

*Pierre Lerat**

La combinatoire des termes. Exemple : *nectar de fruits*

Abstract

Terminology is henceforth based on textual analysis, although its methodology remains paradigmatic. Concordances need to be interpreted, and a theory of combinations is needed. In this article, Harris' (1968, 1976, 1988) « operator-argument » theory is used, though words, concepts and terms are strictly distinguished. Specialised texts are made up of words as are every text, but some of them denote special concepts, and these are terms. The case of *nectar de fruits*, as used in EU regulations and FAO *Codex alimentarius* and on corporate websites, is used to show the relevance of combinatory analysis of concepts and of terms in establishing the ontology of the fruit industry.

0. Introduction

L'accès à des corpus textuels spécialisés étant devenu facile, le besoin d'une « terminologie textuelle » (Slodzian 2000) impose une méthodologie appropriée. Or celle de la terminologie reste paradigmatique. Pour avancer, il faut une réflexion théorique, car le pragmatisme des collocations ou de la phraséologie montre vite ses limites aux linguistes, et l'intérêt des informaticiens pour des relations « transversales » (Grabar/Hamon 2004: 63) manque de critères de la transversalité, tout comme chez Sager (1990: 34-35), où la définition des « complex relations » « ne s'appuie sur aucun fondement théorique » (Grabar/Hamon 2004: 63).

La proposition qui est présentée ici est fondée sur une approche productive en linguistique générale, le modèle « **opérateur-argument** » de Harris (1971, 1976, 2007). Ce qui est souvent retenu de Harris est l'analyse distributionnelle, applicable à toutes les chaînes de caractères

* *Pierre Lerat*
Université Paris XIII
Laboratoire Lexiques, Dictionnaires, Informatique
Avenue J.B. Clément
F-93430 Villetaneuse
pierre.lerat@wanadoo.fr

res ; son autre apport majeur, le modèle « opérateur-arguments », n'a pas seulement un grand intérêt pour « les langues de la science » (Harris 2007: 56-59), mais aussi pour les terminologies en général. En gros, ce mathématicien considère le verbe, dans la phrase simple, comme l'expression linguistique du prédicat dans la proposition élémentaire, et les groupes nominaux sujet(s) et objet(s) comme les arguments logiques du prédicat.

La terminologie constitue l'ossature conceptuelle des textes spécialisés. Il importe donc de bien préciser ce que sont les unités de compte : concepts et termes.

L'exemple étudié, celui de *nectar de fruits*, est peu technique mais fait néanmoins l'objet de stipulations dans la réglementation de l'UE¹, dans celle de l'OMS et de la FAO², et aussi de mentions normatives sur les étiquettes et les sites de professionnels³ (anonymés en l'occurrence). Chaque institution a sa « communication d'entreprise » (ou d'organisme), sa « politique linguistique » et sa « politique de communication » (Leroyer/Simonsen 2006: 288), et le terminologue est aussi concerné en tant que consommateur de produits manufacturés, mais ils ont en commun un accord (durable ou provisoire) sur ce dont ils parlent. Toute terminologie est en équilibre instable, et c'est pourquoi elle est nécessairement normative, qu'elle le dise ou non.

1. La définition communautaire de *nectar de fruits*

Selon D01JUS (I-I-a), le nectar de fruits est « le produit fermentescible mais non fermenté obtenu en ajoutant de l'eau et des sucres et/ou du miel aux produits définis aux points 1,2,3, à de la purée de fruits ou à un mélange de ces produits, et qui est en outre conforme à l'annexe IV ».

1 Avertissement : « Seule fait foi la version de la législation européenne telle que publiée dans les éditions du Journal officiel de l'Union européenne ». Le corpus principal est une directive de 2001 « relative aux jus de fruits ». Les abréviations utilisées visent la traçabilité philologique, indispensable dans le cas d'une terminologie qui se veut réemployable : *D01* pour « directive de 2001 », *JUS* pour « jus de fruits », *c 1* pour « premier considérant », *11-4-a-ii* pour « article 11, section 4, sous-section a, tiret ii », *A IV-2* pour « annexe IV, point 2 ».

2 Le *Codex alimentarius* [en ligne]. <http://www.fao.org>.

3 Énoncés suivis ici de la simple mention « pro », du fait de la difficulté d'obtenir des autorisations.

En tant qu'énoncé, cette définition illustre les difficultés de la terminologie textuelle et les limites de ce qu'on peut attendre de l'« extraction de termes » (voir notamment L'Homme 2004: 166 et Bourigault et al. 2004: 100).

- 1) Elle suppose des connaissances accessibles dans des « ressources externes » (L'Homme 2004: 223), plus ou moins techniques, pour comprendre *fermentescible*.
- 2) Elle utilise une anaphore générique (*ces produits*) et des renvois textuels internes (*points 1,2,3 et annexe IV*).
- 3) Elle utilise un verbe non spécialisé, *ajouter*, pour décrire le mode de fabrication et mettre en relation les ingrédients.

2. Le mot *nectar*

Le signifié du mot *nectar* est attaché, dans notre mémoire collective, à la boisson des dieux de l'Olympe : dire d'une boisson que c'est un vrai nectar n'est pas un petit éloge. Le nectar de fruits bénéficie de ce fait d'une connotation très laudative, alors que c'est un produit industriel faisant l'objet de normes professionnelles plus ou moins consensuelles, qui tentent de concilier les intérêts économiques et la protection du consommateur, et où sont pertinentes des considérations d'origine (CODEXALI, 4-5-2), de risques de rayonnements ionisants (D00ETI, 5-3) et d'OGM (P06ADD), à côté des tolérances ici ou là en matière d'additifs divers (P06ADD, A I) et même des « limites maximales de résidus de pesticides » (CODEXJUS, 6-1).

L'emploi technique du mot *nectar* est reflété par ses concordances, par exemple celle-ci :

- (1) fr. *la fabrication des nectars sans addition de sucres* = en. *fruit nectars manufactured without added sugar* = it. *la fabbricazione di nettari di frutta senza zuccheri aggiunti* (D01JUS, A I-I-4-a)

La polysémie interne au vocabulaire professionnel se perçoit également grâce aux cooccurents, comme le montre le cas d'*étiquetage* :

- (2) fr. « la liste d'ingrédients figurant sur l'*étiquetage* » = de. *Etikett* = es. *etiqueta* = it. *etichetta* = nl. *etiket* (D01JUS, c 7)
- (3) fr. « les modalités d'*étiquetage* » = de. *Etikettierung* = es. *etiquetado* = it. *etichettatura* = nl. *etiketteren* (D00ETI, c 16)

Pour un lecteur humain, la différence entre opération et objet est claire dans ces énoncés, mais il n'existe pas de logiciels permettant d'émuler cette compétence : *sirop de fruits* et *fabrication de sirops* ne sont pour eux que des segments nominaux (répétés ou non), comme *combustible des réacteurs* et *concentration du minerai* (voir Pavel/Nolet 2001: 44). Or avec *sirop de fruits* ou *combustible des réacteurs* il s'agit de « dénomination classificatrice » (Gross/Prandi 2004: 231), catégorématique, c'est-à-dire concernant un objet « référentiellement autonome » (Kleiber 1981: 39), dans le cas de *fabrication de sirops* ou de *concentration du minerai* on a un premier mot syncatégorématique : la dénomination d'une action (la fabrication) ou d'une propriété (la concentration) nécessairement applicable à un objet référentiel. Il importe de ne pas confondre les deux sortes de concepts.

3. Le concept de « nectar de fruits »

3.1. Concept et définition

Contrairement au signifié « en langue », la définition n'intéresse la linguistique que par sa mise en forme lexicographique ; ce n'est pas une affaire de langue, mais de logique (voir Mounin 1990: 134). En outre, contrairement à « une couche primordiale de concepts » qui appartient à « notre ontologie naturelle partagée » (Gross/Prandi 2004: 12), comme l'idée de cause ou celle de but, le concept spécialisé est un artefact : « une unité de connaissance créée par une combinaison unique de caractères » (ISO 1987, 2001).

Le caractère n'est pas un « trait sémantique » mais une propriété reconnue par les spécialistes à une entité concrète ou abstraite ; ainsi, ce n'est pas au mot *nectar* qu'est reconnue la propriété d'être « fermentescible », mais à un produit dont la composition est normalisée et qu'il est convenu, entre professionnels, associations de consommateurs et pouvoirs publics, d'appeler *nectar de fruits* (et, par abréviation, fr. *nectar*, de. *Nektar*, it. *nettare* etc.).

La terminologie repose nécessairement sur un consensus d'experts, faute de quoi elle se dissout dans une lexicologie aléatoire, comme l'a bien vu Alain Guillet :

Les substantifs classifieurs, bien que sensiblement plus précis et opératoires que les traits sémantiques, connaissent donc les mêmes limi-

tes, comme probablement toute tentative pour faire correspondre directement le réel et la langue. Dans une telle correspondance en effet trois niveaux semblent intervenir : la connaissance objective du monde, l'image que s'en fait le locuteur (sa représentation du monde) et l'usage linguistique proprement dit (comment il combine les mots pour décrire (1986 : 100).

Non seulement c'est une « combinaison unique de caractères » qui fait la prévisibilité des boissons non alcoolisées commercialisées ou de tout autre produit industriel, mais « toutes les versions d'une définition dans les différentes langues doivent comprendre les mêmes caractères » (ISO 860, 1996 in Depecker 2000: 117), et seule « la formulation des définitions dépend des règles propres à chaque langue » (ISO 860).

3.2. Les définitions de « nectar de fruits »

La définition communautaire de 2001 est plus complète que la précédente (de 1993) où il fallait attendre un article complémentaire lointain (D93JUS, 6-2-c) pour apprendre que le miel était acceptable « en remplacement total des sucres ». Elle l'est moins que celle de l'OMS et de la FAO (*Codex alimentarius*):

« (...) le produit non fermenté, mais fermentescible, obtenu en ajoutant de l'eau, avec ou sans adjonction de sucres tels que définis à la section 3-1-2 (a), de miel et/ou de sirops tels que décrits à la section 3-1-2 (b), et/ou d'édulcorants parmi ceux énumérés dans la *Norme générale pour les additifs alimentaires (NGAL)*, à des produits visés dans les sections 2-1-1, 2-1-2, 2-1-3, 2-1-4 et 2-1-5 ou à un mélange de ces produits »

En somme, un nectar de fruits comprend nécessairement un pourcentage (variable selon les espèces de fruits : c'est ce que précise l'annexe IV de la directive de 2001) de produits tirés de fruits (jus de fruits, jus de fruits concentré, jus de fruits déshydraté, respectivement, selon les points 1,2,3 de la directive), plus ou moins d'eau, et en général du sucre ou un substitut de sucre (miel ou édulcorant) jusqu'à 20% (selon l'annexe I de la directive). En outre, colorants et conservateurs sont interdits mais « acidifiants et anti-oxygène sont autorisés » (campagne publicitaire des Centres Édouard Leclerc en 1998, voir de Vecchi 2002: 90, n.1). Voilà la somme des caractères que devrait contenir une définition terminologique cumulative disant toute la vérité sur l'offre du marché mondial en la matière.

3.3. Le réseau conceptuel de « nectar de fruits »

3.3.1. Les relations paradigmatiques

L'intelligence artificielle prend en compte depuis longtemps les liens suivants :

– générique/spécifique (en intension)

- (4) « nectar de fruits » / G(énérique) : « boisson non alcoolisée » (D00ETI, 9-5) / G : « boisson » (D00ETI, 3-1-10) / G : « produit alimentaire » (D01JUS, c 1) = « denrée alimentaire » (D00ETI, 6-6) = « aliment » (D00ETI, 6-5) / G : « produit » (D01JUS, A IV-a)

– comprenant/compris (en extension)

- (5) « nectar de fruits » / C(omprenan)T : « ingrédient » (D01JUS, c 7)

– exemple (en extension)

- (6) « nectar de fruits » / E(xemple) : « nectar d'abricots »

3.3.2. Les preuves horizontales

– les définitions

Les définisseurs génériques peuvent être de niveau hiérarchique variable, comme le montre l'exemple (4), et certains ne sont que des introducteurs passe-partout (à commencer par *sorte de*). Voici un cas de trop grande généralité (pour un traitement systématique sous forme de réseau conceptuel) :

- (7) « édulcorant » / D(éfinition) : « Substance qui ... » (P06ADD, A-I-I)

C'est le spécifique « substance chimique » qui serait précis⁴.

– les formulations de liens paradigmatiques

Il existe beaucoup de façons d'exprimer la hiérarchie « générique/spécifique ». Pour ne prendre qu'un exemple, « c'est le cas de l'expression *entre autres* » (Baneyx et al. 2005: 3.3).

4 En terminologie, la bonne définition est celle qui recourt au définisseur générique immédiatement supérieur, « *nächsthöherer Oberbegriff* » (Wüster 1985: 30). Par exemple, dans les textes concernant le machinisme agricole, un hyponyme comme *outil* ou *machine* est préféré à l'hyperonyme *instrument* (voir Messaoudi 2000: 426).

Il en va de même pour la relation « comprenant / compris » :

- (8) fr. *produits ne comportant qu'un seul ingrédient* = en. *products comprising a single ingredient* = it. *prodotti costituiti da un solo ingrediente* (D00ETI, 6-2-c)
- (9) fr. *ingrédients composants d'une autre denrée alimentaire* = it. *ingredienti che rientrano nella composizione di un altro prodotto alimentare* (D00ETI, 6-6)
- (10) fr. *ce terme recouvre les boissons (...) il inclut l'eau* = it. *sono comprese le bevande (...) esso include l'acqua* (R2SEC, 2)

Des mots tels que *composant*, au sens chimique, relèvent de la terminologie du domaine, l'agro-alimentaire, mais *comporter* et *inclure* sont d'usage plus général, ce qui ne les empêche pas d'exprimer des liens entre des termes. Leur choix semble toutefois influencé par des habitudes d'écriture ; ainsi, si « *est trouvé dans* est le converse de *contient* » (Harris 2007: 56), c'est que le grand linguiste américain en juge d'après des rapports de laboratoires d'immunologie.

4. Le terme *nectar*

Un terme est le nom donné dans une langue à une entité conceptualisée par une communauté de travail. Cette dénomination est souvent un nom ou un groupe nominal, mais elle peut aussi appartenir à une nomenclature alphanumérique :

- (11) *CO2 = E290* = fr. *dioxyde de carbone* (D01JUS, A I-II-1, et aussi IU-PAC) = fr. *anhydride carbonique* (D93JUS, 10-4-b)

C'est « une unité définie dans les textes de spécialité » (Kocourek 1991: 180). Il synthétise souvent une information, en concurrence avec une périphrase descriptive :

- (12) en. *sales name* (D00ETI, 5-1-b) = en. *name under which the food additive is sold* (P06ADD, 24-1-a) = fr. *dénomination de vente* = de. *Verkehrszuweisung*

Tout comme les ontologies de domaines, les terminologies, au sens d'« ensembles cohérents de termes », reposent sur une « conceptualisation des objets du monde que partage une communauté de pratiques » (Depecker/Roche 2007: 112). Comme les communautés de pratiques

sont variées, une même conceptualisation n'exclut pas la concurrence de synonymes selon la géographie et selon les milieux de travail :

(13) fr. (de France) *saccharose* (UE) = fr. (canadien) *sucrose* (CODEXJUS)
= *C12H22O11*

(14) fr. (de France) *jus de fruits gazéifié* (D93JUS, 10-4-b) = fr. (anglicisé)
jus de fruits carbonaté (CODEXJUS, 8-1-2-9)

De tels exemples donnent du crédit à la thèse selon laquelle le terme est une « variante fonctionnelle du nom commun » (Sager 2000: 53). Il s'agit en effet d'expressions référentielles sujettes à l'anaphore générique, comme les noms propres et les descriptions définies, et la désignation d'un objet technique est toujours un acte de langage d'un énonciateur, même si l'existence de caractères distinctifs garantit la spécificité du référent. Par exemple, un amortisseur arrière reste un amortisseur arrière pour tout le personnel de la même usine italienne, qu'il soit appelé techniquement *ammortizzatore posteriore* ou *monoammortizzatore*, dialectalement *amortisator de drio* ou, familièrement, *mono* (voir Bertaccini/Matteucci 2006: 324). Mais si une partie des expressions prédicatives peuvent être des termes, il n'y a pas de raisons pour exclure des terminologies les verbes et les adjectifs spécialisés, qui introduisent un autre acte de langage majeur : la prédication.

5. Les mots de relation

Harris (1971: 229) appelle « noms de relation » des mots tels que *frère* : un frère est nécessairement le frère de quelqu'un, ce qui s'exprime classiquement en logique par la notation *frère* (x,y). Ce n'est pas une expression référentielle, mais un opérateur nominal qui a besoin d'être saturé pour que l'information transmise soit complète ; les noms de relation sont donc dignes d'intérêt quand on a en vue les relations « transversales » ou « horizontales » dans le discours. Ainsi, *administration* peut être un nom d'entité référentielle s'il s'agit de désigner des bureaux, mais c'est un nom d'action dans « l'administration de l'hôpital par le directeur ». Dans un corpus de réanimation chirurgicale, c'est un autre scénario de réalité qui sera exprimé à l'aide du même mot, par exemple dans « l'administration de morphine par l'infirmier ». Ce n'est donc pas un simple critère numérique (le « nombre de voisins en tête et en expansion », selon Baneyx et al. 2005: 3.2.2) qui fait la pertinence

ce terminologique du mot *administration* dans le travail de Le Moigno (voir Le Moigno et al. 2002: 4.2 et/ou Bourigault et al. 2004: 105), mais bien le typage conceptuel des compléments.

5.1. Noms de relation à tout faire

Beaucoup de noms, dans les textes spécialisés, n'ont pas de spécificité par eux-mêmes mais fonctionnent comme des dénominations de concepts généraux. C'est le cas d'*utilisation* dans l'exemple ci-dessous, et pourtant ce nom peut contribuer indirectement à un réseau conceptuel en mettant en relation les termes *additif* et *denrée alimentaire*.

- (15) « utilisation d'*additifs* dans des *denrées alimentaires* » (P06ADD, c 13)

5.2. Verbes de relation à tout faire

Si l'on observe les verbes introducteurs du terme médical *lésion* (bien différents de ceux de *lésion* en droit) dans le corpus de Le Moigno, on est frappé par leur absence de spécificité conceptuelle : « *avoir; confirmer; éliminer; évoquer; exister; montrer; noter; présenter; retrouver; révéler* » (Le Moigno et al. 2002: 5.1). Cet usage particulier de verbes courants a été observé dans « les langues de la science » par Harris : « *apparaît dans* et *est trouvé dans* peuvent tous deux être considérés comme des opérateurs sur la paire d'arguments *molécule, cellule*, comme dans *un anticorps apparaît dans les cellules du plasma* et *un anticorps est trouvé dans les cellules du plasma* » (2007: 50).

On peut également considérer comme des verbes de relation des mots comme *mettre* ou *procéder à*, qui sont moins des mots « vides » que des mots polyvalents. C'est le cas de *mettre, bringen, to place, immittere* et *colocar* dans l'exemple suivant :

- (16) fr. *les additifs alimentaires (...) peuvent être mis sur le marché* = de. *dürfen in Verkehr gebracht werden* = en. *may be placed on the market* = es. *podrán comercializarse* = it. *possono essere immessi sul mercato* = pt. *poden ser colocados no mercado* (P06ADD, 4-1)

On le voit, les locutions verbales sont strictement équivalentes, conceptuellement, au verbe simple espagnol *comercializar*. En outre les verbes concernés sont susceptibles de nominalisations comme beaucoup d'autres :

- (17) fr. *mise sur le marché* = de. *Inverkehrbringen* = en. *placing on the market* = es. *comercialización* = it. *immissione sul mercato* = pt. *colocação no mercado* (P06ADD, ch. III)

5.3. Adjectifs, participes et prépositions en concurrence

- (18) fr. *directive relative aux jus de fruits* = de. *Richtlinie über Fruchtsäfte* = en. *directive relating to fruit juices* = it. *direttiva concernente i succhi di frutta* = pt. *directiva relativa aos sumos de frutos* (D01JUS, titre)

Dans une même langue, on peut observer tour à tour une locution prépositive, un adjectif et une préposition simple pour exprimer la même relation, comme on le voit dans les exemples suivants :

- (19) fr. *législation en matière d'additifs alimentaires* = *législation relative aux additifs alimentaires* = *législation sur les additifs alimentaires* (P06ADD, exposé des motifs)
- (20) it. *allo scopo di conferire loro un gusto zuccherato* (D01JUS, 3-4) = it. *per dolcificare il prodotto* (D01JUS, A 1) = fr. *à des fins d'édulcoration* (bis)

5.4. Les chaînes de relations

Ce qui est souvent marginalisé sous le nom de *phraséologie* et parfois dénoncé comme langue de bois est exploitable en tant que marqueur de relations :

- (21) fr. *exigences en matière d'information concernant l'identification des additifs alimentaires* (P06ADD, 20)

6. Les termes de relation

Il n'est pas rare, dans les textes spécialisés, que les mots de relation dénotent des concepts terminologiques ; on peut alors parler de termes de relation. Harris a été l'un des premiers à observer que, « dans les langues de la science », « il existe des séries spécifiques d'arguments qui apparaissent sous des séries particulières d'opérateurs » (2007: 49). Voici son exemple :

- (22) *Les polypeptides ont été dissous dans de l'HCl* (Harris 2007: 49)

Dans cet emploi, *dissoudre* est aussi chimique que *polypeptides* ou *HCl*, et c'est en l'occurrence un terme insaturé qui en relie deux saturés. Il peut donc être pertinent de partir d'opérateurs verbaux pour construire un réseau conceptuel (voir de Vecchi/Estachy 2008). C'est aussi ce que propose la méthodologie des schémas d'arguments à base de classes d'objets, dont les exemples les plus convaincants, comme celui-ci, sont tirés, en fait, des langues spécialisées :

(23) « N°<humain, automate> REMPLIR N1<formulaire> » (Le Pesant / Mathieu-Colas 1998: 16)

6.1. Terme et discours

Le repérage des opérateurs ne peut être que manuel⁵. Une liste de concordances de termes référentiels fondée sur la connaissance préalable du domaine est utile pour tout ce qui est explicité par des moyens lexicaux (et des éléments de nomenclature alphanumérique le cas échéant), mais il lui manquera toujours le non dit (voir les ellipses de l'exemple 24), auquel il importe de faire une place⁶, et le « dit grammaticalement » (voir les anaphores de l'exemple 25). Examinons un sous-titre hors contexte :

(24) « Obligation d'information » (P06, 26)

Ce titre d'article est triplement implicite : information par qui ? pour qui ? sur quoi ? Seul le recours au contexte fournit les références de cette séquence elliptique. Quant à l'anaphore, elle peut prendre des formes variées :

(25) fr. *publicité faite* à leur égard = de. *Werbung* hierfür = en. *advertising of* foodstuffs = es. *publicidad* de los productos alimenticios = it. *la relativa pubblicità* (D00ETI, titre)

6.2. Substantifs déverbaux

La nominalisation est, dans les textes spécialisés, la façon la plus fréquente d'exprimer les opérations, les états, les processus et les proprié-

⁵ « Il n'existe pas d'outils dédiés spécifiquement à l'acquisition de ces relations » (Grabar/Hamon 2004: 63).

⁶ Voir Costa/Silva 2008.

tés. Cette particularité a été observée par divers auteurs, notamment Halliday et al. (1966), Kocourek (1991: 87-89) et Condamines (2003) ; elle a été vérifiée sur divers corpus en français (Lerat 2008b), en espagnol (Lerat 2008a) et dans plusieurs autres langues de l'UE (Lerat 2007). Ainsi, *obligation* et *information* sont des nominalisations dans l'exemple 24. Aussi bien, quand un spécialiste de la recherche assistée d'information médicale évoque « des relations (...) comme *manifestation_de*, non hiérarchique mais utile en particulier pour la recherche d'information » (Zweigenbaum 2004: 117), c'est d'une nominalisation qu'il s'agit, et d'un opérateur insaturé dont le sujet attendu est un nom de symptôme et le complément attendu un nom de pathologie.

Les nominalisations sont les meilleurs candidats pour désigner canoniquement les concepts : d'une part leur fréquence dans les textes spécialisés est un bon argument statistiquement, d'autre part il est techniquement (lexicographiquement) plus facile de définir un nom que le verbe ou l'adjectif correspondant.

6.3. Substantifs non déverbaux

Les noms de propriétés à base d'adjectifs sont plus ou moins nombreux selon les domaines. Voici un exemple pour l'agro-alimentaire :

- (26) fr. *la traçabilité des denrées alimentaires* = de. *die Rückverfolgbarkeit bei Lebensmitteln* = en. *the traceability within food and feed business* = it. *la rintracciabilità dei prodotti* = pt. *a rastreabilidade nas empresas do sector alimentar* (R02SEC, c 28)

L'histoire des langues fait que les opérateurs nominaux, même quand ils sont morphologiquement parents de verbes ou d'adjectifs, n'ont pas toujours une origine transformationnelle indiscutable ni un suffixe repérable automatiquement, comme le montre l'exemple suivant :

- (27) fr. *teneur en sucre inverti* = de. *Gehalt* = en. *content* = es. *contenido* = it. *tenore* = nl. *gehalte* = pt. *teor* (D93JUS, 1-4-b-bb)

6.4. Verbes

Le choix entre formulation verbale et formulation nominale varie avec les langues, les rédacteurs et ce qui vient d'être énoncé :

- (28) fr. *l'addition de vitamine* = de. *Vitamine zusetzen* (D01JUS, c 8)

Dans le corpus, le verbe l'emporte sur le nom pour dénommer cette opération ; certes, on trouve les trois synonymes *addition* (exemple 28), *adjonction* (D93JUS, c 2) et *ajout* (CODEXETI, 4-2-1-5), mais surtout *additionner*, *adjoindre* et *ajouter*. Au demeurant, la grande majorité des opérations concernées est exprimée par des nominalisations.

6.5. Adjectifs

Les adjectifs propres aux textes spécialisés (y compris aux étiquettes apposées sur les produits) sont des termes. En particulier, les participes passés désignant des états fournissent des indications sur la fabrication. Voici quelques exemples :

- (29) fr. *déshydraté* (D01JUS, A I-I-3), *irradié* (D00ETI, 5-3), *lyophilisé* (D00ETI, 5-3), *reconstitué* (CODEXJUS, 8-1-2-2), *sucré* (D00ETI, 3-4)

Ces propriétés pertinentes font aussi l'objet de formulations concurrentes :

- (30) *déshydraté* a officiellement pour synonyme *en poudre* (D01JUS, A-I-I-3), *non sucré* se dit aussi *sans sucre ajouté* (P06ADD, 3-2-e) et *non additionné de sucre* (D93JUS, 3-2-b), *irradié* = *traité par ionisation* (D00ETI, 5-3)

7. Complétude et relation combinatoire

7.1. La pensée et la langue

L'idée de prendre en compte les relations « transversales » (Grabar/Hammon 2004: 63) dans les textes spécialisés respecte la discursivité en tant que telle, mais quels cooccurrents sélectionner ? On peut envisager de s'attacher à la « collocation conceptuelle », au sens de « relation typique entre un prédicat et un objet auquel il est applicable » (Lerat 1995: 104), mais l'inconvénient d'un tel projet est épistémologique : c'est mêler du logique (« un prédicat ») et du grammatical (« un objet »).

Le risque de confondre scénario de réalité (ontologique) et schéma actanciel (grammatical) est vieux comme la confusion entre la pensée et la langue, et la linguistique mathématique n'y a pas échappé, pas plus que les WordNets. En parlant de « prédicats sémantiques » (M. Gross 1981), on peut faire un excellent travail descriptif, comme celui du la-

boratoire de M. Gross, mais dans les limites d'une langue donnée. De même, en utilisant « les résultats de l'analyse distributionnelle pour dégager les grandes classes de concepts » (Bourigault et al. 2004: 105), on postule la transparence conceptuelle des mots, et il suffit que certains soient fréquents pour qu'ils deviennent des « candidats termes ».

Il y a chez Harris lui-même un flou terminologique que masque en partie l'usage du générique *opérateur*, qui vaut à la fois pour *prédicat* (logique) et pour *expression prédicative*. Tantôt il s'agit du premier emploi : « l'opérateur est un constructeur d'assertions, approximativement un prédicat par rapport à ses arguments » (1976: 26). Tantôt l'approche serait strictement grammaticale, n'était l'usage d'*argument* : « opérateur et arguments sont des mots, un verbe comme *manger* est un opérateur à deux arguments » (1976: 7).

Il importe, pour éviter toute confusion, de ne pas considérer comme synonymes *concept* (le concept relève de la cognition et, par la définition et les inférences possibles, de la logique) et *signifié* lexical (voir notamment la mise au point de Depecker 2000), c'est à dire le sens d'un mot en tant que tel (avec sa polysémie, ses connotations, sa motivation morphologique etc.). L'emprunt terminologique comble un vide lexical d'une langue quand les spécialistes s'efforcent d'éviter la dénomination en langue étrangère d'une innovation faite à l'étranger : ils connaissent le concept, parce qu'ils sont du métier, ils connaissent le terme étranger, parce qu'ils lisent et voyagent, mais ils n'ont pas encore « les mots pour le dire » dans leur langue. Bien plus, il n'est pas nécessaire qu'un concept émergent soit nommé dans une langue naturelle ; ainsi, au moment de la création de *rayons x*, il suffisait d'être assez bon physicien pour savoir ce qu'étaient les « rayons appelés *x* » (Harris 1976: 224), et c'est durablement le cas. Pour résumer, « une structure conceptuelle peut être conçue indépendamment des formes linguistiques, qui se limitent à la porter à l'expression (...) une structure sémantique, par contre, est le contenu spécifique d'une forme linguistique, indissociable de celle-ci » (Gross/Prandi 2004: 119, n. 3).

Le concept terminologique, faisceau de « caractères » permettant des inférences telles que « si « nectar de fruit », alors « denrée alimentaire » », relève de l'ontologie de domaine, et les phrases dans lesquelles ses dénominations trouvent place sont des propositions logiques en même temps que des énoncés linguistiques. Dans le domaine de l'agro-

alimentaire, la terminologie vitivinicole illustre l'importance prioritaire des concepts par rapport aux dénominations, au point que les réalités rendues spécifiques par l'histoire et la géographie y sont dénommées dans toutes les langues par emprunt à la langue d'origine. *Retzina* est définitivement un emprunt au grec, parce que c'est une réalité grecque. Il en va de même pour *Qualitätswein mit Prädikat* dans toutes les versions de R08VIN, parce que c'est une conceptualisation allemande, et aussi pour l'espagnol *crianza*, quand il dénomme un vin rouge de la Rioja opposable, dans un paradigme commercial espagnol (et non pas de l'espagnol), à *reserva* et *gran reserva* (voir Lerat 2008a).

Le même mot, *crianza*, largement polysémique en espagnol, fournit un autre terme dans le même domaine, parce qu'il dénomme aussi une opération : l'élaboration des vins de qualité destinés à vieillir. Cette fois, il s'agit d'un concept partagé par l'Europe de la vitiviniculture, et la langue de chacun des principaux pays producteurs fournit un terme équivalent, comme le montre l'exemple suivant :

- (31) es. *crianza del producto* = de. *Ausbau* = el. τελεοποίηση = fr. *élevage*
 = it. *affinamento* (préféré en 2008 à *invecchiamento*, que le Parlement européen proposait en 2007) = pt. *apuramento* (R08VIN, 27-1-2)

L'exemple 31 illustre assez clairement la différence entre concept technique et signifié lexical : chaque terme est un mot (*affinamento* appartient à une « famille de mots », *invecchiamento* à une autre), mais il dénomme le concept partagé dans une « communauté de pratiques » professionnelle (travail sur le vin en barrique) en énonçant l'un de ses caractères (élaboration, préparation, finition, amélioration, vieillissement).

7.2. Insaturation conceptuelle et insaturation grammaticale

– les verbes

Une verbe transitif appelle par définition un sujet et un complément grammatical : c'est ce qui permet à Harris de dire que *manger* est un verbe à deux arguments. La différence entre *manger* et *ajouter*, au sens chimique, est que le premier verbe accepte une liste difficilement limitable culturellement (le mangeable est relatif, tout voyageur le sait) alors que le second, dans le corpus, n'accepte que les noms d'additifs autorisés.

– les noms

Les nominalisations d'adjectifs sont plus fréquentes ici que les adjectifs correspondants, qui parfois sont hors d'usage (comme *traçable*) ou de sens différent (comme *intègre*, qui n'a de lien qu'étymologique avec *intégrité* dans cet exemple) :

(32) *la sécurité ou l'intégrité des denrées alimentaires* (R02SEC, 19-2)

– les adjectifs

Au sein d'une unité terminologique polylexicale, un adjectif de relation tient lieu de complément de nom et devient lui-même constituant de terme. Il s'agit alors de détermination, non de prédication :

(33) fr. *sécurité alimentaire* (R02SEC, 20-2) = fr. *sécurité des aliments* (R02SEC, titre) = de. *Lebensmittelsicherheit* (bis) = en. *food safety* (bis) = it. *sicurezza alimentare* (bis)

– les prépositions

Pour Harris, les prépositions qui introduisent des objets indirects sont des « indicateurs d'arguments » (1976: 26). Les locutions prépositives diffèrent des prépositions simples en ceci qu'elles s'organisent, au moins génétiquement, autour d'un « nom prédicatif » (Gross/Prandi 2004: 143). Pour la terminologie, cette différence n'est pas essentielle puisque globalement les unes et les autres jouent le rôle d'opérateurs du fait qu'elles mettent en relation des expressions référentielles qui peuvent être des termes :

(34) fr. *jus de fruits obtenu à partir d'un concentré* = de. *Fruchtsaft aus Fruchtsaftkonzentrat gewonnener* = en. *fruit juice from concentrate* = es. *zumo a base de concentrado* = it. *succo di frutta ottenuto da un succo concentrato* (D01JUS, c 7)

8. Relations combinatoires et héritage des propriétés

8.1. Héritage des propriétés et vérité

La validité de l'héritage des propriétés dans les réseaux conceptuels repose entièrement sur la constance des caractères définitoires des concepts. Il s'agit de propriétés reconnues à des objets référentiels

concrets ou abstraits, et non pas de traits sémantiques constitutifs de signifiés de mots⁷. L'exemple de « concentré » est éclairant :

(35) jus de fruits concentré / solution concentrée à 10% / vin concentré

Dans les 3 cas, le mot *concentré* est un terme différent.

- 1) dans *jus de fruits concentré*, *concentré* veut dire « obtenu (...) par élimination physique d'une partie déterminée de l'eau de constitution. Lorsque le produit est destiné à la consommation directe, cette élimination est d'au moins 50% » (D01JUS, A I-I-2).
- 2) *solution concentrée à 10%* désigne la quantité d'une substance chimique contenue dans l'unité de volume (d'après *IATE* et *ETB*).
- 3) Un vin est dit *concentré* s'il a « une teneur élevée en extrait sec (selon l'âge de la vigne, le rendement, le millésime) » (Lerat 2008a).

8.2. L'héritage des propriétés combinatoires

Puisque tout nectar de fruits est, par définition, un produit alimentaire, tout ce qui est vrai des produits alimentaires en général l'est aussi nécessairement de ce concept subordonné. La multiplicité des relations horizontales entretenues entre l'entité conceptualisée « produit alimentaire » et les opérations auxquelles elle est sujette enrichit grandement le réseau conceptuel de « nectar de fruits » :

(36) « nectar de fruits » / C(ombinatoire) C(onceptuelle) : « adjonction », « commerce », « commercialisation », « conditionnement », « consommation », « distribution », « étiquetage », « exportation », « fabrication », « importation », « présentation », « production », « stockage », « traitement », « transformation », « transport », « vente » etc.

On obtient un réseau terminologique en remplaçant les guillemets par l'italique et en prenant en compte les variables grammaticales et les synonymes :

(37) *nectar de fruits* / C(ombinatoire) T(erminologique) : *addition, additionner, adjoindre, adjonction, ajout, ajouter, commerce, commercialisation, commercialiser, conditionnement, conditionner, consommation, consommer, distribuer, distribution, emballage, emballer, entreposage,*

⁷ Voir notamment la différence entre « élément de pensée » (*Denkelement*) et « signification lexicale » (*wörtliche Bedeutung*) chez Wüster (1985, 7 et 34).

entreposer, étiquetage, étiqueter, exportation, exporter, fabrication, fabriquer, importation, importer, mettre sur le marché, mise sur le marché, présentation, stockage, stocker, traitement, transformation, transformer, vendre, vente etc.

En tant que mots, plusieurs de ces formes sont polysémiques (*conditionnement, conditionner, traitement*, à tout le moins), d'où la tendance de la terminologie à suivre le modèle des « ontologies de tâches » (un fichier par application), ce qui limite gravement les possibilités de réemplois. Pour que leur valeur de terme soit repérable, il existe plusieurs critères :

- 1) le domaine. Malheureusement, « la définition d'un « domaine » est arbitraire et relative aux objectifs définis dans un projet terminologique » (L'Homme *et al.* 2003 : 152).
- 2) une traduction validée est un indicateur appréciable, notamment sur les sites destinés à la clientèle et les sites communautaires, du fait des enjeux économiques, mais un texte est par définition unilingue.
- 3) la combinatoire des termes apparaît donc comme le critère le plus sûr, à condition que l'on ait vérifié qu'il s'agit bien d'un concept et non d'un autre, et que la dénomination retenue comme terme soit effectivement gagée sur la pratique textuelle d'une communauté de spécialistes. Par exemple, en matière d'agro-alimentaire, *conditionnement* n'est pertinent que quand il admet les équivalents suivants :

(38) fr. *conditionnement* = de. *Verpackung* = en. *packaging* = es. *envasado* = it. *imballaggio* = pt. *embalagem* (P006ADD, 3-2-a)

9. Conclusions

- 1) La terminologie est nécessaire pour la communication professionnelle interne et externe, la traduction spécialisée, la recherche d'information, la veille industrielle et l'élaboration de réseaux conceptuels destinés à la gestion des connaissances, qui nécessitent une ontologie « partageable, réutilisable et calculable » (Roche 2007: 7).
- 2) Les collocations sont précieuses parce qu'elles signalent des relations combinatoires, mais elles demandent à être interprétées pour

plusieurs raisons très fortes (pertinence inégale des fréquences⁸, polysémie, anaphores, connaissances présupposées).

- 3) La terminologie a besoin d'être multilingue pour ne pas se réduire à une lexicologie spécialisée. Seuls les concepts spécialisés sont de son ressort, et la seule façon d'être sûr que l'on parle bien de la même chose est le recours à une définition comportant les mêmes caractères d'une langue à l'autre.
- 4) La diversité des dénominations au sein d'une langue et d'une langue à l'autre (voir notamment l'exemple 31) est un bon révélateur de celle des caractères à prendre en compte (voir Thoiron 1994) dans beaucoup de cas, mais dans les règlements communautaires le choix de la langue source (anglais ou français) influence visiblement les langues de traduction.
- 5) Au total, les relations combinatoires entre expressions insaturées et expressions saturées ne dispensent nullement d'un travail d'analyse des concepts, bien au contraire, mais leur utilité ne semble pas négligeable : elles valident syntagmatiquement des relations paradigmatiques entre termes, surtout elles signalent dans la chaîne du discours des relations non hiérarchiques mais fonctionnelles, et enfin la combinatoire conceptuelle est héritable.

Bibliographie

- Baneyx, Audrey/Malaisé, Véronique/Charlet, Jean/Zweigenbaum, Pierre/Bachimont, Bruno 005 : Synergie entre analyse distributionnelle et patrons lexico-syntaxiques pour la construction d'ontologies différentielles. In *TIA-2005*, 31-42 [en ligne] <http://www.test.biomath.jussieu.fr>.
- Bertaccini, Franco/Matteucci, Alessandra 2006 : La terminologie d'entreprise et ses contextes d'usage. In D. Blampain/Ph. Thoiron/M. Van Campenhoudt (eds), *Mots, termes et contextes*. Paris : AUF, 317-326.
- Bourigault, Didier/Aussenac-Gilles, Nathalie/Charlet, Jean 2004 : Construction de ressources terminologiques ou ontologiques à partir de textes. Un cadre unificateur pour trois études de cas. In *Revue d'intelligence artificielle* 18-2, 87-110.
- Cabré, Teresa 2003 : Theories of Terminology. Their Description, Prescription and Explanation. In *Terminology* 9/2, 163-199.

8 « Le candidat terme *rapport de vraisemblance* a la plus forte fréquence d'apparition (177) dans le corpus. Or il est sémantiquement pauvre pour le domaine de la pneumologie, il n'est donc pas caractéristique » (Baneyx et al. 2005: 3.2.1).

- Condamines, Anne 2003 : *Sémantique et corpus spécialisés : constitution de bases de connaissances terminologiques*. Habilitation à diriger des recherches, Université de Toulouse 2 [en ligne] <http://w3.univ-tlse2.fr/erss/textes/pagespersos/acondami/HDR-2.pdf>.
- Costa, Rute/Silva, Raquel 2008 : De la typologie à l'ontologie de textes. À paraître dans les actes du colloque TOTh 2008 (Annecy : Institut Porphyre).
- Depecker, Loïc 2000 : Le signe entre signifié et concept. In H. Béjoint/Ph. Thoiron (eds.), *Le sens en terminologie*. Lyon : Presses de l'Université de Lyon, 86-126.
- Depecker, Loïc/Roche, Christophe 2007 : Entre idée et concept : vers l'ontologie. In *Langages* 168, 106-114.
- ETB : EuroTermBank Consortium [en ligne] <http://www.eurotermbank.com>.
- Grabar, Natalia/Hamon, Thierry 2004 : Les relations dans les terminologies structurées : de la théorie à la pratique. In *Revue d'intelligence artificielle* 18-1, 57-85.
- Gross, Gaston/Prandi, Michele 2004 : *La finalité. Fondements conceptuels et genèse linguistique*. Bruxelles-Louvain La Neuve : De Boeck-Duculot.
- Gross, Maurice 1981 : Formes syntaxiques et prédicats sémantiques. In *Langages* 63, 7-52.
- Guillet, Alain 1986 : Représentation des distributions dans un lexique-grammaire. In *Langue française* 69, 85-107.
- Halliday, Mikael A. K./Mac Intosh, Angus/Strevens, Peter D. 1966 : *Linguistic Sciences and Language Teaching*. London : Longman.
- Harris, Zellig S. 1971 : *Structures mathématiques du langage* (1968), trad. C. Fuchs. Paris : Dunod.
- Harris, Zellig S. 1976 : *Notes du cours de syntaxe*, trad. M. Gross. Paris : Seuil.
- Harris, Zellig S. 2007 : *Langue et information* (1988), trad. Amr Helmy Ibrahim. Paris : Cellule de recherche en linguistique.
- IATE : *Inter Active Terminology in Europe* [en ligne] <http://iate.eu/interdiff>.
- Kleiber, Georges 1981 : *Problèmes de référence : descriptions définies et noms propres*. Paris : Klincksieck.
- Kocourek, Rostislav 1991 : *La langue française de la technique et de la science* (1982), 2^{ème} éd.. Wiesbaden : Oscar Brandstetter.
- L'Homme, Marie-Claude 2004 : *La terminologie : principes et techniques*. Montréal : Presses de l'Université de Montréal.
- L'Homme, Marie-Claude/Heid, Ulrich/Sager, Juan C. 2003 : Terminology during the past decade (1994-2004) : An Editorial Statement, *Terminology* 9-2, 151-161.
- Le Moigno, Sylvie/Charlet, Jean/Bourigault, Didier/Jaubert, Marie-Christine 2002 : Construction d'une ontologie à partir de corpus : expérimentation et validation dans le domaine de la réanimation chirurgicale, *ICO2*, 229-238.
- Le Pesant, Denis/Mathieu-Colas, Michel 1998 : Introduction aux classes d'objets. In *Langages* 131, 6-33.

- Lerat, Pierre 1995 : *Les langues spécialisées*. Paris : Presses Universitaires de France.
- Lerat, Pierre 2007 : Les nominalisations en *-tion* dans un texte techno-administratif. In *Terminologie et Ontologie. Théories et applications*. Annecy, Institut Porphyre, 79-92.
- Lerat, Pierre 2008a : La terminologie communautaire des opérations vitivinicoles. À paraître dans les actes du *II Congreso sobre la lengua de la vid y el vino y su traducción* (Soria, avril 2008).
- Lerat, Pierre 2008b : Propositions pour un réseau conceptuel des instruments de mesure œnologiques. À paraître dans les actes du colloque TOTh 2008 (Annecy, Institut Porphyre).
- Leroyer, Patrick/Simonsen Henrik Køhler 2006 : Contexte d'utilisation, contexte de communication – La double identité du dictionnaire d'entreprise. In D. Blampain/Ph. Thoiron/M. Van Campenhoudt (eds.), *Mots, termes et contextes*. Paris : AUF, 283-298.
- Messaoudi, Leila 2000 : Opacité et transparence dans les technolèctes bilingues français-arabe. In *Meta* 45-3, 424-435.
- Mounin, Georges 1990 : *Problèmes théoriques de la traduction* (1963). Paris : Gallimard.
- Pavel, Silvia/Nolet, Diane 2001 : *Précis de terminologie / Handbook of Terminology*. Canada : Ministère des Travaux publics et Services gouvernementaux.
- Pearson, Jennifer 1998 : *Terms in Context*. Amsterdam/Philadelphie : John Benjamins.
- Roche, Christophe 2007 : Le terme et le concept : fondements d'une ontologie. In *Terminologie et Ontologie : Théories et applications*. Annecy : Institut Porphyre, 1-22.
- Sager, Juan C. 1990 : *A Practical Course in Terminology Processing*. Amsterdam : John Benjamins.
- Sager, Juan C. 2000 : Pour une approche fonctionnelle de la terminologie. In H. Béjoint/Ph. Thoiron (eds.), *Le sens en terminologie*. Lyon : Presses Universitaires de Lyon, 40-60.
- Slodzian, Monique 2000 : L'émergence d'une terminologie textuelle et le retour du sens. In H. Béjoint/Ph. Thoiron (eds.), *Le sens en terminologie*. Lyon : Presses Universitaires de Lyon, 61-85.
- Thoiron, Philippe 1994 : La terminologie multilingue : une aide à la maîtrise des concepts. In *Meta* 39-4, 765-773.
- de Vecchi, Dardo 2002 : *Vous avez dit jargon* Paris : Eyrolles.
- de Vecchi, Dardo/Estachy, Laurent 2008 : Pragmaterminologie : les verbes et les actions dans les métiers. À paraître dans les actes du colloque TOTh 2008 (Annecy : Institut Porphyre).
- Wüster, Eugen 1968 : *Dictionnaire multilingue de la machine-outil. Volume de base Anglais-Français*. Londres : Technical Press.

Wüster, Eugen 1981 : L'Étude scientifique générale de la terminologie, zone frontalière entre la linguistique, la logique, l'ontologie, l'informatique et la science des choses. In G. Rondeau/H. Felber (eds.), *Fondements théoriques de la terminologie*. Québec : Université Laval, 57-114.

Wüster, Eugen 1985 : *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie* (1979). Copenhague : Handelshøjskolen.

Zweigenbaum, Pierre 2004 : L'UMLS entre langue et ontologie : une approche pragmatique dans le domaine médical. In *Revue d'intelligence artificielle* 18-1, 111-137.

Corpus

CODEXETI : Norme générale CODEX pour l'étiquetage des denrées alimentaires préemballées (CODEX STAN 1-1985 (rév. 1-1991)) [en ligne] <http://www.fao.org>

CODEXJUS : Norme générale CODEX pour les jus de fruits et les nectars de fruits (CODEX STAN 247-2005) [en ligne] <http://www.fao.org>

D00ETI : Directive 2000/13/CE du Parlement européen et du Conseil du 20 mars 2000 relative au rapprochement des législations des États membres concernant l'étiquetage et la présentation des denrées alimentaires ainsi que la publicité faite à leur égard.

D01JUS : Directive 2001/112/CE du Conseil du 20 décembre 2001 relative aux jus de fruits et à certains produits similaires destinés à l'alimentation humaine.

D93JUS : Directive 93/77/CEE du Conseil du 21 septembre 1993 relative aux jus de fruits et à certains produits similaires.

P06ADD : Proposition de Règlement du Parlement européen et du Conseil sur les additifs alimentaires (COM/2006/0428 final).

R02SEC : Règlement (CE) n° 178/2002 du Parlement européen et du Conseil du 28 janvier 2002 établissant les principes généraux et les prescriptions générales de la législation alimentaire, instituant l'Autorité européenne de sécurité des aliments et fixant des procédures relatives à la sécurité des denrées alimentaires.

R08VIN : Règlement (CE) n° 479/2008 du Conseil du 29 avril 2008 portant organisation commune du marché vitivinicole.