

# Web et terminologie philologique

**Pierre Lerat**

LLI-Université Paris XIII

*A linguistic method for using the Web consists in making requests on strings of characters as long as possible between quotation marks, then confronting the obtained results with other sources. One knows that the Web is “transitory” and is “varying greatly in value”; nonetheless using this method will be efficient. Useful requests are made through classes of objects and sets of predicates (Gross & Guenther 2002). Comparing frequencies is a good means of evaluating the respective vitality of concurrent terminological phrases and their social uses (official, technical, professional or journalistic texts). So neologisms and archaisms become obvious. If terminology’s aim is to work on specialized names, professional contents and linguistic knowledge about terms, the Web is a very good tool to validate technical translations, provided philological precautions.*

## 1. Introduction

Le Web est avant tout une ressource documentaire considérable, grâce à ses moteurs de recherche, ses portails et ses sites spécialisés. Or ce qui est bon pour la documentation l’est toujours plus ou moins pour la terminologie. Le problème qu’il pose n’est pas le silence mais le bruit. Comment le résoudre? Par l’expertise du domaine, pour une part. Par des critères linguistiques, surtout, et c’est ce que je voudrais montrer, en mettant en évidence ce qui est réalisable sur la base de connaissances solides en grammaire, en lexicologie et en traduction. Il faut pour cela se donner une méthode générale d’investigation, prendre la mesure du possible, observer l’intérêt particulier des fréquences relatives et voir les gains possibles pour la terminographie. Le recours à des méthodes philologiques va avec un souci d’explicitation maximale des conditions de ce type de recherche. Dans le cas présent, le moteur utilisé est Google, et la date des requêtes celle de la dernière consultation opérée au terme de l’expérience, du fait des variations possibles d’un jour à l’autre et même d’un instant à l’autre.

## 2. Le besoin de partir de dénominations longues

Si l’on admet que les termes sont des dénominations spécialisées, ce sont à l’écrit des chaînes de caractères qui se rencontrent principalement dans des textes destinés à une communication plus ou moins professionnelle

par son objet (un domaine de connaissances) et par ses partenaires (présûmés compétements ou désireux d'apprendre). Rien ne permet malheureusement de les distinguer, les termes étant au demeurant des mots comme les autres. Ce qui est possible en revanche, et très rentable, c'est de chercher sur le Web ceux que l'on connaît et de les utiliser comme voies d'accès à d'autres chaînes de caractères récurrentes dans leur voisinage.

On peut bien entendu trouver les chaînes de caractères les plus dignes d'intérêt auprès de professionnels, dans des documentations, dans des encyclopédies ou des dictionnaires spécialisés. Une certaine familiarité avec le domaine permet toutefois des gains de temps et de sécurité appréciables; bien plus, les autres aides à attendre d'Internet, comme les prises de contacts, les consultations par courrier électronique et les téléchargements, supposent pour être efficaces un minimum de culture spécialisée. En toute hypothèse l'usage d'un moteur de recherche généraliste en terminologie n'est pertinent qu'à partir d'une connaissance préalable de la nomenclature de base du domaine.

En quoi peut-on parler d'une méthode philologique? La philologie traite traditionnellement des textes anciens, la terminologie est classiquement synchronique. En dépit de cette différence, les points communs sont nombreux: importance de l'écrit, de la validation des textes, d'une signification non évidente des expressions (justifiant d'un côté des glossaires, de l'autre des dictionnaires spécialisés). Ce qui est bon pour la première l'est aussi en partie pour la seconde: critique des sources, représentativité des textes, pertinence des gloses.

Les enjeux justifient des précautions dans ces matières, d'autant que le Web charrie des flots d'information non pertinente pour la terminologie: mercantile, ludique, anecdotique etc.. La perte des majuscules, des accents et des signes diacritiques de toutes sortes constitue déjà un obstacle suffisant. Mais une étude philologique de l'actualité et de la nouveauté terminologiques bien conduite a toutes chances de déboucher sur une connaissance à la fois vaste et précise des expressions usuelles, des connaissances dénommées, des catégories d'usagers et des profils linguistiques d'usages.

Pourquoi privilégier les dénominations longues? Pour limiter le bruit, qui est assourdissant si l'on n'y prend pas garde. Par exemple, l'euthanasie est un problème qui est au cœur de débats juridico-éthico-médicaux, et sa dépénalisation est matière à controverse, comme celle de diverses drogues un peu partout ou de la prostitution là où elle est interdite. Si l'on s'intéresse au cas de la dépénalisation de l'euthanasie, il faut se garder de requêtes sur "dépénalisation" (19600 occurrences sous Google le jour de la dernière relecture du présent article, soit le 14 juin 2004) et sur "euthanasie" (302000): de telles quantités sont ingérables et inégalement pertinentes. Avec une requête (entre guillemets, car il s'agit d'expressions linguistiques) concernant "dépénalisation de l'euthanasie" (672), on est encore face à des dizaines de pages, mais au bout des pre-

nières on dispose déjà d'indications importantes sur l'invariabilité de la forme, les contenus, les enjeux et les prises de positions.

Bien entendu, il s'agit ici d'un cas particulièrement facile, du fait de la monosémie de *dépénalisation* et d'*euthanasie*, et aussi de l'invariabilité en nombre du seul déterminant usuel, l'article défini. Il en va de même pour la dénomination verbale du même contenu de pensée: *dépénaliser l'euthanasie* (253), qui autour de l'infinitif cristallise une prédication du deuxième ordre (objet de prédications au moyen de *falloir*, *refuser de*, *accepter de*, *viser à*, *décision de* etc..).

De même que le philologue a besoin de connaissances solides en langues, linguistique et civilisation, le terminologue se conçoit difficilement sans compétences en ces matières. Il est de bon ton chez les traducteurs de s'amuser du nombre de philologues reconvertis dans la terminologie; on peut aussi penser que tant de convergences ne sont pas simplement fortuites ou opportunistes.

### 3. Le besoin de critères linguistiques

Il est clair qu'un nom qui se décline, un verbe qui se conjugue, un adjectif ou un déterminant qui s'accordent, posent des problèmes d'identification des formes morphologiques et des classes grammaticales qui ont motivé la mise au point de divers logiciels (voir par exemple Pavel et Nolet 2001). Leur exploitation n'est techniquement possible que sur un sous-ensemble de textes téléchargés, ce qui fait perdre le bénéfice quantitatif d'une information maximale, globalement significative par son caractère massif.

L'un des instruments les plus classiques est le concordancier. Sur le Web, il faut substituer à son exploitation, devenue impossible, un attachement systématique aux entourages syntagmatiques, en compensant les pertes d'automatismes par une interprétation syntactico-sémantique, qui de toute façon serait requise en aval de l'utilisation de concordanciers. Voyons donc le parti qui peut être tiré des classes d'objets et des schémas d'arguments.

#### 3.1. Pertinence des classes d'objets

La définition de ce qu'est le vin relève d'une ontologie de sens commun: en gros, on s'accorde pour y voir un produit de la fermentation alcoolique de raisins. Linguistiquement, *vin* a une forme phonique et graphique, appartient à la catégorie des noms, a des hyperonymes selon plusieurs facettes (*liquide*, *boisson*, *marchandise*), des hyponymes (*rosé*, *primeur*, *chardonnay*, *bourgogne* etc., selon la couleur, la vinification, le cépage ou l'origine), des synonymes selon les niveaux de langue (*nectar*, *pinard* etc.) et entretient des relations méronymiques

avec les noms des composants du vin (*acide, tanin, alcool, glycérine, arômes*).

*Vin* est aussi caractérisé par une combinatoire syntactico-sémantique (cooccurents les plus appropriés): *fût de vin, bouteille de vin, cubi(tainer) de vin* etc., mais aussi *boucher, embouteiller, filtrer, soutirer, tirer* etc.. Les contenants et les opérations appropriés sont applicables aux noms de certains autres produits, comme *huile*, mais une prédiction sûre est possible : de telles combinaisons valent en particulier pour tout vin, quels qu'en soient la couleur, les méthodes de vinification, le cépage ou l'origine. Ces noms de vins ne sont pas seulement des désignations d'objets du monde, ce sont aussi des substantifs dotés de cooccurents prédictifs appropriés comme *bouchage, embouteillage, filtrage, soutirage, tirage* etc., et c'est en pareil cas qu'on parle de "classes d'objets" (au sens de Gross et Guenther 2002: 181).

Le degré de finesse dans l'étude des compatibilités syntagmatiques à base syntactico-sémantique atteint par les classes d'objets va bien plus loin que les "traits" généraux comme "humain", "animé", "matériel" etc., qui sont en usage dans les ontologies. L'utilisation cumulative de ces traits et de classes d'objets permet d'éclairer des différences de comportements syntaxiques tels que ceux de l'adjectif *tranquille*, ordinairement qualificatif mais non dans le cas d'un <vin>:

- N <humain> *très tranquille*
- N <lieu> *très tranquille*
- N (vin) \* *très tranquille* -> *vin tranquille* (antonyme de *mousseux* et *effervescent*)

Quel est l'intérêt des classes d'objets dans l'exploitation du Web à des fins de recherche terminologique? Celui de ne pas se perdre dans une masse énorme d'informations en sélectionnant des chaînes de caractères assez longues pour lever la polysémie des mots et assez typiques pour contenir une information restreinte à un domaine (considéré comme un ensemble de classes d'objets et d'expressions prédicatives appropriées). Par exemple, si je veux savoir quelle différence il y a entre une *pièce de vin* et une *barrique de vin*, il y a deux écueils à surmonter: surtout le trop général (pour *une pièce*, 447000 occurrences), mais aussi le trop particulier (pour *pièce de meursault*, 3). Le bon niveau est celui du générique, donc *pièce de vin* (202) et *pièces de vin* (322). Les contextes rencontrés permettent d'apprendre ou de vérifier que la *pièce* équivaut à peu près à une *barrique bordelaise* (mais celle-ci fait précisément 225 litres) et aussi que le terme *pièce* est surtout vivant en Bourgogne. En outre, on obtient en prime la chaîne *demi-pièce de vin* (2), qui en retour vient compléter la liste de la classe d'objets des <fûts>.

### 3.2. Pertinence des schémas d'arguments

Un <vin> se stocke, se vend etc. comme tout produit, il se met en bouteille comme d'autres denrées liquides, mais il a aussi des prédicats plus typiquement appropriés, tels que *chambrier* ou *champagner*. Bien plus, on ne sert *chambré*, professionnellement, qu'un vin appartenant à la sous-classe du <vin rouge>. Un schéma d'arguments est la suite constituée par une expression prédicative (comme *chambrier* ou *chambré*) et les arguments auxquels elle s'applique typiquement (classes d'objets ou non).

La polysémie des verbes a beaucoup à voir avec les schémas d'arguments. Ainsi, *chambré* se dit d'un <cylindre métallique>, d'un <humain> ou d'un <vin rouge>, avec des sens entre lesquels on trouverait difficilement une intersection. Les synonymes, aussi bien, sont clairement distincts: respectivement *mis au point*, *raillé* et *à température ambiante*.

L'intérêt heuristique des schémas d'arguments est qu'il conduit à "la suite la plus longue" (Gross et Guenther 2002: 180) des arguments. Ainsi, *vendanger* est une expression prédicative qui suppose un sujet (des vendangeurs), un objet (des raisins), un résultat (la vendange comme récolte), un lieu (un vignoble), un temps (la saison des vendanges) et un moyen (à la main ou à la machine). Un tel schéma, inspiré de la grammaire des cas mais enrichi par des classes d'objets, rend explicite la polysémie de *vendange* (nom d'action, de résultat et de temps) et prend en compte la connaissance cognitivo-linguistique du domaine: <raisin: cabernet franc, merlot ...>, <vignoble: clos, château, domaine ...> etc..

Il n'est pas facile, pour autant, de procéder à des tris automatisés. Par exemple, on vendange des raisins, mais on vendange aussi, par métonymie, leur lieu de production (ex.: 18 fois *vendanger les vignes*). De même, s'il est globalement satisfaisant de distinguer les schémas "*acquitter* un <humain>" (it. *assolvere*), "*acquitter* une <somme d'argent>" (it. *quietanzare*) et "*s'acquitter* d'une <obligation>" (it. *adempiere*), dans le détail on ne peut pas se contenter de telles indications (comme dans *Quadrirédacteur*). Quand on multiplie les contextes, comme le permet le Web, on rencontre des cas où l'on acquitte des montants (3 fois *acquitter en euros la taxe*, par exemple), d'autres (9) où l'on doit *s'acquitter de droits de succession* (qui sont eux aussi monétaires, mais également obligatoires, en sorte que l'on dit aussi (109 fois) *acquitter des droits de succession*). Ainsi, le Web pousse à la faute les modèles, ce qui est scientifiquement stimulant même si c'est inconfortable (mais l'important est l'explicitation des difficultés, si l'on a en vue la traduction la plus idiomatique possible). On est loin des annonces de merveilles à court terme, mais on respecte ainsi les faits, qui auront toujours le dernier mot. Rien n'interdit une double appartenance pour les droits de succession: à la classe des <sommes d'argent> et à celle des <obligations>. Mais aucun automate ne permet de lever à coup sûr la polyvalence de *droit* en pareil cas.

#### 4. Le besoin de contextes définitoires

Bien loin que le Web soit une panacée, il faut au contraire avoir bien cerné au préalable un objet de requête soigneusement délimité avant de l'interroger. Autant le texte unique permet (laborieusement) de sous-traiter l'analyse grammaticale, la lemmatisation et les concordances, qui ne dispensent pas, au demeurant, d'une interprétation à ses risques et périls, autant l'espace textuel éclaté d'Internet (le contraire d'un corpus au sens classique) impose des stratégies de recherche ad hoc.

##### 4.1. Réduire la polysémie

L'univocité a toujours été pour les terminologues un idéal, même si les plus lucides ont toujours su que c'était un "vœu pieux" (Wüster 1985: 79). On n'a pas toujours à traduire une chaîne de caractères aussi facile que *dépénalisation de l'euthanasie*, pour reprendre un exemple choisi délibérément pour sa commodité. Ainsi, si l'on a à trouver la meilleure traduction de *rafle* au sens végétal (la grappe de raisin moins les grains) dans les dictionnaires d'entreprises consultables en ligne, il faut avant tout se débarrasser d'un homonyme, la forme verbale relevant de *rafler*. Il faut aussi écarter le sens policier. On peut obtenir assez vite de bons résultats, vu la rapidité des réponses, mais à condition de prévoir toutes les chaînes distinctives : *rafle de raisin* (11), *rafles de raisin* (39), *rafle de raisins* (3), *rafles de raisins* (15), *rafles des raisins* (6), *rafle du raisin* (23). La précision des requêtes est alors suffisante pour que tous les contextes (botaniques, vitivinicoles et lexicographiques) soient pertinents, alors que la forme *rafle* est extrêmement bruyante (56600).

La polysémie à base de métaphore ne pose pas de problème particulier là où une classe d'objets permet de la traiter à part. Il est bien vrai que d'un point de vue cognitif "polysemy is functional. It is the consequence of meaning evolution" (Temmerman 1997: 62); en revanche, une image comme celle des perles, qui est usuelle en allemand et en anglais pour caractériser tout <vin> que le français appelle *effervescent* ou *pétillant*, est lexicalisée par les dénominations *Perlwein*, *pearl wine* et *pearly wine*, qui constituent de bonnes chaînes de caractères pour des requêtes à partir de ces langues.

L'extension d'une notion "humaine" à de l' "inanimé" ne crée pas non plus de difficulté insurmontable si l'on prend en compte les schémas d'arguments. En français, par exemple, le mot *séjour* est usuel dans la vie quotidienne (avec *salle de séjour* ou, par métonymie, *séjour*), en droit (avec *l'interdiction de séjour*) et dans le domaine du vin (*séjour en fût / en bouteille*). En droit, les équivalents sont respectivement *Aufenthalt* et *residence* (voir notamment *Quadrirédacteur*); en matière de vinification, ce sont des termes différents: *Lagerung* et *stay*. Pour le droit, la dénomination usuelle *titre de séjour* est un bon accès à l'usage des traités euro-

péens, qui conduit à *Aufenthaltstitel* et à *residence permit*, donc il est possible et sûr à la fois de faire l'économie du recours à un dictionnaire spécialisé en partant d'une chaîne longue. Pour le vin, il faut avoir présent à l'esprit le scénario du séjour en pareil cas: il s'agit d'un <vin>, et le lieu approprié du prédicat *séjour* est soit un <fût> soit une <bouteille>. D'où la pertinence d'une chaîne comme *séjour en fûts de chêne* (251), grâce à quoi l'on accède à *Lagerung* et à *stay* via des formulations usuelles dans la profession: *Lagerung in Eichenholzfässern* (230) et *stay in oak barrels* (17).

#### 4.2. Confronter les définitions

Pour comprendre avec précision ce qu'est la rafle du raisin, il y a bien sûr les dictionnaires spécialisés. Malheureusement les lexiques multilingues se contentent souvent d'étiquettes sans justifications ni modes d'emplois, les bilingues numérotent des sens faute de place pour faire plus, et les unilingues ne sont pas forcément à jour.

A jour, le Web l'est parfaitement, par nature. Pour la langue ordinaire, il fournit plus d'exemples d'emplois que d'énoncés définitoires, mais en sélectionnant les sources on y trouve un accès à des travaux lexicographiques de qualité comme, pour l'italien, le dictionnaire de Tullio de Mauro. Pour les langues spécialisées, les sources professionnelles, officielles, techniques et médiatiques concourent à fournir des profils conceptuels complémentaires dans lesquels il y a place pour des informations encyclopédiques.

De fait, les connaissances spécialisées appellent des informations encyclopédiques. Par exemple, avant de décider de traduire l'allemand *Qualitätswein mit Prädikat* ("with distinction"? "con calificación"?), ou plutôt de conserver l'expression allemande telle quelle, comme intraduisible, il faut regarder suffisamment d'occurrences (assez peu, on sait vite de quoi il s'agit), pour apprendre non seulement que c'est le haut de gamme en pays germanophone, mais aussi que c'est parce qu'il ne contient pas de sucre ajouté dans le moût. L'explication va ici normalement avec la dénomination, car il s'agit précisément de dénommer des connaissances techniques.

#### 5. Le besoin de validations

Si un terme est une "unité lexicale définie dans les textes de spécialité" (Kocourek 1991: 180), ce qui semble être une définition réaliste, la grande affaire est la détection de textes représentatifs de la spécialité considérée. Or il faut reconnaître que les premières pages fournies par un moteur de recherche généraliste tendent à favoriser l'information commerciale et promotionnelle. Toutefois, ce qui importe est la fréquence globale et la

répartition entre catégories d'usagers. Par exemple, les publicités pour des balances hydrostatiques italiennes à usage viticole sont à prendre en compte en même temps qu'un récent règlement de la Commission européenne (du 23 janvier 2004) rendant obligatoire cet instrument de densimétrie : *balance hydrostatique*, du nom d'une invention de Galilée, est ainsi doublement pertinent dans le secteur du vin.

Influencé par la culture juridique, pour compléter ma base *Quadrirédacteur* là où mes sources communautaires (européennes) ou fédérales (suisses) ne suffisaient pas, j'ai pris en compte leur distinction classique entre textes de la norme (législative et réglementaire), de la doctrine (des professeurs de droit) et des pratiques (juges, notaires etc.). J'ai dû ajouter les médias, car plus j'allais à la pêche dans l'océan du Web, plus je m'apercevais que des expressions courantes mais encore récentes comme *des incivilités* (3920) ou *les sans papiers* (30900) se lisent surtout dans la presse.

La rencontre plus tardive d'une autre terminologie, celle du vin (voir Lerat 2004), confirme la pertinence d'une distinction entre textes officiels (directives européennes, traité entre l'UE et l'Australie, normes nationales), scientifiques et techniques (œnologie et autour), professionnels (production et commercialisation) et médiatiques (sites de groupements, presse spécialisée, publicité rédactionnelle) et l'intérêt particulier des expressions communes à plusieurs de ces catégories.

Ce modèle général à base de types d'émetteurs et de supports semble propre à une évaluation pondérée de la vitalité des dénominations. Ainsi, *dépénalisation de l'euthanasie* retient l'attention à un quadruple titre: l'expression est officialisée par le Législateur belge depuis 2001, c'est la matière de travaux de biologistes et de juristes, de débats dans les professions de la santé et aussi dans les médias.

La validation ne peut en effet pas résulter simplement d'une fréquence brute, pour plusieurs raisons. D'abord les mêmes références sont souvent largement récurrentes. Ensuite, des ressources terminologiques souvent utiles, comme *Eurodicautom*, n'apparaissent pas dans les requêtes au moyen de moteurs de recherches généralistes. Enfin, le nombre élevé des homonymes, noms propres ou raisons sociales, conduit à relativiser les quantités.

Il faut aussi tenir compte du fait que la notion même de texte est relativisée par le support. Ainsi, la diffusion de textes officiels sur des sites multiples pose le problème de la validité des reproductions, synthèses et citations. Pour les traités européens, seules les "versions consolidées" font foi, ce qui est sage vu les enjeux. Pour les mises à jour de dictionnaires d'entreprises ou de groupements, la précaution philologique à prendre est de signaler les dates de consultations, qui ne seraient pas pertinentes dans le cas d'un texte imprimé mais le sont dans celui d'un fichier numérique.

## 6. Le rendement d'une terminologie à base de Web

### 6.1. Indice de vitalité des expressions

Naguère, on disait *les sans domicile fixe*. Les administrations et la presse en ont fait *les SDF*. Une dénomination moins juridique au départ, mais soumise à son tour à une réglementation, est maintenant *les sans abri*. Si l'on veut avoir une juste idée du degré de vitalité de chacune de ces dérivations impropres, rien de plus facile: respectivement 978, 8100 et 14600.

Ce n'est pas seulement la vie sociale qui est sujette à la variation synonymique. En fait, on trouve aussi des concurrences durables dans les sciences et dans les techniques. Ainsi, *capsuleuse* (746) a plus de succès que *capsulateur* (65), *filtration du vin* (63) plus que *filtrage du vin* (9) (voir Lerat 2004). La chimie ne fait pas exception: *le lévulose* (27) ou *la lévulose* (5) n'est rien d'autre que *le fructose* (3100) ou *la fructose* (187).

Il ne suffit pas qu'une opération soit normalisée pour que la dénomination soit unique. Par exemple, les équivalents de *sulfitage* sont en anglais *sulphitation* (386) et *sulphiting* (309), en espagnol *sulfitación* (49) et *sulfitado* (87, précédé de *el* pour désambiguïser), en italien *solfitaggio* (48) et *solfitazione* (293), en portugais *sulfitagem* (22) et *adição de sulfito* (33).

La question de l'idiomaticité, importante en traduction spécialisée, est souvent affaire, ici comme ailleurs, de choix lexicaux entre synonymes. C'est vrai pour un vocabulaire très concret, comme celui du vin : *embouteillage du vin* ou *mise en bouteilles du vin* ? La réglementation européenne préfère le premier terme (50), mais le second est le plus vivant chez les professionnels, sous les formes *mise en bouteille du vin* (210) et *mise en bouteilles du vin* (35). C'est vrai aussi pour le vocabulaire juridique au sens large, où existent des registres de langues (*allouer une aide*: 371 / *accorder une aide*: 6780) et même des doubles, triples, voire quadruples emplois officiels (*avantage fiscal*: 36200 / *déduction fiscale*: 12000 / *abattement fiscal*: 2530 / *réduction fiscale*: 2350).

Les contacts de langues mettent en concurrence des formes autochtones, des emprunts et des calques. Voici quelques exemples qui montrent une situation équilibrée de l'allemand à cet égard:

- fr. *macération carbonique* = de. *Kohlensäuremazeration* (126) / *Kohlen-säuremaischung* (84)
- fr. *oenologie* = de. *Oenologie* (170, discriminé par l'article *die*) / *Weinwissenschaft* (427)
- fr. *vin de garde* = de. *Reservewein* (67) / *lagerfähiger Wein* (392)
- fr. *vinification* = de. *Vinifikation* (4400) / *Weinbereitung* (12200), *Weinerzeugung* (4450), *Weinbehandlung* (1230)

## 6.2. Indice de vitalité des réalités

Contrairement au droit, la vigne ne prospère pas sur n'importe quel sol. L'Europe communautaire doit accepter, pour ne pénaliser personne, que chaque Etat conserve dans toute la mesure du possible ses traditions, y compris les palliatifs au manque d'ensoleillement. L'Europe non méditerranéenne a donc droit à un "enrichissement" de ses jus de raisin qui utilise en particulier une méthode de sucrage des moûts mise au point par le savant français Chaptal, d'où son nom: *chaptalisation*.

En appliquant les règles de dérivation de quelques langues, on peut trouver sans dictionnaire les formes correctes: de. *Chaptalisierung*, en. *chaptalization*, es. *chaptalización*, it. \**chaptalizzazione*, pt. *chaptalização*. Ce qui est intéressant, en revanche, c'est que le nombre des attestations est conforme à la géographie: respectivement 11300, 839, 18, 0, 1.

Le cas de l'italien mérite un commentaire: la forme est disponible, mais personne n'en veut, et la traduction officielle est *zuccheraggio*. Linguistiquement, on peut considérer qu'il s'agit seulement de hiérarchie sémantique: à défaut d'équivalent on recourt à l'hyperonyme, selon un mécanisme traductif décrit sous le nom d'*hyperonomase* (voir Van Campenhout 2001: 13). Economiquement, comment ne pas penser à la dénonciation d'une pratique réduite polémiquement à un vulgaire "sucrage"?

Une autre pratique traditionnelle dans une partie de l'aire de production vinicole européenne, notamment dans le Bordelais, est l'assemblage de vins issus de plusieurs cépages de même origine et de la même année. Cette fois l'italien est présent, avec *assemblaggio di vini* (66), mais c'est l'anglophonie qui est hésitante: face au petit nombre d'attestations des expressions *assemblage of wines* (38) et *wine assemblage* (27), les dénominations courantes sont *wine blend* (2460) et *wine blending* (679), par une hyperonomase qui met sur le même plan la typicité (par exemple merlot + cabernet sauvignon + cabernet franc très souvent dans le Médoc) et le simple mélange de petit vin et de vin "médecin".

Ces deux exemples le montrent, le Web n'est pas seulement un bon révélateur des réalités géographiques, mais aussi des prises de position. Un néologisme cristallise bien l'importance du lobbying et de la médiatisation dans le commerce des vins: *parkerisation*. Le mot existait déjà, mais pour désigner un procédé de protection superficielle de pièces métalliques par traitement thermo-chimique, du nom de l'inventeur. Cette fois il s'agit de l'influence à l'échelle mondiale de Robert Parker, dégustateur américain de très grande réputation, d'où les emplois suivants: pour *parkerisation*, 13 sur 59; pour *parkerization*, 15 sur 200; pour *parkerizzazione*, 2 sur 32.

### 6.3. Indice de vitalité des néologismes

La néologie spécialisée est un phénomène étroitement dépendant de la circulation des connaissances spécialisées. Celle-ci ne peut être contrariée durablement sur Internet, lieu où la liberté va jusqu'à la licence.

Pour les traducteurs et rédacteurs spécialisés, une information sur l'usage réel de mots "possibles" est indispensable, notamment pour passer d'une formulation verbale à une formulation nominale. Ainsi, *chamber un vin* est usuel, mais *chambrage* n'a ce sens qu'une fois (et dans une publicité) sur 623 attestations de ce nom, qui n'est guère vivant que dans le secteur de l'armement. A éviter, donc.

Inversement, un néologisme de sens très vague mais devenu courant peut être facile à traduire si la conceptualisation incertaine est partagée comme un fait de civilisation internationale.

Ex.: *incivilités* (17700), intersection flottante de faits de société appréhendés par les sociologues, les juristes, les philosophes et les politiques, se traduit sans difficultés par en. *incivilities* (3690), de. *Unhöflichkeiten* (699) et it. *le inciviltà* (110).

Appliquée aux néologismes, la méthode des chaînes longues est révélatrice des contenus et permet de mieux cerner ce qui est en cause. Par exemple, pour l'italien, la chaîne *inciviltà d* fait apparaître des comportements situables entre *maleducazione* et *microcriminalità*: *inciviltà dei automobilisti, della scuola, dei tifosi* etc..

### 6.4. Indice de vitalité des archaïsmes

Il est souvent possible de dater l'apparition d'un mot spécialisé, mais les obsolescences sont plus difficiles à consigner; prudemment, les lexicographes ne parlent jamais de disparition, mais se contentent de mentions comme "vieux" ou "vieilli". Un avantage du Web est de fournir des indices numériques interprétables.

En matière de terminologies, les dénominations de connaissances sont particulièrement évolutives, y compris quand les connaissances sont relativement stables. Ainsi, la séparation de la rafle et des grains de raisin est une opération incontournable chez les producteurs de vin depuis l'Antiquité. De ce fait, elle a reçu plusieurs dénominations plus ou moins dialectales. On pourrait parler de *déraflage*, mais ce mot a laissé la place à *éraflage* (785), en dépit de la polysémie d'*érafler*, avec le soutien d'un nom d'appareil, *érafleuse*; on peut aussi conceptualiser l'opération autrement, et parler d'*égrappage*, ce qui était déjà le cas au XVIII<sup>e</sup> siècle et se fait toujours (1790).

Dans le même domaine, la survivance des dénominations traditionnelles s'observe là même où la technologie a évolué, c'est-à-dire un peu partout. Ainsi, la pièce permettant d'ouvrir et de fermer l'écoulement du vin hors d'un fût n'est plus toujours en bois comme autrefois, mais on

utilise encore pour la nommer *cannelle*, *canule*, *champlure*, *chantepleure* etc.. Il est difficile d'utiliser le Web pour évaluer la vitalité de *cannelle* et de *canule* à cause de leur polysémie, mais on peut observer que *chantepleure* (614) et *champlure* (103) sont encore bien vivants.

Dans le domaine juridique, la méthode des chaînes longues trahit le caractère routinier non pas du français, mais d'administratifs et d'administrés francophones. Voici quelques coexistences tranquilles d'archaïsmes connotés et d'expressions équivalentes mais non endimanchées:

- *acquitter une redevance* (248) / *payer une redevance* (2150)
- *octroyer un crédit* (434) / *accorder un crédit* (2390)
- *ajournement du procès* (457) / *renvoi du procès* (1540)
- *appréhender un suspect* (75) / *arrêter un suspect* (130)
- *créancier chirographaire* (268) / *créancier ordinaire* (210)

## 6.5. Une aide à l'harmonisation

La connaissance la plus large possible des usages de mots spécialisés est le préalable scientifique à toute harmonisation terminologique. En matière vinicole, l'Europe se fait d'un côté sur fond de traditions nationales et locales, où le goût s'exprime avec les mots du terroir, de l'autre dans le cadre d'une mondialisation où les groupes de pression font ce qu'ils peuvent soit pour maintenir la diversité soit pour promouvoir des standards planétaires.

Le souci de la protection du consommateur conduit ici, comme dans la normalisation industrielle, à celui de la prévisibilité des dénominations. Il en résulte une terminologie officielle (ci-dessous O) qui ne périclète pas les usages professionnels (ci-dessous P).

Ex: fr. *vin pétillant* (O, 3410), *vin effervescent* (P, 2530) = de. *Perlwein* (O, 14400), *spitziger Wein* (P, 18), *schäumender Wein* (P, 110) = en. *semi-sparkling wine* (O, 305), *gassy wine* (P, 50), *effervescent wine* (P, 526), *pearl wine* (P, 1700); *pearly wine* (P, 38) = es. *vino de aguja* (O, 555), *vino efervescente* (P, 65) = it. *vino frizzante* (O, 9610), *vino effervescente* (P, 108) = pt. *vinho frisante* (O, 154), *vinho efervescente* (P, 5)

Ces séquences visant la distinctivité par rapport aux mousseux rappellent l'importance de la mémoire (individuelle et collective) dans la dénomination des choses et des concepts. Un cas très caractéristique à cet égard est celui des expressions comportant un nom propre comme, pour le vin, *degré Gay-Lussac*, *degré Oechsle*, *pasteurisation*, *méthode Charmat* ou, pour le droit, *espace Schengen*, *traité d'Amsterdam*, *convention de Genève* etc..

C'est encore la mémorisation qui entretient des routines énonciatives et rédactionnelles comme les suivantes : *apposer sa signature* (3410)

est dix fois plus attesté que *mettre sa signature* (327), *affranchir l'enveloppe* (90) cinq fois plus que *timbrer l'enveloppe* (28)!

Cela ne veut pas dire que toute action sur les langues spécialisées soit vouée à l'échec, mais au contraire que c'est une question multidimensionnelle qui doit être traitée, si l'on veut qu'elle le soit, en prenant en compte ce que révèle le Web pour chaque langue, dans sa géographie, sa sociologie et sa syntagmatique. Ex.: entre *transfrontière* (17300 au singulier, 30300 au pluriel) et *transfrontalier* (46900 au masculin singulier, 48200 au pluriel, 66800 au féminin singulier, 35300 au pluriel), le second a pour lui la fréquence et le premier l'invariabilité en genre.

## 7. Conclusion

Ainsi, le Web, à la fois par la quantité et par la diversité des textes et énoncés qu'il rend immédiatement accessibles, apparaît comme une ressource incomparable pour la terminologie. Certes, il est bien vrai qu'Internet "fournit un large éventail de sources documentaires mais plutôt éphémères et de valeur très inégale" (Pavel et Nolet 2001: 8); toutefois, il est possible de procéder à des requêtes à intervalles plus ou moins éloignés (c'est la bonne façon d'observer l'évolution de la vitalité des dénominations spécialisées) et de n'exploiter pour des études de contextes que ce qui bénéficie d'une signature digne de foi.

L'un des problèmes majeurs de la terminologie étant la validation (que les pouvoirs publics visent par la normalisation, l'harmonisation et l'officialisation), le Web apporte à cette question la réponse de l'usage effectif à grande échelle. Cet usage demande à être interprété, mais il en va de même pour toute étude de texte, depuis la philologie la plus ancienne.

La traduction spécialisée peut bénéficier d'investigations conduites langue par langue, notamment dans les textes officiels où chaque version fait foi (pour les équivalents et parfois pour les définitions), mais aussi dans les écrits techniques, professionnels et médiatiques (pour les connaissances spécialisées et pour les cooccurrents typiques).

En outre, l'histoire des vocabulaires spécialisés peut désormais dépasser plus facilement une limite épistémologique de la terminologie classique: son caractère achronique. Encore faut-il prendre des précautions philologiques: critique des sources, chaînes de caractères désambiguïsées, pluridisciplinarité.

## Bibliographie

- Gross, G. et F. Guenther (2002). "Comment décrire une langue de spécialité?." *Cahiers de lexicologie* 80 (1), 179-199.
- Kocourek, R. (1991). *La langue française de la technique et de la science* (2<sup>ème</sup> éd.). Wiesbaden: Oskar Brandstetter.

- Lerat, P. (1995). *Les langues spécialisées*. Paris: Presses Universitaires de France. Trad. esp. *Las lenguas especializadas* (1997). Barcelone: Ariel.
- Lerat, P. (2004). "Dénominations spécialisées, connaissances professionnelles et connaissances linguistiques. Le cas du vin". A paraître dans *Actas del primer congreso sobre la traducción del lenguaje especializado de la vid y el vino* (Soria, 3-6 mars 2004).
- Pavel, S. et D. Nolet (2001). *Manuel de terminologie / Handbook of Terminology*. Ottawa: Travaux publics et Services gouvernementaux du Canada.
- «Quadrirédacteur». En ligne sur <http://www-lli.univ-paris13.fr/ressources/quadri> (consulté le 14-06-2004)
- Temmerman, R. (1997). "Questioning the univocity ideal. The difference between sociocognitive Terminology and traditional Terminology." *Hermes. Journal of Linguistics* 18, 51-91.
- Van Campenhoudt, M. (2001). "Pour une approche sémantique du terme et de ses équivalents." *International Journal of Lexicography* XIV (3), 181-209.
- Wüster, E. (1985). *Einführung in die Terminologielehre und terminologische Lexikographie*. Copenhague: Ecole des Hautes Etudes Commerciales.