espagnol, parce que le panorama lexicographique catalan, si l'on exclut les dictionnaires bilingues avec l'espagnol, est extrêmement pauvre.

#### 4 Aspects typologiques

L'analyse en détail de plusieurs types de dictionnaires qui apparaissent entre 1940 et 1990 révèle des manques inexplicables de certains types de dictionnaires, et des répétitions superflues d'autres types.

Si nous examinons les types de dictionnaires bilingues catalan et espagnol on constate tout d'abord qu'il y a deux grands groupes de dictionnaires, dont un peut être divisé en sept sous-groupes différents. En effet, le panorama lexicographique catalan a deux grands types de dictionnaires : les dictionnaires d'orientation générale (que nous avons définis ci-dessus) et les dictionnaires d'orientation terminologique. Ce dernier groupe est homogène, parce qu'il peut être défini par les mêmes caractéristiques (que nous avons déjà données). Le premier, par contre, n'est pas aussi fermé. On peut distinguer plusieurs sous-types différents.

Dans ce but, nous avons établi 8 types lexicographiques, correspondant à des publications existantes.<sup>7</sup> Chaque type représente une sorte de dictionnaires,

<sup>7</sup> L'absence dans la classification de types déterminés de dictionnaires est un indicateur de leur inexistence dans la lexicographie catalane jusqu'en 1990.

définie à partir d'une seule dimension, selon l'aspect le plus important qui permet d'attribuer un dictionnaire à un type déterminé :<sup>8</sup>

- G dictionnaires généraux
- T dictionnaires terminologiques
- D dictionnaires dialectaux
- EN dictionnaires encyclopédiques
- SC dictionnaires scolaires
- L dictionnaires de locutions et phraséologismes
- I dictionnaires idéologiques
- R dictionnaires de rimes

L'analyse du degré de corrélation entre la richesse quantitative des dictionnaires bilingues édités et la diversité que le panorama lexicographique catalan présente dans cette période, nous permettra de voir si cette richesse correspond à une diversité typologique de dictionnaires. Pour aborder cette question, nous nous fondons sur trois observations :

- a la distribution des deux types lexicographiques les plus fréquents
  - b la période à laquelle apparaît le premier exemplaire de chaque type
  - c la fréquence de chaque type
- a Nous partons de l'idée que les dictionnaires bilingues catalan et espagnol sont distribués fondamentalement sous deux formes : les dictionnaires bilingues généraux et les dictionnaires bilingues terminologiques. Cette double distribution pourrait être corrigée par d'autres subdivisions correspondant à des publications d'autres types.

L'évolution de plusieurs types de dictionnaires, par périodes de 10 ans, offre les résultats suivants :

période / type	1940-50	1951-60	1961-70	1971-80	1981-90	TOTAL
généraux		1	8	14	16	39
terminologiques	0	0	2	13	55	70
dialectaux	0	1	4	4	5	14
encyclopédiques	0	0	1	0	0	1
scolaires	0	0	0	1	2	3
locutions	0	0	0	2	0	2
idéologiques	0	0	0	1	0	1
rimes	0	1	0	0	0	1

où la moyenne (X) est 16,4.

8 Il semble qu'une classification multidimensionnelle des types de dictionnaires qui tiendrait compte de la diversité d'informations que chaque dictionnaire contient, au lieu du type prioritaire de dictionnaire qu'il constitue conduirait à des résultats substantiellement différents.

Les données de ce tableau montrent que, sur un total de 8 types de dictionnaires évalués, seulement 2 types dépassent la moyenne, à savoir, par ordre décroissant :

- 1 dictionnaires terminologiques
- 2 dictionnaires généraux

Il paraît logique que les terminologies et les dictionnaires de type général constituent les types les plus fréquents, parce que l'on peut supposer que chaque dictionnaire offre une thématique spécifique (les terminologies), ou contient des langues diverses. Par contre, il faut mentionner la rareté des autres types. Ces résultats permettent de suggérer, si l'on analyse les données indépendamment du contexte dans lequel les dictionnaires ont été produits, un certain manque de rationalité dans la production lexicographique.

L'observation du rythme d'apparition de plusieurs types lexicographiques au long de cette période montre que les deux types de dictionnaires les plus fréquents apparaissent progressivement suivant une courbe ascendante :

période / type	1940-50	1951-60	1961-70	1971-80	1981-90
terminologiques	0	0	2	13	55
généraux	0	1	8	14	16

Pour en terminer avec cette question, nous soulignerons que le panorama lexicographique bilingue catalan est relativement pauvre du point de vue typologique, parce qu'il se manifeste de façon presque exclusive dans seulement deux types de dictionnaires. Cette pauvreté est claire si nous la comparons avec la diversité de types lexicographiques que possède la lexicographie catalane en général. En effet, la langue catalane dispose de types variés comme les dictionnaires visuels, les correctifs, les dictionnaires d'homonymes, de synonymes, d'antonymes, d'argot, de vocabulaire fondamental, orthographiques, de prononciation, entre autres.

- b L'analyse de l'apparition du premier dictionnaire de chaque type (repéré par un astérisque dans le tableau suivant) peut être utile pour déterminer dans quelle période s'est produite une diversification lexicographique plus importante. En effet, les données montrent qu'entre 1960 et 1980 apparaissent tous les types de dictionnaires bilingues catalan et espagnol dont la langue catalane dispose aujourd'hui :

période / type	1940-50	1951-60	1961-70	1971-80	1981-90
généraux		g*	g	g	g
terminologiques			t*	t	t
dialectaux		d*	d	d	d
encyclopédiques			en*		
scolaires				sc*	sc
locutions				l*	
idéologiques				j*	
rimes		r*			

Une interprétation de ces données peut nous faire penser que la capacité de concevoir la publication de nouveaux types de dictionnaires, s'agissant des auteurs et des éditeurs, paraît presque épuisée, et ce n'est pas parce que tous les besoins ont été couverts. La raison est peut-être à rechercher dans les règles du marché : on produit seulement les dictionnaires dont la vente est assurée, bien qu'ils soient déjà sur le marché (souvent sans différence significative). Le conservatisme commercial est imposé, dans ce cas, par le risque économique que pourrait représenter le fait de sortir sur le marché un nouveau type de dictionnaire. Cette position conservatrice s'impose aussi à la volonté de doter la langue catalane d'un inventaire lexicographique bilingue semblable à celui d'autres langues de culture, bien que proportionnelle à la population parlante.

- c Nous observerons, enfin, avec quelle fréquence sont produits les différents types de dictionnaires au long des divers segments de temps pour établir le degré de constance de chaque type lexicographique et, ultérieurement, nous comparerons ces données avec les indices d'importance et de répétition :

LES DICTIONNAIRES BILINGUES

période / type	1940	1945	1951	1956	1961	1966	1971	1976	1981	1986	Total
	1945	1950	1955	1960	1965	1970	1975	1980	1985	1990	
généraux				g	g	g	g	g	g	g	7
terminologiques						t	t	t	t	t	5
encyclopédiques						en					1
dialectaux				d	d	d	d	d	d	d	7
scolaires						sc				sc	2
idéologiques								i			1
locutions							l				1
rimes				r							1

Selon ces données, nous pouvons établir trois séries de types lexicographiques, à partir du degré de constance d'apparition :

1 types constants :

répétition 7 :

dictionnaires bilingues généraux  
dictionnaires bilingues dialectaux

répétition 5 :

dictionnaires bilingues terminologiques

2 types inconstants :

répétition 1 :

dictionnaires bilingues de rimes  
dictionnaires bilingues de locutions  
dictionnaires bilingues idéologiques  
dictionnaires bilingues encyclopédiques

répétition 2 :

dictionnaires bilingues scolaires

Le degré maximum d'inconstance d'un type lexicographique est un indicateur de la nécessité de renforcer ce type lexicographique avec de nouvelles publications de caractéristiques semblables, si l'indice d'importance ne fonctionne pas comme neutralisant. Le degré maximum de constance suppose une répétition gratuite, si l'indice d'importance ne fonctionne pas comme neutralisant.

L'indice d'importance de la répétition est un indicateur qui permet de caractériser les divers types de dictionnaires, selon la redondance de la répétition. Un type de dictionnaire sera redondant si la répétition, au lieu de conduire à un dictionnaire d'information nouvelle, conduit simplement à un dictionnaire du même type avec une augmentation ou une réduction de l'information. Ainsi, deux terminologies, qui appartiennent au même type lexicographique, seront redondantes si elles traitent de la même thématique, bien qu'ayant été conçues pour des usagers différents, et que l'une contienne seulement une sélection des données de l'autre. Par contre, ne seront pas redondantes deux terminologies qui traitent des thématiques différentes dans des dictionnaires bilingues de langues elles aussi différentes.

L'indice d'importance permet de distinguer entre les dictionnaires répétés, ceux qui sont répétés de façon importante, et pourtant, introduisent un nouveau dictionnaire, et ceux qui sont répétés de façon non importante, et sont généralement nécessaires comme type lexicographique.<sup>9</sup>

- 1 dictionnaires bilingues de répétition importante :  
terminologies
- 2 dictionnaires bilingues de répétition non importante :  
catalan > espagnol; espagnol > catalan

Il semble, donc, que les terminologies sont des types importants, parce qu'on peut supposer que chaque exemplaire, dans le cas des dictionnaires bilingues catalan et espagnol, traite une thématique différente. En règle générale, nous pouvons affirmer que cette hypothèse est vérifiée, mais qu'un bon nombre de terminologies bilingues concerne plus d'un domaine en même temps, même si évidemment, ces domaines sont liés. Les informations des tableaux suivants confirmeront, donc, l'hypothèse que la majorité (au moins) des terminologies sont destinées à un seul domaine :

TERMINOLOGIES	CHIFFRE ABSOLU	%
1 domaine sémantique	50	71,4 %
2 domaines sémantiques	12	17,1 %
3 domaines sémantiques	5	7,1 %
7 domaines sémantiques	1	1,4 %
plus de 10 domaines sémantiques	2	2,8 %
TOTAL	70	100 %

<sup>9</sup> Pour cet indice d'importance nous n'avons pas pris en considération la différenciation à partir des usagers auxquels le dictionnaire est destiné. De ce point de vue, quelques-unes des répétitions que nous considérons comme redondantes pourraient ne pas l'être. De toute façon, ce détail n'affecte pas la relation importance-non importance des dictionnaires bilingues catalan et espagnol.

C'est-à-dire:

TERMINOLOGIES I DOMAINE SÉMANTIQUE <sup>10</sup>	CHIFFRE ABSOLU	%
commercial	5	10 %
automobile/automation	4	8 %
mathématiques	3	6 %
philosophie	3	6 %
juridique	2	4 %
construction	2	4 %
bancaire	2	4 %
sciences naturelles	2	4 %
géologie	1	2 %
métallurgie	1	2 %
gaz	1	2 %
maritime	1	2 %
nautique	1	2 %
médical	2	4 %
téléphonique	1	2 %
administration municipale	1	2 %
informatique	1	2 %
image et son	1	2 %
bars	1	2 %
restaurants	1	2 %
outillage de cuisine et salle à manger	1	2 %
électroménager	1	2 %
entreprise électrique	1	2 %
signalisation routière	1	2 %
sécurité sociale	1	2 %
vocabulaire général	1	2 %
assurances	1	2 %
mercerie	1	2 %
entreprise textile	1	2 %
plongée sous-marine	1	2 %
bijouterie	1	2 %
outillage	1	2 %
<b>TOTAL</b>	<b>50</b>	<b>100 %</b>

10 Il faut dire que les terminologies catalan et espagnol représentent plus de 50 % du total des terminologies éditées, en catalan, sur un total de 142.

TERMINOLOGIES 2 DOMAINES SÉMANTIQUES	CHIFFRE ABSOLU	%
physique/chimie	2	16,6 %
électricité/électronique	2	16,6 %
banque/bourse	1	8,3 %
art/construction	1	8,3 %
histoire/philosophie	1	8,3%
confection/chaussure	1	8,3 %
langue/littérature	1	8,3%
radio/télévision	1	8,3 %
dessin technique/dessin industriel	1	8,3 %
administration/relations publiques	1	8,3 %
TOTAL	12	100 %

TERMINOLOGIES 3 DOMAINES SÉMANTIQUES	CHIFFRE ABSOLU	%
hôtels/bars/restaurants	1	20 %
bois/ameublement/décoration	1	20 %
dessin industriel/construction/bois	1	20 %
administration/économie/droit	1	20 %
commercial/administration/production	1	20 %
TOTAL	5	100 %

TERMINOLOGIES 7 DOMAINES SÉMANTIQUES	CHIFFRE ABSOLU	%
métiers d'art	1	100 %
TOTAL	1	100 %

TERMINOLOGIES 10 OU PLUS DOMAINES SÉMANTIQUES	CHIFFRE ABSOLU	%
sports	1	50 %
professions	1	50 %
TOTAL	2	100 %

En relation avec les dictionnaires bilingues catalan et espagnol de type général, et du point de vue de la constance, il faut conclure que les dictionnaires terminologiques constituent un type de dictionnaire très fréquent, surtout pendant la période 1965-1990. Du point de vue de l'importance de leur répétition, ils constituent un type de dictionnaire clairement redondant.

## 5 Quelques considérations autour des résultats

De l'analyse qualitative de la production lexicographique bilingue catalan et espagnol on peut extraire quelques conclusions qui paraissent assez claires :

- 1 La corrélation entre les résultats quantitatifs et les types de dictionnaires bilingues catalan et espagnol montre qu'il y a une très grande disproportion entre la quantité d'œuvres éditées pendant la période étudiée et la diversification typologique des œuvres, surtout dans le cas des dictionnaires de type général.
- 2 L'analyse de l'apparition du premier exemplaire de chaque type montre que dans la période intermédiaire (1965-1980) la langue catalane dispose presque des mêmes types de dictionnaires bilingues catalan et espagnol qu'actuellement. Cette constatation traduit l'intérêt relatif des maisons d'édition de dictionnaires pour compléter le panorama lexicographique bilingue catalan, probablement à cause de l'incertitude du marché. On observe une tendance plus accusée à répéter des types de dictionnaires déjà existants pour lesquels le marché est assuré, qu'à diversifier l'inventaire de types, en dotant la langue catalane de la diversité d'outils lexicographiques bilingues que toute langue moderne de culture possède.
- 3 La relation catalan-espagnol constitue l'un des principaux axes de la conception et de l'édition de dictionnaires de langue catalane depuis 1965, bien qu'il semble qu'il pourrait se produire une réduction de publications présentant ces caractéristiques. Ce paramètre spécifique de la lexicographie catalane est explicite de façon spectaculaire si au nombre de dictionnaires proprement bilingues catalan et espagnol on ajoute d'autres types de dictionnaires qui incluent le catalan et l'espagnol ensemble, comme les terminologies bilingues catalan > espagnol ou espagnol > catalan. Le nombre total de ce type de dictionnaires représente pratiquement la moitié de toute la production lexicographique catalane contemporaine. Par ailleurs, la relation catalan-autres langues est très pauvre.

En conclusion, on peut donc affirmer que le panorama lexicographique bilingue catalan est très riche. Riche mais pas diversifié. Et de plus, cette richesse ne veut pas dire qu'il soit adapté aux besoins des usagers. Il faut travailler pour établir les paramètres efficaces d'évaluation lexicographique, qui aideront à situer la lexicographie bilingue catalane à la place qu'elle mérite.

# BIBLIOGRAPHIE

- AL, B. P. F. et A. SPA (Eds.) (1983) : *Le dictionnaire : Actes du Colloque Franco-Néerlandais, avril 1981*, Lille, Presses Universitaires de Lille.
- AL, B. P. F. (1991) : «L'organisation microstructurelle dans le dictionnaire bilingue» in F. J. Hausmann et al. (Eds.), Vol. 3, pp. 2828-2837.
- ALVAR EZQUERRA, M. (Ed.) (1992) : *Proceedings of Euralex '90*, Barcelone, Bibliograf SA.
- ARNAUD, P. et H. BÉJOINT (1992) : *Vocabulary and Applied Linguistics*, London, Macmillan.
- ATKINS, B. T. S. (1985) : «Monolingual and Bilingual Learners' Dictionaries : a Comparison» in R. F. Ilson (Ed.), *Dictionaries, Lexicography and Language Learning*, Oxford, Pergamon Press, pp. 15-24.
- ATKINS, B. T. S. (1990) : «Corpus Lexicography : The Bilingual Dimension» in L. Cignoni et C. Peters (Eds.), *Computational Lexicology and Lexicography*, Linguistica Computazionale, vol. VI, pp. 43-64.
- ATKINS, B. T. S. (1992) : «Tools for computer-aided corpus lexicography», in F. Kiefer et al., *Papers in Computational Lexicography*, Linguistics Institute, Hungarian Academy of Sciences, Budapest, pp. 5-68.
- ATKINS, B. T. S. et al. (1994) : «Deliverable II» of the Compass Project 62-080, Chapter 2, EC document.
- ATKINS, B. T. S. et A. ZAMPOLLI (Eds.) (1994) : *Computational Approaches to the Lexicon*, Oxford, Oxford University Press.
- BALDINGER, K. (1984) : *Vers une sémantique moderne*, Paris, Klincksieck.
- BALDINGER, K. (1985) : «Semasiologie und Onomasiologie», in L. Zgusta, *Probleme des Wörterbuchs*, Darmstadt.
- BALLY, C. (1965) : *Linguistique générale et linguistique française*, Berne, Francke.
- BÁRDOSI, V. (1992) : «Problèmes posés par le traitement lexicographique des figés», *Cahiers d'Etudes Hongroises*, N° 4, Paris, pp. 107-113.
- BÉJOINT, H. et P. THOIRON (1993) : «Macrostructure et microstructure dans un dictionnaire de collocations en langue de spécialité» in *Terminologie et traduction*, N°s 2/3, Luxembourg, Office des publications officielles des communautés européennes, pp. 513-522.
- BÉJOINT, H. (1994) : *Tradition and Innovation in Modern English Dictionaries*, Oxford, Clarendon Press.
- BLANCO, X. (1995a) : *El ejemplo en el diccionario bilingüe. Tipología y funciones del ejemplo en el marco de la lexicografía bilingüe general contemporánea francés-español, español, francés*. Thèse de doctorat, Universitat Autònoma de Barcelona.
- BLANCO, X. (1995b) : "Description et traduction d'opérateurs dans la lexicographie bilingue français-espagnol, espagnol-français", in J. Labelle (Ed.), *Lexiques-Grammaires comparés et traitements automatiques*, Montréal, Université du Québec, pp. 299-310.
- BOISSON, C., KIRTCHUK P. et H. BÉJOINT (1991) : «Aux origines de la lexicographie : les premiers dictionnaires monolingues et bilingues», *International Journal of Lexicography*, 4-4, pp. 261-315.

- BURGER, H. (1989) : «Phraseologismen im allgemeinen einsprachigen Wörterbuch», in F. J. Hausmann et al. (Eds.), vol. 1, pp. 593-599.
- BUSSE, W. et J.-P. DUBOST (1977) : *Französisches Verblexicon. Die Konstruktion der Verben in Französischen Stuttgart*, Klett.
- CABRÉ, M. T. (à paraître) : «La lexicografia catalana contemporània», *Caplletra*, València.
- CABRÉ, M. T. et M. LORENTE (1991) : *Els diccionaris catalans. De 1940 a 1988*, Barcelona, Publicacions de la Universitat de Barcelona.
- CAGNI, L. (Ed.) (1984) : *Il bilinguismo a Ebla*, Napoli, Istituto Universitario Orientale.
- Cambridge Word Routes Anglais-Français. Lexique thématique de l'anglais courant* (1994), Cambridge, Cambridge University Press.
- CANDEL, D., BERNADORFF A. et C. NIEF (1994) : «Sur quelques dictionnaires récents du français scientifique et technique», *Le Français moderne*, Paris, CILF, pp.164-180.
- CAVIGNEAUX, A. (1976) : *Die sumerisch-akkadischen Zeichenlisten, Ueberlieferungsprobleme*, Inaugural-Dissertation, München, Ludwig-Maximilians-Universität.
- CAVIGNEAUX, A. (1980-1983) : «Lexikalische Listen», *Reallexicon der Assyriologie und Vorderasiatischen Archäologie*, vol. 6, Berlin, New York, Walter de Gruyter, pp. 609-641.
- CHICHE, M. et al. (Eds.) (1989) : *Graines de parole : puissance du verbe et traditions orales : Textes offerts à Geneviève Calame-Griaule*, Paris, Editions du Centre National de la Recherche Scientifique.
- CHIEN, D. (1986) : *Lexicography in China : Bibliography of Dictionaries and Related Literature*, Exeter, University of Exeter Press.
- CHURCH, K. et P. HANKS (1990) : «Word Association Norms, Mutual Information, and Lexicography», *Computational Linguistics*, 16 :1, pp. 22-29.
- CHURCH, K. et W. GALE (1991) : «Concordances for Parallel Text», *Using Corpora. Proceedings of the 7th Annual Conference of the New Oxford English Dictionary and Text Research*, Oxford.
- CHURCH, K., W. GALE et al. (1991) : «Using Statistics in Lexical Analysis», in Y. Zernik (Ed.), *Lexical Acquisition : Exploiting On-line Resources to Build a Lexicon*, Hillsdale, New Jersey, Lawrence Erlbaum Associates.
- CIVIL, M. (1975) : «Lexicography», *Sumerological Studies in Honor of Thorkild Jacobsen on His Seventieth Birthday*, Assyriological Studies N° 20, University of Chicago Press, pp. 123-157.
- CLAS, A. et H. SAFAR (1992) : «L'environnement traductionnel : la station de travail du traducteur de l'an 2001", *Actes du Colloque de Mons 1991*, Actualité Scientifique, Montréal, Sillery (Québec), AUPELF-Presses de l'Université du Québec.
- CLAYTON, M. L. (1989) : «A Trilingual Spanish-Latin-Nahuatl Manuscript Dictionary Sometimes Attributed to Fray Bernardino de Sahagún», *International Journal of American Linguistics*, 55-4, pp. 391-416.
- COHEN, B. (1986) : *Lexique de cooccurrents; Bourse – conjoncture économique*, Montréal, Lingua tech.
- COLLIGNON, L. et M. GLATIGNY (1978) : *Les Dictionnaires : Initiation à la lexicographie*, Paris, Cedic.
- COLON, G. et A. J. SOBERANAS (1991) : *Panorama de la lexicografia catalana*, Barcelona, Biblioteca Universitària, Enciclopèdia Catalana.
- COP, M. (1990) : *Babel unravelled : An annotated world bibliography of dictionary bibliographies, 1658-1988*, Tübingen, Max Niemeyer.

- COWIE, A. P. (1981) : «The Treatment of Collocations and Idioms in Learners' Dictionaries», *Applied Linguistics*, vol. II, pp. 223-235.
- COWIE, A. P. (Ed.) (1987) : *The Dictionary and the Language Learner*, Lexicographica Series Maior, n° 17, Tübingen, Max Niemeyer Verlag.
- CSÉCSY, M. (1969) : *Problèmes de l'apprentissage du français pour des élèves de langue hongroise*, Paris, BELC.
- DAGNEAUD, R. (1965) : *Le vocabulaire grammatical*, Paris, Sedes.
- DESCAMPS, J. L. et A. PHAL (1968) : «La recherche linguistique au service de l'enseignement des langues de spécialité», in *Le Français dans le monde*, n° 61, décembre 1968, Paris, Hachette.
- DHAMOTHARAN, A. (1978) : *Tamil Dictionaries : A Bibliography*, Südasiens-Institut Universität Heidelberg, B and 50, Wiesbaden, Franz Steiner.
- DUBOIS, J. et al. (1960) : «Le mouvement général du vocabulaire français de 1949 à 1960 d'après un dictionnaire», in *Le Français moderne*, vol. 2, 3, pp. 86-106 et pp. 196-210.
- DUBOIS, J. et al. (1971) : *Introduction à la lexicographie : le dictionnaire*, Paris, Larousse, Coll. «Langue et Langage».
- DUMESTRE, G. (1989) : «Note à propos du poème de Yerowal», in M. Chiche et al. (Eds.), *Graines de parole : puissance du verbe et traditions orales. Textes offerts à Geneviève Calame-Griaule*, Paris, Editions du Centre National de la Recherche Scientifique, pp. 329-332.
- DUVAL, A. (1991) : «L'équivalence dans le dictionnaire bilingue», in F. J. Hausmann et al. (Eds.), pp. 2817-2823.
- EDMONSON, B. (1984) : «Linguistic ability in sixteenth-century Mexico», *International Journal of American Linguistics*, 50-3, pp. 343-352.
- ENGEL, U. et H. SCHUMACHER (1976) : *Kleines Valenzlexicon deutscher Verben*, Tübingen, Narr.
- FÓNAGY, I. (1982) : *Situation et signification*, Amsterdam, Philadelphia, John Benjamins.
- FONTENELLE, T. (1992) : «Collocation acquisition from a corpus or from a dictionary : a comparison» in H. Tommola et al., *Euralex '92 Proceedings*. *Studia Translatologica*, pp. 221-228.
- FRAENCKEL, B. B. (1987) : *Glossario inglês-português de termos médicos : termos médicos, termos correlatos, expressões idiomáticas*, 3500 verbetes, Rio de Janeiro : B. B. Fraenckel.
- FRIEDRICH, J. (1969) : «Churritsch», in J. Friedrich et al., *Altkeleinsiasiatische Sprachen*, Handbuch der Orientalistik 1-2-1/2-2, Leiden, Köln, Brill, pp. 1-30.
- FRIEDRICH, J. et al. (1969) : *Altkeleinsiasiatische Sprachen*, Handbuch der Orientalistik 1-2-1/2-2, Leiden, Köln, Brill.
- FU-MIEN YANG, P. (1985) : *Chinese Lexicology and Lexicography : A Selected and Classified Bibliography*, Hong-Kong, Chinese University Press.
- GALISSON, R. (1989) : «La culture partagée : une monnaie d'échange interculturelle», *Lexiques*, Paris, Hachette, pp. 113-117.
- GLÄSER, R. (1990) : *Phraseologie der englischen Sprachen*, Leipzig, Verlag Enzyklopädie.
- GOETZE, A. (1945) : «The vocabulary of the Princeton Theological Seminary», *Journal of the American Oriental Society*, 65, pp. 223-227.
- GOFFIN, R. (1985) : «La science terminologique», *Terminologie et traduction*, 2, Luxembourg, CEE, pp. 9-29.

- GOODY, J. (1979) : *La Raison graphique*, Paris, Editions de Minuit. [Traduction française de *The Domestication of the Savage Mind*, Cambridge, 1977].
- GOODY, J. (1986) : *La Logique de l'écriture : Aux origines des sociétés humaines*, Paris, Armand Colin.
- GROSS, G. (1992) : «Reconnaissance des emplois à l'aide d'un dictionnaire électronique», *Etudes de linguistique appliquée*, n° 85-86, pp. 89-99.
- GROSS, M. (1975) : *Méthodes en syntaxe*, Paris, Hermann.
- GROSS, M. (1981) : «Les bases empiriques de la notion de prédicat sémantique», *Langages*, n° 63, pp. 7-52.
- GROSS, M. (1982) : «Une classification des phrases figées du français», *Revue québécoise de linguistique*, vol. 11, n° 2, pp. 151-185.
- GROSS, M. (1987) : «La traduction automatique, bilan des descriptions en cours», *Le Français dans le monde*, pp. 30-49.
- GROSS, M. (1988) : «Les limites de la phrase figée», *Langages*, n° 90, pp. 7-23.
- GUILBERT, L. (1970) : «La dérivation syntagmatique dans les vocabulaires scientifiques et techniques» in *Actes du stage de St-Cloud*, 1967, AIDELA, pp. 116-131.
- GUILLET, A. (1986) : «Représentation des distributions dans un lexique-grammaire», *Langue française*, n° 69, pp. 85-108.
- GÜTERBOCK, H. G. et C. W. CARTER (1978) : *Keilschrifttexte aus Boghazköi*, 27 Heft, Berlin, Gebr. Mann.
- HANKS, P. (1988) : «Typicality and Meaning Potentials» in M. Snell-Hornby (Ed.), pp. 37-47.
- HARTMANN, R. R. K. (1983) : *Lexicography : Principles and Practice*, London, Academic Press.
- HARTMANN, R. R. K. (Ed.) (1984) : *LEXeter '83 Proceedings*, Tübingen, Max Niemeyer Verlag.
- HAUSMANN, F. J. (1979) : « Un dictionnaire des collocations est-il possible ? », *Travaux de Linguistique et de Littérature*, XVII, 1, pp. 187-195.
- HAUSMANN, F. J. (1988) : «Grundprobleme des zweisprachigen wörterbuchs» in K. Hylgaard-Jensen et A. Zettersten, *Symposium on Lexicography III*, Tübingen, Niemeyer, pp. 137-154.
- HAUSMANN, F. J. et al. (Eds.) (1989 - 1990 - 1991) : *Wörterbücher, Dictionaries, Dictionnaires. Encyclopédie internationale de lexicographie*, Berlin, Walter de Gruyter.
- HAUSMANN, F. J. (1989) : «Die gesellschaftlichen Aufgaben der Lexikographie in Geschichte und Gegenwart», in F. J. Hausmann et al. (Eds.), Vol. 1, pp. 1-19.
- HAUSMANN, F. J. (1989) : «Le dictionnaire de collocations», in F. J. Hausmann et al. (Eds.), vol. 1, pp. 1010-1019.
- HAUSMANN, F. J. (1991) : «Collocations in Monolingual and Bilingual English Dictionaries», in V. Ivir et D. Kalogjera, *Languages in Contact and Contrast*, Berlin, New York, Mouton de Gruyter, pp. 225-236.
- HAUSMANN, F. J. et al. (1994) : *Portugiesische und portugiesisch-deutsche Lexikographie*, Tübingen, Niemeyer.
- HEID, U. (1992) : «Décrire les collocations – deux approches lexicographiques et leur application dans un outil informatisé», *Terminologie et traduction*, 2-3.
- HEID, U. (1994) : «Relating Lexicon and Corpus : Computational Support for Corpus-Based Lexicon Building in DELIS», in W. Martin et al. (Eds.), pp. 459-471.

- HOIJER, H. (1976) : «History of American Indian Linguistics», in T. A. Sebeok (Ed.), *Native Languages of the Americas*, vol. 1, New York, London, Plenum Press, pp. 3-22.
- HOUSEHOLDER, F. W. et S. SAPORTA (1962 - 1967) : *Problems in Lexicography*, Bloomington, Indiana University.
- IANNUCCI, J. E. (1962) : «Meaning Discrimination in Bilingual Dictionaries» in F. W. Householder et S. Saporta (Eds.), pp. 201-216.
- ILSON, R. F. (Ed.) (1985) : *Dictionaries, Lexicography and Language Learning*, London, Pergamon.
- ILSON, R. F. (Ed.) (1987) : *A Spectrum of Lexicography*, Amsterdam, Philadelphia, John Benjamins.
- ISABELLE, P. (1992) : «La bi-textualité : vers une nouvelle génération d'aides à la traduction et à la terminologie», *Meta*, 37 :4, pp.721-737.
- ISABELLE, P. et al. (1993) : «Translation Analysis and Translation Automation», *Proceedings of the Fifth International Conference on Theoretical and Methodological Issues in Machine Translation*, Kyoto, Japon, pp. 201-217.
- IVIR, V. (1988) : «Collocations in Dictionaries - Monolingual and Bilingual», *Lexicographical and Linguistic Studies. Essays in Honour of G.W. Turner*, Cambridge, D.S. Brewer, pp. 43-50.
- JACOBSEN, J. R. et al. (1991) : «Examples in the Bilingual Dictionary», in F. J. Hausmann et al. (Eds.), vol. 3, pp. 2782-2789.
- JAMES, G. (1989) : «Landmarks in the development of modern Tamil lexicography», in R. R. K. Hartmann, *Lexicographers and their Works*, University of Exeter, pp. 117-148.
- JOHNSON, S. (1755) : *Preface to the Dictionary of the English Language*.
- JOHNSON, S. (1747) : *The Plan of a Dictionary of the English Language*, London.
- KASSAI, G. (1974) : «Syntagmes figés et attirances entre lexèmes», *Studia Romanica*, Fasc. III., pp. 23-37.
- KASSAI, G. (1979) : «De la réceptivité des langues», Paris, *Etudes Finno-Ougriennes*, n° XVI., pp. 37-72.
- KASSAI, G. et T. SZENDE. (Sous presse) : *Perfectionnement hongrois*, Paris, Assimil.
- KASSER, R. (1972) : «Les dictionnaires coptes», in *Textes et langages de l'Égypte pharaonique : Cent cinquante années de recherches, 1822-1972. Hommage à Jean-François Champollion*, Institut français d'archéologie orientale du Caire, Bibliothèque d'Etudes, LXIV-1, pp. 209-218.
- KATRE, S. M. (1965) : *Lexicography (A Course of Three Special Lectures Delivered at the Linguistics Department of the Annamalai University)*, Annamalai University.
- KOZLOWSKA, C. D. (1991) : *English Adverbial Collocations*, Warsaw, WN PWN.
- KRECHER, J. (1969) : «Schreiberschulung in Ugarit : die Tradition von Listen und Sumerischen Texten», *Ugarit-Forschungen*, Band I, Kevelaer, Butzon & Bercker, pp. 131-158.
- KROMAN, H. et al. (1984) : «'Active' and 'Passive' Bilingual Dictionaries : the Šcerba Concept Reconsidered», in R. R. K. Hartmann (Ed.), pp. 207-215.
- KURTBÖKE, P. (1994) : «17th-century Italian-Turkish dictionaries», W. Martin et al. (Eds.), pp. 536-541.

- LAFLEUR, B. (1966) : *Recueil de tournures et expressions idiomatiques*, Québec, Presses de l'Université Laval.
- LAINE, C. (1993) : *Combinatory Vocabulary of CAD/CAM in Mechanical Engineering*, Ottawa, Canada.
- LAMBERT, P.-Y. (1994) : *La Langue gauloise*, Paris, Errance.
- LAMBERT, W. G. (1983) : «The Earliest Bilingual Dictionary», *UR, The International Magazine of Arab Culture*, 1, pp. 44-46.
- LAMBERT, W. G. (1984) : «The Section AN», in L. Cagni (Ed.), *Il bilinguismo a Ebla*, Napoli, Istituto Universitario Orientale, pp. 393-401.
- LANDAR, H. (1977) : «Historiography of Native Ibero-American linguistics», in T. A. Sebeok (Ed.), *Native Languages of the Americas*, vol. 2, New York, London, Plenum Press, pp. 185-203.
- LANDAU, S. I. (1989) : *Dictionaries. The Art and Craft of Lexicography*, Cambridge, Cambridge University Press.
- LANDSBERGER, B. (1956) : «Babylonian Scribal Craft and its Terminology», *Proceedings of the Twenty-Third International Congress of Orientalists*, London, The Royal Asiatic Society, pp. 123-127.
- LANDSBERGER, B. et R. T. HALLOCK (1955) : «S<sup>o</sup> Vocabulary», *MSL*, III, pp. 47-87.
- LAPORTE, E. (1988) : «Reconnaissance des expressions figées lors de l'analyse automatique», *Languages*, n° 90, pp. 117-126.
- LAROCHE, E. (1966) : «Etudes de linguistique anatolienne, II», *Revue Hittite et Asiatique* n° 79, pp. 160-184.
- LAROCHE, E. (1968) : «Documents en langue hourrite provenant de Ras-Shamra», in J. Nougayrol et al. (Eds.), *Textes suméro-accadiens des archives et bibliothèques privées d'Ugarit, Ugaritica V*, Mission de Ras Shamra, Tome XVI, Paris, Imprimerie Nationale/Paul Geuthner, pp. 1-446.
- LAUFER, B. et L. MELAMED (1994) : «Reconnaissance des expressions figées lors de l'analyse automatique», *Languages*, n° 90, pp. 117-126.
- LEECH, G. (1991) : «The State of the Art in Corpus Linguistics», in K. Aijmer et B. Altenberg (Eds.), *English Corpus Linguistics*, Londres, Longman.
- LÉPINETTE, B. (1989) : «Vers un dictionnaire explicatif et combinatoire bilingue» *Cahiers de lexicologie*, n° 54-1, pp. 105-162.
- LÉPINETTE, B. (1989) : «La définition des émotions en français et en espagnol, étude sémantique et lexicographique», *Revue québécoise de linguistique*, vol. 17, n° 2, pp. 95-131.
- LÉPINETTE, B. (1992) : «L'information ajoutée codifiée dans deux dictionnaires bilingues. Contribution à l'étude de la lexicographie français-espagnol», in M. A. Ezquerro (Ed.), pp. 483-509.
- LÉPINETTE, B. (1994) : «Dictionnaire bilingue et syntaxe», in *Linguisticae Investigationes Supplementa*, Amsterdam, Philadelphia, John Benjamins (à paraître).
- LINDEMANN, M. (1994) : *Die französischen Wörterbücher von den Anfängen bis 1600. Entstehung und typologische Beschreibung*, Tübingen, Niemeyer.
- MACKLOVITCH, E. (1989) : «An Off-the-shelf Workstation for Translators», *Proceedings of the 30th Annual Conference of the ATA*, Washington, D.C., pp. 491-498.
- MACKLOVITCH, E. (1992) : «PTT-2 : la deuxième version du poste de travail du traducteur du CCRIT», *Actes des deuxièmes journées scientifiques du réseau LTT*, Mons, Belgique, pp. 171-182.

- MACKLOVITCH, E. (1993) : «PTT-3 : la troisième version du poste de travail du traducteur», *Rapport technique du CITI*, 18 p.
- MAGAY, T. et J. ZIGANI (Eds.) (1990) : *BudaLEX '88 Proceedings*, Budapest, Akadémiai Kiadó.
- MAKKAI, A. (1972) : *Idiom Structure in English*, La Haye, Mouton.
- MALBLANC, A. (1968) : *Stylistique comparée du français et de l'allemand*, Paris, Didier.
- MARELLO, C. (1989) : *Dizionario bilingui con schede sui dizionari italiani per francese, inglese, spagnolo, tedesco*, Bologna, Zanichelli.
- MARTIN, W. (1992) : «Remarks on Collocations in Sublanguage», *Terminologie et Traduction*, 2-3, pp. 157-164.
- MARTIN, W. et al. (Eds.) (1994) : *Euralex 1994 Proceedings*, Amsterdam, Euralex International Congress on Lexicography.
- MCQUOWN, N. A. (1967) : «History of Studies in Middle American Linguistics», *Handbook of Middle American Indians*, Vol. 5, Linguistics, Austin, University of Texas Press, pp. 3-7.
- MEL'ČUK, I. et al. (1983) : «Trois principes de description sémantique d'une unité lexicale dans un dictionnaire explicatif et combinatoire», *Revue canadienne de linguistique*, n° 28-2, pp. 108-21.
- MEL'ČUK, I. et L. WANNER (1994) : «Towards an Efficient Representation of Restricted Lexical Cooccurrence», in Martin et al. (Eds.), pp. 325-338.
- MEYER, I. et K. MACKINTOSH (1994) : «Phraseme Analysis and Concept Analysis : Exploring a Symbiotic Relationship in the Specialized Lexicon», in W. Martin et al. (Eds.), pp. 339-348.
- MITCHELL, L. C. (1994) : «Inversion of grammar books and dictionaries in the seventeenth and eighteenth centuries», in Martin et al. (Eds.), pp. 548-554.
- MOUNIN, G. (1963) : *Problèmes théoriques de la traduction*, Paris, Gallimard.
- MOUNIN, G. (1972) : *Clefs pour la sémantique*, Paris, Seghers.
- NORMAN, J. (1988) : *Chinese*, Cambridge, Cambridge University Press.
- NOUGAYROL, J. et al. (1968) : *Textes suméro-accadiens des archives et bibliothèques privées d'Ugarit, Ugaritica V*, Mission de Ras Shamra, Tome XVI, Paris, Imprimerie Nationale/Paul Gauthier, pp. 1-446.
- NYÉKI, L. (1988) : *Grammaire pratique du hongrois contemporain*, Paris, Ophrys-P.O.F.
- OTTEN, H. et W. VON SODEN (1968) : *Das akkadisch-hethitische Vokabular KBo I 44 + KBo XIII 1*, Wiesbaden, Otto Harrassowitz.
- PÁLFY, M. (1982) : «A polyszémia és a homonímia szótári kérdéscsöröl» (Aspects lexicographiques de la polysémie et de l'homonymie), *Nyelvtudományi közlemények*, Budapest, 84/2, pp. 404-409.
- PAVEL, S. (1993) : «Bibliographie de la phraséologie (1905-1992)», Montréal, Réseau International de Néologie et de Terminologie.
- PAVEL, S. et M. BOILEAU (1994) : *Vocabulaire des systèmes dynamiques et de l'imagerie fractale*, Montréal, Ministère des Approvisionnements et Services, Canada.
- PERGNIER, M. (1980) : *Les fondements sociolinguistiques de la traduction*, Paris, Champion.
- PERGNIER, M. (1986) : *Le Mot*, Paris, Presses Universitaires de France.
- PERROT, J. (1992) : «Soixante ans après : encore «enfiler des mots», *Cahiers d'Etudes Hongroises*, N° 4, Paris, pp. 93-97.
- PETTINATO, G. (1981) : «I vocabolari bilingui di Ebla. Problemi di traduzione e di lessicografia sumerico-ebaita», in L. Cagni (Ed.), *La lingua di Ebla*, Napoli, Istituto Universitario Orientale, pp. 241-276.

- PETTINATO, G. et al. (1982) : *Testi lessicali bilingui della biblioteca L. 2769. Parte 1 : Traslitterazione dei testi e ricostruzione del VE*, Materiali epigrafici di Ebla, 3, Napoli, Istituto Universitario Orientale di Napoli.
- PICCHIONI, S. A. (1984) : «Le varianti sumerique delle liste bilingue di Ebla», in L. Cagni (Ed.), *Il bilinguismo a Ebla*, pp. 295-308.
- PICOCHÉ, J. (1986) : *Structures sémantiques du lexique français*, Paris, Fernand Nathan.
- POPPE, N. (1965) : *Introduction to Altaic Linguistics*, Wiesbaden, Otto Harrassowitz Press, Paris, Hachette.
- QUEIXALOS, F. (1989) : «Les conceptions linguistiques des Indiens Américains», *Histoire des idées linguistiques, Tome 1 : La naissance des métalangages en Orient et en Occident*, Liège, Bruxelles, Pierre Mardaga, pp. 45-64.
- QUEMADA, B. (1967) : *Les dictionnaires du français moderne (1539-1863). Etude sur leur histoire, leurs types et leurs méthodes*, Paris, Bruxelles, Montréal, Didier.
- QUEMADA, B. (1978) : «Technique et langage», *Histoire des techniques*, Encyclopédie de la Pléiade, Paris, pp. 1146-1239.
- REY, A. (1970) : *La Lexicologie*, Paris, Klincksieck.
- REY, A. (1973) : *Théories du signe et du sens*, 2 vol., Paris, Klincksieck.
- REY, A. (1977) : *Le lexique : images et modèles. Du dictionnaire à la lexicologie*, Paris, Armand Colin.
- REY, A. (1982) : *Encyclopédies et dictionnaires*, Paris, Presses Universitaires de France.
- REY-DEBOVE, J. (1971) : *Etude linguistique et sémiotique des dictionnaires français contemporains*, La Haye, Mouton.
- REY-DEBOVE, J. (1978) : *Le Métalangage*, Paris, Le Robert.
- SATO, S. et M. NAGAO (1990) : «Toward Memory-Based Translation», *Proceedings of the 13th International Conference on Computational Linguistics, COLING-90*, Helsinki, pp. 247-252.
- SAUVAGEOT, A. (1964) : *Portrait du vocabulaire français*, Paris, Larousse.
- SCHUTZ, R. (1994) : «VLIS : Van Dale Lexicographic Information System», in W. Martin et al (Eds.), pp. 511-522.
- SCIARONE, B. (1983) : «Analyse sémantique : l'homme et la machine» *Lexique*, n° 2, pp. 133-147.
- SEBEEK, T. A. (1976-1977) : *Native Languages of the Americas*, 2 vol., New York, London, Plenum Press.
- SEYDOU, C. (1989a) : *De l'art d'accommoder les mots. Lexiques bilingues en forme de poèmes*, pp. 300-328.
- SEYDOU, C. (1989b) : «Raison poétique contre raison graphique», *L'Homme*, 110, pp. 50-68.
- SINCLAIR, J. M. (1987) : *Introduction to the Collins Cobuild English Language Dictionary*, London, Glasgow, Collins.
- SINCLAIR, J. M. (Ed.) (1987) : *Looking Up : An Account of the Cobuild Project in Lexical Computing*, London, Glasgow, Collins.
- SINCLAIR, J. M. (1991) : *Corpus Concordance Collocation*, Oxford, Oxford University Press.
- SMADJA, F. (1994) : «Retrieving Collocations from Text : Xtract», *Using Large Corpora*, Cambridge, Massachusetts, MIT, pp. 143-177.

- SMITH, S. et C. J. GADD (1925) : «A cuneiform vocabulary of Egyptian words», *Journal of Egyptian Archaeology*, 11, pp. 230-240.
- SNELL-HORNBY, M. (1984) : «The Bilingual Dictionary - Help or Hindrance», in R. R. K. Hartmann (Ed.), pp. 274-281.
- SNELL-HORNBY, M. (Ed.) (1988) : *ZuriLEX '86 Proceedings*, Tübingen, Francke Verlag.
- SNELL-HORNBY, M. et G. POEHL (Eds.) (1990) : *Translation and Lexicography*, Paintbrush series (ed. Bennani), Amsterdam, John Benjamins.
- SOLLBERGER, E. (1962) : «Graeco-Babyloniaca», *Iraq*, 24, pp. 63-72.
- SOMMERFELDT, K. E. et H. SCHREIBER (1974) : *Wörterbuch zur Valenz und Distribution deutscher Substantive*, Leipzig, Verlag Enzyklopädie.
- SOMMERFELDT, K. E. et H. SCHREIBER (1977) : *Wörterbuch zur Valenz und Distribution deutscher Adjective*, Leipzig, Verlag Enzyklopädie.
- STAAL, F. (1979) : «Oriental ideas on the origin of language», *Journal of the American Oriental Society*, 99-1, pp. 1-14.
- STEIN, G. (1990) : «The emerging role of English in the dictionaries of Renaissance Europe», *Folia Linguistica Historica*, 9-1, pp. 29-138.
- SUAREZ, J. A. (1983) : *The Mesoamerican Indian Languages*, Cambridge, Cambridge University Press.
- SVENSÉN, B. (1993) : *Practical Lexicography. Principles and Methods of Dictionary-Making*, Oxford, New York, Oxford University Press.
- SWALES, J. M. (1990) : *Genre Analysis*, Cambridge, CUP.
- SZENDE, T. (1991) : «Travaux forcés, artisanat, industrie ? Vers un nouveau dictionnaire bilingue hongrois-français/français-hongrois», *Cahiers d'Etudes Hongroises*, n° 3, 1991, Paris, pp. 55-74.
- SZENDE, T. (1992) : «Sélection des données lexicographiques : considérations méthodologiques à propos du nouveau dictionnaire hongrois-français», *Cahiers d'Etudes Hongroises*, n° 4, Paris, pp. 103-107.
- SZENDE, T. (1993) : «Traduction et lexicographie bilingue», *Cahiers d'Etudes Hongroises*, n° 5, Paris, pp. 73-91.
- SZENDE, T. et E. BORDIER (1988) : «Termes d'adresse et de salutation dans un dictionnaire hongrois-français», *Contrastes*, n° 17, Paris, pp. 92-105.
- TAYLOR, D. (1977) : *Languages of the West Indies*, Baltimore, London, Johns Hopkins University Press.
- THOIRON, P. et H. BÉJOINT (1989) : «Pour un index évolutif et cumulatif de co-occurents en langue techno-scientifique sectorielle», *Méta*, N°XXXIV, Montréal, Presses de l'Université de Montréal, pp. 661-671.
- THOIRON, P. et D. MANIEZ (1989) : «Recherche d'utilisation optimale d'un SGBD en traduction avec aides informatiques : Le dictionnaire informatisé bilingue d'écologie des eaux continentales (DIBEEC)», *Méta*, 34/3, pp. 509-515.
- THUREAU-DANGIN, F. (1931) : «Vocabulaires de Ras-Shamra», *Syria*, 12, pp. 225-266.
- TOMMOLA, H. et al. (Eds.) (1992) : *EURALEX '92 Proceedings*, Department of Translation Studies, University of Tampere, Finland.
- TOURNIER, J. (1985) : *Introduction descriptive à la lexicogénétique de l'anglais contemporain*, Paris, Genève, Champion Slatkine.

- TOURNIER, J. (1988) : *Précis de lexicologie anglaise*, Paris, Nathan.
- ULLMANN, S. (1959) : *Précis de sémantique française*, Berne, Francke.
- VERLINDE, S. et J. BINON (1994) : «The Dictionnaire Contextuel du Français Economique : A Production Oriented Dictionary of Business French», in W. Martin et al. (Eds.), pp. 523-529.
- VIVÈS, R. (1989) : «D'hier à demain. La grammaire dans tous ses états», in S. Moirand et al. (Eds.), *Le Français dans le Monde*, n° spécial.
- VOGEL, C. (1979) : *Indian Lexicography*, Wiesbaden, Otto Harrassowitz.
- VON SODEN, W. (1984) : «Zweisprachige Listen in Ebla und in der altbabylonischen Zeit», in L. Cagni (Ed.), *Il bilinguismo a Ebla*, pp. 51-64.
- Von WARTBURG, W. (1922-1928) : *Französisches etymologisches Wörterbuch. Eine darstellung des galloromanischen Sprachschatzes*, Bonn.
- WAGNER, R.-L. (1967) : *Les Vocabulaires français*, vol I : *Définitions. Les dictionnaires*, vol. II (1970) : *Les tâches de la lexicologie synchronique. Glossaires et dépouillements. Analyse lexicale*. Paris, Didier.
- WALKER, D. et al. (Eds.) (1994) : *Automating the Lexicon*, Oxford, Oxford University Press.
- WARNER, G. F. (1899) : *The Voyage of Robert Dudley [...] to the West Indies, 1594-1595, narrated by Capt. Wyatt, by himself, and by Abraham Kendall, Master*, London, The Hakluyt Society.
- WESTENHOLZ, A. (1985) : «An essay on the Sumerian 'lexical' texts of the third millenium», *Orientalia*, 54, pp. 294-298.
- WINTER, C. (1992) : «Bilingual Dictionaries : Between Language and Speech», in P. Arnaud et H. Béjoint (Eds.), pp. 41-51.
- WOLF et J. CANAVAN (1978) : *Dictionary of English Words in Context*, Dortmund, Lensing.
- WOOLDRIDGE, T. R. (1977) : *Les débuts de la lexicographie française. Estienne, Nicot et le Thresor de la langue françoise (1606)*, Toronto, University of Toronto Press.
- XUE SHIQI (1982) : «Chinese lexicography, past and present», *Dictionaries*, 4, pp. 151-169.
- ZGUSTA, L. (1971) : *Manual of Lexicography*, La Haye, Mouton.
- ZGUSTA, L. (1984) : «Translational Equivalence in the Bilingual Dictionary», in R. R. K. Hartmann (Ed.), pp. 147-154.

## DICTIONNAIRES

- BBI Combinatory Dictionary of English (A guide to Word Combinations)*, (1986), Amsterdam, Philadelphia : John Benjamins.
- Compact Wörterbuch der exacten Naturwissenschaften und der Technik*, (1991), Band II, Deutsch-Französisch, Wiesbaden : Brandstetter.
- Diccionario de la Lengua Española*, (1956), 18<sup>e</sup> édition, Madrid : Real Academia Española / Espasa Calpe S.A.
- Diccionario de la Lengua Española*, (1992), 21<sup>e</sup> édition, Madrid : Real Academia Española / Espasa Calpe S.A.
- Diccionario Manual Francés-Español Español-Francés*, (1990), Barcelona : Vox Bibliograf.
- Diccionario moderno Francés*, (1988), Barcelona : Herder.
- Dictionary of Modern Written Arabic*, (1976), New York : Spoken Language Services.

- Dictionnaire chimique allemand-français : mots et locutions fréquemment rencontrées dans les textes de langue allemande*, Paris : Dunod.
- Dictionnaire compact des sciences et de la technique*, (1995), vol. I, Français-Allemand, Wiesbaden : Brandstetter.
- Dictionnaire de didactique des langues*, (1976), Paris : Hachette.
- Dictionnaire de la langue française, Lexis*, (1987), Paris : Larousse.
- Dictionnaire de la langue française, Lexis*, (1993), Paris : Larousse.
- Dictionnaire de la linguistique*, (1974), Paris : Presses Universitaires de France.
- Dictionnaire de l'anglais d'aujourd'hui. Anglais-français, français-anglais*, (1982), Paris : Presses Pocket.
- Dictionnaire des expressions et locutions figurées*, (1979), Paris : Le Robert.
- Dictionnaire des mots et des idées*, (1956), Paris : Fernand Nathan.
- Dictionnaire du français contemporain*, (1966), Paris : Larousse.
- Dictionnaire Espagnol-Français*, (1968), Paris : Hachette.
- Dictionnaire Espagnol-Français et Français-Espagnol*, (1976/1960), Paris : Hachette.
- Dictionnaire explicatif et combinatoire du français contemporain*, (1984), Montréal : Les Presses de l'Université de Montréal.
- Dictionnaire français-anglais, anglais-français de termes et locutions techniques de génie civil, mécanique, électricité, mines, sciences et industries annexes*, Londres : Routledge & Kegan Paul.
- Dictionnaire Français-Espagnol / Espagnol-Français Hispano Bordas*, (1989), Paris : Bordas.
- Dictionnaire français-russe/frantsuzsko-russkij*, (1991), Moscou : slovar' Aktinavo Tipa.
- Dictionnaire français-latin*, (1539, 1573), Paris : R. Estienne.
- Dictionnaire moderne Français-Espagnol*, (1967), Paris : Larousse.
- Dictionnaire Oxford-Hachette français-anglais, anglais-français*, (1994), Oxford : Oxford University Press.
- Dizionario fraseologico completo italiano-spagnolo e spagnolo-italiano*, (1983, 1986), Milano : Ulrico Hoepli.
- Dizionario Garzanti di francese-italiano/italiano-francese*, (1994), Milano : Garzanti.
- Grand Dictionnaire Français-Espagnol, Espagnol-Français*, (1992), Paris : Larousse.
- Grand Dictionnaire Larousse français-anglais, anglais-français*, (1993), 2 vol., Paris : Larousse.
- Grand Robert de la langue française*, (1987), 9 vol., Paris : Le Robert.
- Handwörter buch English-Deutsch*, (1988), Leipzig : Verlag Enzyklopädie.
- Harrap's New Standard French and English Dictionary*, (1972, 1981), 4 vol., London : Harrap.
- Langenscheidts Kontextwörterbuch. Französisch-Deutsch. Ein neues Wörterbuch zum Schreiben, Lernen, Formulieren*, (1989), Berlin : Langenscheidt.
- Langenscheidts Schulwörterbuch Französisch. Französisch-deutsch/Deutsch-Französisch*, (1988), Berlin : Langenscheidt.
- Larousse moderne français-espagnol, español-francés*, (1987), Paris : Larousse.
- Les faux amis aux aguets. Dizionario di false analogie e ambigue affinità tra francese e italiano*, (1988), Bologna : Zanichelli.

- Nouveau Petit Robert*, (1993), Paris : Dictionnaires Le Robert.
- Nuevo diccionario francés-español*, (1804), Madrid : Sancha.
- Nuovo Dizionario Garzanti di francese*, (1992), Milano : Garzanti.
- Nuovo Dizionario moderno italiano-francese francese-italiano*, (1991), Torino : Società Editrice Internazionale.
- Oxford Dictionary of Current Idiomatic English*, (1975-1983), 2 vol., Oxford : Oxford University Press.
- Oxford-Hachette French Dictionary French-English English-French*, (1994), Oxford, New York, Toronto : Oxford University Press.
- Password. English Dictionary for Speakers of French*, (1989), Montréal : Modulo.
- Petit Robert 1*, (1987), Paris : Le Robert.
- Prisma Woordenboek Italiaans-Nederlands*, (1993), Utrecht : Uitgeverij Het Spectrum B.V.
- Prisma Woordenboek Nederlands-Italiaans*, (1993), Utrecht : Uitgeverij Het Spectrum B.V.
- Robert & Collins. Dictionnaire français-anglais anglais-français Senior*, (1956), 3e édition, Paris : Dictionnaires Le Robert.
- Robert & Collins. Dictionnaire français-anglais, anglais-français Senior*, (1993), Paris : Le Robert et Glasgow : HarperCollins.
- Robert & Signorelli, Dictionnaire français-italien italien-français. Dizionario francese-italiano italiano francese*, (1981), Paris : Société du Nouveau Littre et Milan : Signorelli.
- Robert & Van Dale. Dictionnaire français-néerlandais et néerlandais-français*, (1988), Paris : Dictionnaires Le Robert et Utrecht, Antwerpen : Van Dale.
- Thresor de la langue francoyse*, (1606), Paris : D. Douceur.
- Van Dale Groot woordenboek Engels-Nederlands*, (1989), Utrecht, Antwerpen : Van Dale Lexicografie bv.
- Van Dale Groot woordenboek Frans-Nederlands*, (1983), Utrecht : Van Dale.
- Van Dale Groot woordenboek Frans-Nederlands*, (1990), Utrecht, Antwerpen : Van Dale Lexicografie bv.
- Van Dale Groot woordenboek Nederlands-Frans*, (1991), Utrecht, Antwerpen : Van Dale Lexicografie bv.
- Vocabolario degli Accademici della Crusca, con tre indici delle voci, locuzioni e proverbi latini e greci posti per entro l'opera*, (1612), Venezia : Appresso Giovanni Alberti.
- Vox Manual Dictionnaire Français-Espagnol Espagnol-Français*, (1981), Barcelona : Bordas.

# INDEX

## A

abréviations grammaticales 35  
acception 8, 33, 35, 56, 60, 61, 83, 103, 114, 116, 121, 125, 126, 131, 162, 163, 214  
acronyme 154  
adressage 14  
adresse 31, 33, 34, 35, 37, 39, 40, 42, 43, 44, 46, 200, 202  
aides automatisées à la traduction humaine 169  
aire sémantique 56  
akkadien 18, 19, 20  
'al Zamakhsharî 21  
allemand 22, 31, 34, 35, 41, 42, 43, 59, 73, 97, 114, 208, 209, 211, 221  
amharique 21  
analyse contrastive 44, 143  
analyse distributionnelle 53  
analyse sémantique 41, 54, 56, 59, 140  
analyse sémémique 56, 57  
analyse syntaxico-sémantique 60, 61  
analyse syntaxique 54, 55, 60, 61, 62, 140  
analyse syntaxique contrastive 60  
analyses contrastives 65  
analyseur morphologique 136  
anglais 5, 6, 14, 18, 22, 31, 34, 35, 40, 42, 43, 45, 46, 50, 51, 59, 72, 73, 76, 77, 78, 79, 81, 82, 83, 84, 85, 86, 87, 89, 90, 91, 92, 94, 96, 97, 100, 101, 115, 129, 130, 132, 136, 138, 144, 145, 148, 154, 155, 160, 161, 162, 172, 173, 174, 175, 179, 181, 186, 188, 193, 194, 209, 221  
anglais américain 173  
anisomorphisme 44, 111  
antonyme 13, 26, 224  
approche booléenne 136  
arabe 21, 22, 23, 24, 25, 26, 28, 47, 73, 75, 76, 78, 97, 100  
araméen 20  
arawak 24  
archaïsme 37, 57  
argot 130, 224  
arménien 22, 23

article 9, 10, 13, 53, 54, 56, 58, 64, 66, 67, 82, 94, 103, 109, 110, 111, 118, 119, 121, 123, 127  
article lexicographique  
article 5, 7, 12, 33, 120, 124, 129, 132, 134, 137, 140, 144, 146, 148, 152, 167, 171, 172, 173, 175, 177, 178, 179, 187, 200, 202, 213, 215, 216  
article abrégé 38  
assonance 163

## B

bambara 26  
banalisation 204, 205  
banque de données 203  
banque de terminologie 171  
base de données 11, 152, 155, 156, 158, 177, 179  
base de données de traductions mémorisées 178  
base de données lexicales 158  
base de données terminologiques 11, 169  
base de données textuelles 161  
base documentaire 203  
base d'une collocation 164, 165  
base lexicographique 60  
*BBJ Combinatory Dictionary of English* 59, 194  
bidirectionnalité 34, 35, 44  
bilinguisation 50  
bohairique 21  
boîte de dialogue 173  
British National Corpus 129, 130, 166  
bureau sans papier 172

## C

calque 203  
*Cambridge Word Routes* 40, 232  
catalan 8, 14, 216, 217, 218, 219, 220, 221, 222, 223, 224, 227, 228, 230

- catégorie grammaticale 42, 54, 71, 74, 75, 78, 81, 83, 101, 178  
 catégorie syntaxique 171  
*Catholicon* 55  
 CD-ROM 152, 156, 157, 168, 172, 173, 174, 175, 176  
 chaghataï 23  
 champ analogique 103  
 champ sémantique 30, 35, 143  
 chinois 19, 23, 27, 28, 51, 73  
 citation 33, 126, 129, 130, 131, 140, 142, 143, 146, 154, 159, 160, 166, 168, 172, 173, 174, 176, 202, 211  
 classe de noms 63, 64  
 classe morphologique 56  
 classement 10, 13, 27, 28, 41, 58, 65, 123, 183  
     classement acrographique 27  
     classement alphabétique 27  
     classement macrostructurel à nids 41  
 COBUILD 14, 128, 167  
 code numérique 94, 96  
*Collins Cobuild Dictionary of the English Language* 71  
*Collins English Dictionary* 143  
 collocabilité 12, 163  
 collocatif 12, 13, 164, 165  
     collocatif typique 166  
 collocation 9, 10, 12, 13, 39, 40, 42, 59, 81, 85, 94, 95, 96, 97, 101, 129, 134, 135, 136, 137, 138, 143, 147, 148, 163, 164, 165, 166, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197  
 combinaison 26, 47, 53, 63, 74, 77, 80, 81, 83, 88, 93, 94, 95, 98, 99, 101, 123, 124, 182, 183, 189, 193, 221  
 commentaire métalinguistique 195  
 communauté linguistique 113, 114, 133, 203  
 communication interlinguistique 6  
 communication techno-scientifique 6  
 commutation 114  
 comparatisme linguistique 28  
 COMPASS 15, 35  
 complément d'objet 144  
 comportement (grammatical, syntaxique) 8, 12, 60, 71, 82, 101, 131, 135, 143, 145, 146, 209  
 composé 20, 23, 34, 41, 42, 43, 59, 105, 107, 117, 128, 188  
 composition 27, 29, 42, 117, 118, 128, 170, 211  
 concaténation orthographique 155  
 concept 10, 11, 15, 92, 112, 116, 122, 126, 164, 167, 169, 170, 177  
 concordance 119, 137, 140, 141, 162  
     concordance KWIC 135  
 concordancier bilingue 177, 178, 179  
 conjugaison 74, 77, 78, 80, 81, 88, 89, 90, 92, 93, 94, 97, 98, 101, 174  
 connotation 85, 120, 143, 209, 210, 215  
     connotation collective 120  
     connotation individuelle 120  
 constituant 14, 56, 57, 58, 60, 155, 196  
 consultation 14, 28, 138, 146, 148, 152, 169, 171, 190, 196  
 contenu sémantique 58, 60, 61, 125  
 contexte 7, 15, 35, 39, 42, 43, 44, 47, 48, 49, 118, 119, 120, 123, 124, 125, 126, 127, 128, 132, 134, 136, 137, 138, 139, 140, 142, 143, 144, 145, 146, 147, 148, 155, 160, 162, 165, 166, 167, 177, 178, 185, 193, 194, 195, 210, 213, 221, 224  
 contrôle croisé 33  
 cooccurrence 59, 103, 237  
 cooccurrent 12, 58, 165  
 copte 21, 28, 29, 235  
 corpus 6, 10, 14, 39, 104, 106, 107, 108, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 142, 143, 144, 145, 146, 147, 148, 149, 153, 154, 155, 156, 158, 159, 160, 161, 162, 163, 166, 167, 168, 176, 178, 179, 182, 184, 185, 186, 216  
 couper-coller 174  
 cunéiforme 18, 19, 20

**D**

- danois 24  
 décodage 7, 10, 36, 64, 100, 133, 135, 147, 157, 191, 199  
 décoder 35, 36, 38, 133, 146, 157, 181, 215

- décomposition 58  
 découpage sémantique 33, 44  
 définition 8, 9, 10, 11, 12, 17, 27, 32, 40, 44, 49, 50, 55, 56, 57, 58, 104, 123, 135, 169, 171, 173, 175, 182, 183, 194, 195, 200, 201, 202, 203, 206, 214, 215, 218  
 délimitation des sens 57  
 dénomination 32, 111, 201, 204, 213  
 dénotation 210  
 dépouillement manuel 39  
 dérivation 117  
 dérivé 23, 41, 74, 77, 81, 83, 143, 215  
 description sémantique 54  
 diachronie 28  
   diachronique 8, 55, 173  
 dialecte 18, 21, 27, 28, 29, 34  
 diastratique 55  
 diathèse 103, 112  
*Diccionario de la Real Academia Española* 106  
*Diccionario Manual Francés-Español* 104  
*Diccionario Manual Francés-Español Español-Francés* 104  
*Diccionario moderno Francés* 104  
*Dictionary of Modern Written Arabic* 97, 100, 102  
 dictionnaire (bilingue) de version 33, 34, 35, 36, 37, 42, 44  
 dictionnaire abrégé 38, 43  
 dictionnaire bi-directionnel 175  
 dictionnaire bilinguisé 51  
 dictionnaire bilingue abrégé 37  
 dictionnaire bilingue de thème 33, 35, 42, 44, 48  
 dictionnaire bilingue d'encodage 64, 159  
 dictionnaire de production 94, 101  
 dictionnaire d'encodage 159, 167, 199  
 dictionnaire bilingue en-ligne 169  
 dictionnaire bilingue inter-dialectes 27  
 dictionnaire bilingue sur CD 35  
 dictionnaire combinatoire 59  
 dictionnaire conceptuel 18  
 dictionnaire d'antonymes 224  
 dictionnaire d'argot 224  
 dictionnaire de collocations 165, 184  
 dictionnaire de décodage 100  
 dictionnaire de difficultés 199  
 dictionnaire de faux-amis 39  
*Dictionnaire de l'anglais d'aujourd'hui* 46  
 dictionnaire de langue de spécialité 5, 7, 10, 12  
 dictionnaire de locutions et phraséologismes 223  
 dictionnaire de poche 92  
 dictionnaire de prononciation 19, 27  
 dictionnaire de rimes 223  
 dictionnaire de synonymes 199  
 dictionnaire de valences 59  
 dictionnaire de vocabulaire fondamental 224  
 dictionnaire d'expressions 194  
 dictionnaire d'homonymes 224  
 dictionnaire d'homophones 27  
 dictionnaire diachronique 173  
 dictionnaire dialectal 223  
 dictionnaire d'orthographe (orthographique) 199, 224  
*Dictionnaire du français contemporain* 57, 61  
 dictionnaire électronique 171, 172, 173  
 dictionnaire en hypertexte 157  
 dictionnaire en-ligne 169, 170, 171, 172  
 dictionnaire encyclopédique 198, 223  
*Dictionnaire Espagnol-Français et Français-Espagnol* 104  
 dictionnaire étymologique 30  
 dictionnaire évolutif 14  
*Dictionnaire Français-Espagnol / Espagnol-Français* 104  
*Dictionnaire Français-Russe* 167  
*Dictionnaire français-latin* 31, 32, 50  
 dictionnaire général encyclopédique 37  
 dictionnaire homoglosse 34  
 dictionnaire idéologique 223  
 dictionnaire informatisé 163, 165, 166, 168, 169  
 dictionnaire interlinguistique 200, 202, 203  
 dictionnaire monodirectionnel 46  
 dictionnaire monolingue 5, 9, 12, 17, 27, 32, 37, 44, 45, 50, 55, 58, 71, 79, 101, 129, 134, 143, 172, 174, 199, 215  
 dictionnaire monolingue bilinguisé 32  
 dictionnaire monolingue général 5, 37, 134  
 dictionnaire monolingue pour apprenants 50, 71

dictionnaire onomasiologique 40  
 dictionnaire oral 26  
 dictionnaire pédagogique 52  
 dictionnaire phraséologique spécialisé bi-  
 lingue 39  
 dictionnaire plurilingue 28  
     dictionnaire plurilingue comparatif 28  
 dictionnaire polytechnique 207  
 dictionnaire scolaire 37, 223  
 dictionnaire sémasiologique 40  
 dictionnaire semi-bilingue 32  
 dictionnaire spécialisé 39  
 dictionnaire synonymique 40  
 dictionnaire terminologique 5, 216, 222,  
     223, 224, 226, 230  
 dictionnaire unilingue 82, 157, 160, 171  
 dictionnaire visuel 224  
 dictionnaires généraux unilingues 184, 185  
 diglossie 29  
 domaine thématique 56  
 données terminologiques 11, 169

**E**

éblaïte 19  
 échange dynamique de données 156  
 écriture 7, 18, 24, 26, 29  
 Egede, Paul 24  
 égyptien 20  
 élément «pré-codé» 122  
 emprunt 203, 210  
 en-ligne 12, 158, 169, 170, 171, 172, 175  
 encodage 7, 36, 59, 61, 64, 100, 133, 135,  
     146, 147, 148, 157, 159, 167, 181, 191,  
     193, 199  
 encyclopédie 8, 18, 213, 215, 234, 238  
 encyclopédique 11, 18, 37, 123, 134, 158,  
     161, 166, 199, 210, 211, 223, 225, 226  
 enrichissement lexical 42  
 entrée 12, 55, 58, 62, 66, 79, 83, 140, 163,  
     189, 191, 195, 202  
     entrées homonymiques 39  
 environnement lexical 60  
 équivalence 8, 9, 10, 11, 49, 60, 85, 107,  
     111, 119, 125, 126, 135, 139, 140, 142,  
     144, 147, 148, 160, 161, 162, 167, 174,  
     208, 209, 210, 211, 214

équivalent 8, 9, 11, 12, 18, 19, 23, 28,  
     32, 33, 34, 37, 39, 42, 43, 44, 45, 46,  
     47, 49, 53, 54, 55, 56, 58, 60, 61, 62,  
     63, 65, 79, 82, 83, 87, 89, 90, 91, 112,  
     113, 114, 115, 118, 119, 120, 121, 122,  
     123, 125, 126, 130, 134, 140, 142, 144,  
     145, 146, 147, 148, 154, 159, 160, 162,  
     166, 167, 169, 171, 174, 175, 177, 178,  
     182, 189, 192, 193, 194, 195, 197, 200,  
     202, 203, 210, 211, 214, 215, 217, 218  
 espagnol 10, 24, 28, 31, 43, 54, 55, 61, 65,  
     73, 103, 104, 105, 106, 108, 109, 216,  
     217, 218, 219, 220, 221, 222, 223, 224,  
     227, 228

Estienne, Robert 31, 32, 50  
 étiquetage SGML 127  
 étiquette syntaxique 136  
 étymologie 171  
 exemple 10, 19, 36, 37, 39, 40, 42, 45, 46,  
     49, 50, 73, 75, 78, 79, 83, 84, 85, 88,  
     92, 94, 95, 97, 99, 100, 103, 104, 105,  
     106, 107, 108, 109, 110, 111, 124, 125,  
     126, 129, 134, 139, 141, 143, 144, 145,  
     148, 158, 159, 162, 163, 166, 169, 172,  
     177, 178, 179, 185, 187, 188, 190, 193,  
     194, 207, 208, 210, 214, 215, 216  
 expression figée 66, 85, 120, 185, 186, 187,  
     189, 236  
 expression figurée 176, 179  
 expression idiomatique 9, 10, 43, 162, 181,  
     182, 183, 184, 185, 186, 187, 188, 189,  
     190, 192, 193, 194, 195, 197  
 expression vedette 142

**F**

faux ami 116, 143, 146  
 fiche 175  
 flexion 42, 78, 80, 88, 93, 98, 171  
 fonction lexicale 183  
 fonctionnement syntaxique 60, 65  
 forme dérivée 134  
 forme fléchie 71, 136, 140, 143  
 forme lemmatisée 190, 191, 192, 194  
 fragments répétés 14

français 10, 19, 20, 31, 32, 33, 34, 35, 36, 37, 40, 42, 43, 44, 45, 46, 47, 49, 50, 54, 55, 57, 59, 61, 63, 65, 73, 76, 78, 87, 88, 89, 90, 91, 92, 94, 96, 103, 104, 105, 106, 108, 109, 111, 112, 113, 114, 115, 116, 117, 118, 122, 123, 124, 125, 129, 132, 136, 139, 143, 145, 148, 160, 161, 162, 166, 167, 172, 173, 174, 175, 182, 185, 186, 188, 194, 202, 206, 208, 209, 211, 218

francophonie 6, 15

*Französisches etymologisches Wörterbuch* 31

fréquence 12, 44, 75, 81, 86, 91, 96, 99, 100, 124, 131, 136, 137, 140, 143, 147, 155, 161, 166, 170, 176, 186, 210, 214, 223, 225

---

## G

---

gallicisme 107

gaulois 22

genre 5, 27, 40, 71, 72, 74, 75, 77, 78, 79, 80, 81, 82, 83, 85, 88, 89, 91, 93, 94, 97, 98, 101, 115, 159, 179, 185, 215

géorgien 23

glose 22, 29, 119, 202

glossaire 22, 34, 171, 175, 214

grammatical 10, 71, 75, 79, 82, 83, 84, 85, 87, 89, 92, 93, 115, 123, 124, 168, 183

*Grand Dictionnaire Larousse* 144

graphie 37, 42

grec 18, 19, 20, 21, 22, 28, 51, 221

grec ancien 19

groënlandais 24

groupements lexicaux 161

groupes sémantiques 139, 148

guèze 19, 21

---

## H

---

*Harraps Shorter* 144

hébreu 21

hittite 20

homographe 154

homomorphique 56

homonyme 172

hongrois 111, 112, 113, 114, 115, 116, 117, 118, 121, 122, 123, 124, 125, 221

hourrite 20

hyperonyme 202, 215

---

## I

---

identificateur de sens 215

identité conceptuelle 10

idiomaticité 36, 192

idiome 19, 181

idiotisme 9, 134

illustration 40, 107, 158, 159, 200

index 9, 50, 137, 172

indexation 154

indicateur 35, 36, 42, 44, 46, 49, 77, 80, 83, 88, 92, 93, 98, 101, 119, 134, 137, 162, 222, 226, 227

indicateur de champ sémantique 46

indicateur de collocation 191

indicateur de contexte 131, 137

indicateur sémantique 50

inflexion 75, 78, 81, 89, 219

informateur 184, 185

informatique 10, 11, 128, 136, 159, 163, 166, 228

informatisation 151, 152, 156, 157, 159, 163, 168

interface graphique 156, 177

intransitif 47, 55, 74, 82

introspection 131

intuition 71, 82, 83, 123, 134, 142, 148, 149, 164, 178

intuition grammaticale 82

isomorphisme 203

italien 2, 24, 44, 46, 50, 73, 76, 78, 92, 94, 96, 221

---

## J

---

japonais 23, 27, 36, 50, 221

jargon 130, 134

**K**

kannada 21  
kawi 21  
kouman 23, 24

**L**

L1 7, 33, 36, 107, 108, 109, 132, 159  
L2 33, 107, 108, 109, 132, 159  
langage définitionnel univoque 58  
*Langenscheidts Schulwörterbuch Deutsch-Französisch* 41  
langue altaïque 27  
langue cible 10, 11, 12, 15, 31, 33, 53, 54, 55, 61, 62, 132, 134, 135, 137, 138, 139, 140, 142, 144, 145, 146, 147, 148, 157, 158, 159, 160, 162, 164, 166, 167, 171, 174, 177, 178, 199, 200, 202, 203, 209, 210, 211, 214, 216, 220  
langue d'arrivée 15, 119, 192  
langue de départ 15, 113, 118, 119, 125, 126, 169, 191, 192, 193  
langue de spécialité (langue spécialisée) 10, 12, 13, 204, 206  
langue de travail 134  
langue écrite 130  
langue internationale 20, 34  
langue maternelle 21, 25, 32, 33, 35, 39, 44, 50, 71, 83, 87, 134, 140, 157, 199  
langue nationale 34  
langue professionnelle 205  
langue seconde 181  
langue source 10, 11, 15, 31, 33, 34, 53, 54, 55, 61, 132, 134, 135, 137, 139, 140, 142, 144, 146, 147, 148, 157, 158, 159, 166, 172, 174, 202, 203, 209, 210, 211, 214, 216, 220, 221  
langue technique 205  
langue vulgaire 55  
langue vulgarisée 205  
langues romanes 43, 60, 62  
latin 8, 19, 21, 22, 23, 24, 27, 28, 31, 32, 34, 50, 54, 55, 216, 217, 221  
lemmatisation 154, 190  
lemme 105, 108, 109, 154, 167

lexème 56, 58, 60, 138, 140, 160, 183, 203  
lexicographie unilingue (monolingue) 43, 53, 54, 57, 160, 168  
lexie 211  
lexique central 37  
lexique général 13  
Lexique-grammaire 58, 59, 60, 62, 65  
lien dérivationnel 41  
ligne de concordance 137  
ligne de contexte 185  
linguistique contrastive 65  
linguistique informatique 136  
locution 124, 184, 200  
logiciel 14, 127, 128, 153, 155, 156, 158, 166, 169, 170, 171, 172, 175, 185  
logogramme 19, 20  
*Longman Dictionary of Contemporary English* 71

**M**

macrostructure 7, 8, 37, 39, 49, 55, 59, 127, 134, 152, 200, 201  
macrostructure à articles alphabétisés 42  
macrostructure à nids 41, 42  
macrostructure à nids ou à niches 42  
macrostructure alphabétisée 41, 42  
Mahmûd al-Kâsyarî 22  
majuscules 19, 136, 153, 154  
mandchou 23  
marathe 22  
marque 115, 118  
marqueur 82, 83  
mémoire de traduction 176, 177, 178, 179  
métalangage 146  
métalangue 200, 202, 206  
métalexicographie 5, 7, 17, 29  
métalexicographe 7, 128, 132, 142  
*Micro Robert* 49  
*Micro Signorelli* 49  
micro-contexte 160  
microstructure 7, 39, 42, 44, 47, 48, 49, 103, 104, 110, 127, 152, 181, 187, 200  
minuscules 136, 153, 154  
mixtec 24  
mongol 23

monodirectionnalité 35, 47  
 monolinguisme 6  
 monosémique 56  
 morphème 181  
 morphologie 34, 42, 117, 206, 211  
 mot clef 136, 213  
 mot-vedette 42, 103, 104, 109, 152, 155,  
 163, 167, 187, 192, 193, 194, 195  
 motivation 116, 117, 118  
 moyen mongol occidental 23

---

## N

---

nahuatl 24, 25, 28  
 narraganset 24  
 Nebrija (de), Antonio 24, 54, 217  
 néerlandais 71, 73, 76, 77, 78, 79, 81, 82,  
 83, 84, 85, 86, 87, 88, 89, 90, 91, 92,  
 96, 101  
 néologisme 17, 203  
*New English Dictionary* 130  
 niche 41  
   niche graphique 41  
   niche sémantique 41  
 Nicot, Jean 32, 55  
 nid 40, 41  
 niveau de langue 42, 45  
 nom commun 136  
 nom composé 89  
 nom propre 136, 154  
 nomenclature 34, 137, 158, 159, 200, 206,  
 207, 214  
 normalisation 214  
 notion 63, 64, 126, 164, 166, 206, 213

---

## O

---

occurrence 135, 136, 137, 139, 145, 154  
 onomasiologique 40, 58, 61, 214  
 onomastica 20  
 onomatopée 112  
 ordre alphabétique 37, 41, 79, 136, 214, 217  
 orthographe 172  
 ougaritique 20  
*Oxford English Dictionary* 130, 143, 172,  
 173

*Oxford Pilot Corpus* 129, 139  
*Oxford Student's Dictionary* 51  
*Oxford-Hachette French Dictionary*  
 (OHFD) 36, 43, 127, 129, 148, 176

---

## P

---

paradigmatique 54, 202, 209  
 paraphrase 55  
 parties du discours 81, 136, 164  
*Password* 51  
 patois 34  
 persan 21, 22, 23, 28  
*Petit Robert* 49, 50, 143, 184, 203  
 phonème 29  
 phraséologie 9, 13, 14, 42, 43, 167  
   phraséologique 13, 39, 42, 43  
   phraséologisme 9, 223  
 plurilinguisme 6  
 politique éditoriale 133  
 politique lexicographique 193  
 polysémie 167  
   polysémique 38, 42, 44, 54, 56, 57  
 ponctuation 75, 154  
 portugais 22, 23, 51, 73, 221  
 post-lexicographique 153, 155, 156  
 poste de travail du traducteur (PTT) 169,  
 170, 171, 172, 173, 174, 175, 176, 177  
 pragmatique 11, 55, 140, 160  
 pré-lexicographique 153, 159  
 préfixe 136  
*Prisma Woordenboek* 76  
 Proença (de), António 22  
 progiciel 170  
 pronominal 47, 74  
 prononciation 19, 20, 27, 37, 42, 46, 75,  
 101, 171, 173, 178, 206, 224  
 proverbe 43, 75, 101  
 psycholinguiste 131

---

## Q

---

quasi-synonyme 143, 145

**R**

Rabbi Sa'adia ben Joseph 21  
 réalité extralinguistique 111  
 référent 12  
 reformulation syntaxique 166  
 registre 23, 148, 166, 195  
 renvoi 12  
 restriction 124  
*Robert & Collins* 45, 192  
*Robert & Collins Français - Anglais An-*  
*glais -Français* 45  
*Robert & Signorelli* 49  
*Robert & Van Dale* 48, 49  
 Rottweil (von), Adam 22  
 russe 31, 73, 167, 221

**S**

saïdique 21  
 sanskrit 21, 22, 23, 29  
 schéma syntaxique 61  
 sémantique 8, 9, 12, 13, 25, 26, 33, 41, 42,  
 44, 46, 54, 55, 56, 58, 59, 60, 61, 62,  
 64, 65, 66, 112, 115, 116, 124, 125,  
 134, 139, 140, 142, 145, 178, 182, 184,  
 186, 188, 189, 193, 195, 209, 227  
 sémantisme 54, 55, 125  
 sens figuré 116, 183  
 sens propre 116  
 séquences répétées 105, 106  
*Shorter Oxford English Dictionary* 143  
 signe arbitraire 118  
 signe motivé 117, 118  
 signifiant 119  
 signification 44, 50, 55, 75, 94, 95, 112,  
 118, 123, 124, 137, 160, 168, 214  
 signifié 31, 32, 39, 40, 44, 47, 48, 60, 64,  
 119, 120, 123, 215, 217  
 sociolecte 161  
 sociolexicographique 14  
 sous-entrée 61, 62, 187, 188, 191  
 sous-unité lexicale 61  
 sous-vedette 32  
 statistique 102, 136, 137, 158, 163, 178  
 structure lexico-sémantique 145

structure syntaxique 101, 125, 139, 140  
 style 75, 119, 120, 129  
 suffixe 117, 136  
 sumérien 18, 19, 20, 28  
 syllabaire 20  
 syllabogramme 18, 19  
 symbole d'accentuation 75  
 synchronie 9, 28  
 synonyme 9, 12, 13, 15, 39, 56, 75, 121,  
 122, 144, 145, 162, 173, 186, 193, 194,  
 197, 199, 201, 202, 207, 210, 216, 224  
 synonymie 12, 13, 120, 121, 197  
 syntagmatique 8, 9, 14, 43, 44, 47, 54, 124,  
 160, 166, 202, 209  
 syntagme 32, 39, 42, 43, 44, 54, 62, 64, 124,  
 125, 135, 138, 177  
     syntagme figé 64  
     syntagme fixe 135  
     syntagme lexicalisé 42, 43  
 syntaxe 53, 54, 59, 65, 129, 134, 135, 140,  
 204, 206, 208  
 syntaxique 8, 53, 54, 55, 57, 58, 59, 60, 61,  
 62, 63, 65, 66, 101, 103, 115, 125, 131,  
 135, 136, 139, 140, 141, 142, 143, 147,  
 148, 155, 166, 171, 182, 194, 205, 208,  
 209

**T**

tablette 19  
 taïno 24  
 tarasque 24  
 taxonomies 159  
 technicisme 37  
 technolecte 161  
 télougou 22  
 terme 6, 11, 12, 13, 17, 18, 19, 21, 22, 23,  
 25, 26, 28, 29, 37, 39, 41, 55, 56, 58,  
 59, 85, 104, 112, 113, 114, 115, 116,  
 117, 120, 121, 122, 123, 124, 126, 129,  
 130, 133, 135, 143, 147, 153, 157, 162,  
 163, 175, 204, 207, 208, 210, 211, 213,  
 214  
 terminographe 12  
 terminographie 14  
 terminographie grise 14

terminologie 6, 12, 15, 55, 56, 130, 159, 164, 169, 171, 176, 208, 209, 214, 218, 224, 227, 228, 230  
 terminologisme 207  
 TERMIUM 9, 170  
 théorie sens-texte 58  
 tibétain 21, 23  
 toscan 50  
 tradition lexicographique 216, 217  
 traducteur 14, 87, 120, 146, 147, 165, 169, 171, 172, 173, 174, 175, 176, 177, 179  
 traduction 6, 14, 18, 19, 20, 22, 25, 26, 29, 31, 33, 34, 42, 44, 45, 49, 55, 61, 62, 64, 65, 66, 72, 73, 76, 83, 89, 97, 103, 104, 106, 114, 119, 122, 123, 124, 129, 130, 134, 135, 137, 139, 140, 142, 144, 145, 146, 147, 148, 159, 160, 161, 162, 165, 166, 167, 169, 170, 171, 172, 174, 175, 176, 177, 178, 179, 192, 194, 202, 210, 215  
 traitement automatique du langage naturel 136  
 transitif 47, 74, 82  
 turc 23, 24  
 typologie des dictionnaires 8  
 typologie des dictionnaires bilingues 37, 42  
 zeltal 24

---

## U

---

unité de traduction 160  
 unité de traduction potentielle 160  
 utilisateur 61, 132, 178

---

## V

---

valence 168  
*Van Dale* 44, 47, 48, 49, 72, 76, 78, 79, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91  
*Van Dale Frans-Nederlands* 47  
 variante 6, 41  
 variations syntaxiques 194  
 verbe à particule 194  
 vieux javanais 21  
 vieux slave 19  
 vocable 56, 57, 58, 59, 60, 62, 64  
*Vocabolario degli Accademici della Crusca* 50  
 vocabulaire 25, 41, 113, 117, 126, 130, 134, 158, 173, 176, 181, 204, 224, 228  
 vocabulaire scientifique 206  
 vocabulaire spécialisé 135  
 vocabulaire spécialisé et semi-spécialisé 135  
 vocabulaire technique 171, 206  
*Vocabulario de romance en latín* 54  
 von Wartburg 31

---

## W

---

Williams, Roger 24

---

## Z

---

zapotèque 24



# LISTE DES CONTRIBUTEURS

Henri BÉJOINT

Centre de Recherche en Terminologie et Traduction, Université Lumière-  
Lyon 2  
Lyon, France

Xavier BLANCO

Départament de Filologia Francesa i Romànica, Universitat Autònoma de  
Barcelona  
Barcelone, Espagne

Claude BOISSON

Département d'Anglais, Université Lumière-Lyon 2  
Lyon, France

Maria Teresa CABRÉ

Institut de Linguística Aplicada, Universitat Pompeu Fabra  
Barcelone, Espagne

André CLAS

Département de Linguistique et Philologie, Université de Montréal  
Montréal, Canada

Cristina GELPÍ

Institut de Linguística Aplicada, Universitat Pompeu Fabra  
Barcelone, Espagne

Valerie GRUNDY

Oxford University Press  
Oxford, Grande-Bretagne

Frank KNOWLES

Institute for the Study of Language and Society, Aston University  
Birmingham, Grande-Bretagne

Marcel LEMMENS

Hogeschool Maastricht, Ecole de traduction et d'interprétation  
Maastricht, Pays-Bas

Brigitte LÉPINETTE

Departament de Filologia Francesa i Italiana, Universitat de València  
Valence, Espagne

Elliott MACKLOVITCH

Centre d'Innovation en Technologies de l'Information  
Laval, Canada

Carla MARELLO

Dipartimento di Scienze letterarie e filologiche, Università di Torino  
Turin, Italie

Roda ROBERTS

School of Translation and Interpreters, University of Ottawa  
Ottawa, Canada

Thomas SZENDE

Centre interuniversitaire d'études hongroises, Université de la Sorbonne  
Nouvelle  
Paris, France

Philippe THOIRON

Centre de Recherche en Terminologie et Traduction, Université Lumière-  
Lyon 2  
Lyon, France

# TABLE DES MATIÈRES

Introduction	5
Chapitre 1	
Claude BOISSON	
L'antiquité et la variété des dictionnaires bilingues	17
Chapitre 2	
Carla MARELLO	
Les différents types de dictionnaires bilingues	31
Chapitre 3	
Brigitte LÉPINETTE	
Le rôle de la syntaxe dans la lexicographie bilingue	53
Chapitre 4	
Marcel LEMMENS	
La grammaire dans les dictionnaires bilingues	71
Chapitre 5	
Xavier BLANCO	
Elaboration et réutilisation des exemples dans la lexicographie bilingue	103
Chapitre 6	
Thomas SZENDE	
Problèmes d'équivalence dans les dictionnaires bilingues	111
Chapitre 7	
Valerie GRUNDY	
L'utilisation d'un corpus dans la fabrication des dictionnaires bilingues	127
Chapitre 8	
Frank KNOWLES	
L'informatisation de la fabrication des dictionnaires bilingues	151

Chapitre 9

Elliott MACKLOVITCH

Les dictionnaires bilingues en-ligne et le poste  
de travail du traducteur

169

Chapitre 10

Roda P. ROBERTS

Le traitement des collocations et des expressions idiomatiques  
dans les dictionnaires bilingues

181

Chapitre 11

André CLAS

Problèmes de préparation rédactionnelle  
de dictionnaires bilingues spécialisés : quelques réflexions

199

Chapitre 12

M.Teresa CABRÉ et Cristina GELPÍ

La lexicographie bilingue catalane contemporaine :  
analyse et évaluation

213

Bibliographie

231

Index

243

Liste des contributeurs

253



À une époque où s'intensifient les échanges entre communautés linguistiques, les travaux de recherche sur les dictionnaires bilingues et multilingues généraux et spécialisés revêtent une importance toute particulière. *Les Dictionnaires bilingues*, premier ouvrage en français consacré à ce sujet, présente les réflexions inédites d'une douzaine d'universitaires, européens et nord-américains. Il s'adresse aux linguistes, étudiants et enseignants, aux chercheurs, et à tous ceux qui s'intéressent aux dictionnaires.

---

Henri Béjoint

*Professeur d'anglais au Département de Langues Etrangères Appliquées de l'Université Lumière Lyon 2, où il enseigne la traduction, la lexicologie et la terminologie, il est également Vice-Président d'EURALEX. Il a publié de nombreux articles dans le domaine de la lexicographie et quelques livres, parmi lesquels Tradition and Innovation in Modern English Dictionaries (Oxford University Press, 1994)*

Philippe Thoiron

*Professeur d'anglais au Département de Langues Etrangères Appliquées de l'Université Lumière Lyon 2, où il enseigne la linguistique, la linguistique quantitative et la terminologie, il est Directeur du Centre de Recherche en Terminologie et traduction (C.R.T.T.). Il a publié de nombreux articles dans le domaine de la linguistique quantitative et de la terminologie multilingue et membre du comité scientifique de réseau Lexicologie, Terminologie et Traduction (L.T.T.) de l'AUEPLF-UREF.*

150 FRF - 820 BEF - 40 CDN - Prix préférentiel UREF 60 FRE

ISBN 2-8011-1138-4



9 782801 111383

DICBIL  
C235

Diffusion Edicef ou Ellipses selon pays

59.4758.5