



You have downloaded a document from
RE-BUŚ
repository of the University of Silesia in Katowice

Title: Bases de donnees lexicales electroniques : une approche orientee objets.
Partie I: Questions de modularite

Author: Wiesław Banyś

Citation style: Banyś Wiesław. (2002). Bases de donnees lexicales
electroniques : une approche orientee objets. Partie I: Questions de modularite.
"Neophilologica" (T. 15, (2002), s. 7-29).



Uznanie autorstwa - Użycie niekomercyjne - Bez utworów zależnych Polska - Licencja
ta zezwala na rozpowszechnianie, przedstawianie i wykonywanie utworu jedynie w celach
niekomercyjnych oraz pod warunkiem zachowania go w oryginalnej postaci
(nie tworzenia utworów zależnych).



UNIwersYTET ŚLĄSKI
W KATOWICACH



Biblioteka
Uniwersytetu Śląskiego



Ministerstwo Nauki
i Szkolnictwa Wyższego

Wiesław Banyś
Université de Silésie
Katowice

Bases de données lexicales électroniques – une approche orientée objets

Partie I: Questions de modularité*

0. Que les informations lexico-sémantiques, à côté des informations phonétiques, morphologiques et syntaxiques, doivent être nécessairement prises en compte dans les recherches qui se veulent efficaces dans le cadre du traitement automatique de la langue et de l'intelligence artificielle est une chose, ces derniers temps au moins, admise par la majorité des chercheurs (cf. p.ex. Y. Wilks, B. M. Sator, L. M. Guthrie, 1996).

Plusieurs questions se posent à cet égard dont nous ne mentionnons ici rapidement que quelques-unes qui sont directement liées aux propositions que l'on va présenter ci-dessous.

La première concerne le besoin lui-même de la construction des dictionnaires électroniques. Même pas tellement le besoin lui-même, qui semble être actuellement une chose parfaitement acquise, aussi bien du point de vue des travaux effectués dans le cadre du traitement automatique du langage, ceux qu'on fait dans le cadre de l'intelligence artificielle, que du point de vue d'un utilisateur moyen averti qui voudrait avoir accès à un dictionnaire qui lui permettrait de transgresser l'ordre parfaitement alphabétique des dictionnaires traditionnels et trouver beaucoup plus rapidement (facteur temps) et efficacement (facteur qualité) le mot, l'expression ou la construction qu'il cherche – ces derniers temps, grâce à la conversion des dictionnaires traditionnels en leurs versions électroniques et grâce aux ordinateurs qui effectuent la computation des entrées presque instantanément, cela est devenu en fait possible. Pourtant, le problème est qu'il n'est pas uniquement question

* Ce travail a été effectué dans le cadre du Projet MATCHPAD (IST-12256) du V^e Programme de la Commission Européenne dont le but est de construire un programme de traduction automatique de l'anglais vers le polonais et du polonais vers le français.

de reprendre sous forme électronique les informations que les dictionnaires traditionnels contiennent explicitement telles quelles ; ce qui importe, c'est pouvoir réunir aussi bien les informations explicites des dictionnaires traditionnels, souvent les réorganiser, en tirer, là où c'est possible, les informations implicites (p.ex. sur les hiérarchies de dépendances sémantiques) et en ajouter éventuellement d'autres pour constituer une véritable base de données lexicales polyvalente et non seulement une version électronique de ce que l'on a déjà sous forme papier. Autant dire que l'extraordinaire travail des lexicographes a permis de réunir sans doute, si ce n'est pas toutes au moins leur grande majorité, les informations factuelles utiles concernant les sens des mots et leurs emplois, mais l'organisation des dictionnaires traditionnels autour des entrées alphabétisées et les limitations de place constantes, d'où la nécessité de diminuer autant que possible la redondance descriptive, ne permet pas d'obtenir, là où c'est possible, ni facilement ni rapidement, toutes les informations nécessaires.

En d'autres termes, le problème est donc peut-être davantage de type « structurel » que « factuel », lié plus à l'organisation et la présentation de l'information lexicale qu'à son contenu (cf. aussi p.ex. G. A. Miller, 1990 ; G. A. Miller, R. Beckwith, Chr. Fellbaum, D. Gross, K. Miller, R. Tengi, 1990 ; Chr. Fellbaum, 1998 ; T. Fontenelle, J. Vanandroye, 1989 ; T. Fontenelle, 1997a, 1997b ; Y. Wilks, B. M. Slator, L. M. Guthrie, 1996)(cf. p.ex. à cet égard les bases de données en gros de ce type, proposant une nouvelle organisation des informations lexicales, mais qui diffèrent fort l'une de l'autre, WordNet (cf. Chr. Fellbaum, 1998 ; T. Fontenelle, 1997a)). Puisque les idées et les projets d'exploitation automatique de telles bases de données apparaissent beaucoup plus vite que les idées et la réalisation de ces bases de données, il existe une sorte de « goulot d'étranglement » (*lexical bottleneck*) qui entrave le progrès rapide en la matière (cf. p.ex. Y. Wilks, B. M. Slator, L. M. Guthrie, 1996).

Ce constat de fait nous fait poser (ou plutôt reposer) la question, vu la rapidité et la dispersion à travers le monde avec lesquelles les travaux en question s'effectuent, du mode de description, aussi bien du point de vue contenu que format, qui pourrait satisfaire aux besoins d'une présentation des données extensible et réutilisable, ne constituant donc pas forcément un tout clos, mais tout au contraire ouvert à des modifications et enrichissements éventuels, nécessaires sans doute si l'on prend en compte le facteur temps vu p.ex., d'une part, la possibilité d'exploiter tout un monde d'informations implicites que renferment les dictionnaires traditionnels (cf. p.ex. Y. Wilks, B. M. Slator, L. M. Guthrie, 1996 ; T. Fontenelle, 1997 ; N. Calzolari, 1988 ; C. M. Guo, 1992) et, d'autre part, l'envol des recherches lexicales de corpus (cf. p. ex. T. McEnery, A. Wilson, 1996 ; R. Garside, G. Leech, A. McEnery, eds, 1997 ; V. B. Y. Ooi, 1998 ;

D. Biber, S. Conrad, R. Reppen, 1998 ; B. Habert, A. Nazarenko, A. Salem, 1997), sans que cela nuise à l'intégralité de la description.

Cela présuppose un type de description des informations lexico-sémantiques aussi souple et général qu'il puisse être réintégré éventuellement dans différents cadres théoriques. Comme on va le voir pourtant ci-dessous, ici comme ailleurs, les impératifs de l'extensibilité et de la réutilisabilité des descriptions imposent en même temps, à leur tour, un certain type – pas trop impératif, il est vrai – de prise de position théorique (on l'a vu aussi au cours des développements de la programmation informatique (cf. p.ex. G. Masini, A. Napoli, D. Colnet, D. Léonard, K. Tombre, 1989 ; B. Meyer, 1997 ; B. Stroustrup, 1989 ; B. Cox ; 1986) et de la construction des bases de données (cf. p.ex. C. Delobel, Chr. Lécuse, Ph. Richard, 1991 ; G. Lausen, G. Vossen, 1997) – celle d'une approche de la description lexico-sémantique qui est orientée objet.

Les propositions que l'on va formuler ci-dessous, aussi bien côté structure que contenu, sont, on dirait d'une manière naturelle, conformes aux quatre principes de la bonne méthode de Descartes (cf. aussi B. Meyer, 1997), que je me ferai le plaisir de citer :

Au lieu de ce grand nombre de préceptes dont la logique est composée, je crus que j'aurais assez des quatre suivants, pourvu que je prisse une ferme et constante résolution de ne manquer pas une seule fois à les observer.

Le premier était de ne recevoir jamais aucune chose pour vraie que je ne la connusse évidemment être telle : c'est-à-dire d'éviter soigneusement la précipitation et la prévention, et de ne comprendre rien de plus en mes jugements que ce qui se présenterait si clairement et si distinctement à mon esprit que je n'eusse aucune occasion de le mettre en doute.

Le second, de diviser chacune des difficultés que j'examinerais en autant de parcelles qu'il se pourrait, et qu'il serait requis pour les mieux résoudre.

Le troisième, de conduire par ordre mes pensées, en commençant par les objets les plus simples et les plus aisés à connaître, pour monter peu à peu comme par degrés jusques à la connaissance des plus composés : et supposant même de l'ordre entre ceux qui ne se précèdent point naturellement les uns les autres.

Et le dernier, de faire partout des dénombrements si entiers et des revues si générales que je fusse assuré de ne rien omettre.

René Descartes, *Discours de la méthode*

1. Extensibilité et réutilisabilité des données d'une base de données lexicales

Les bases de données lexicales ne sont en fait *électroniques* que s'il y a un *software* qui permet de manipuler les données qu'ils renferment. Cela fait que leur contenu, dans le sens du modèle de description choisi, renfermant telles et telles informations, et leur «peau» représentationnelle, un programme – logiciel, sont relativement étroitement liés, même s'il est vrai qu'en fait, tout contenu, vu le but qu'on assigne au programme, peut être programmé par tout type de langage de programmation.

En parlant de l'extensibilité, c'est-à-dire de la facilité avec laquelle le «produit» s'adapte à un changement de spécification, le concept pouvant se rapporter aussi bien au contenu qu'à la forme ('la peau'), nous avons à l'esprit aussi bien le contenu d'une base de données lexicales que la forme de la présentation du contenu.

Si sa complexité n'est pas très grande, la chose est relativement facile à régler et les informations qu'un dictionnaire de telle taille renferme peuvent être adaptées à un nouveau format et à une nouvelle utilisation. Le problème commence quand la quantité de données est de loin plus élevée et la nouvelle utilisation demanderait un format descriptif totalement différent de celui qui a été adopté dans le dictionnaire de départ. Dans ce cas, un changement de la représentation de données ou de nos exigences par rapport à ce qu'elles représentent entraînerait une refonte tant soit peu complète de la description de départ.

Pour pouvoir assurer ce type d'extensibilité, on devrait avoir recours à (cf. aussi p.ex. B. Meyer, 1997 ; M. Page-Jones, 1995 ; D. G. Fire-smith, E. M. Eyckholt, 1995 ; C. Baudoin, G. Hollowell, 1996) :

- *une architecture descriptive simple*, puisqu'il est plus facile, si besoin se présente, d'adapter une architecture simple qu'une architecture complexe,
- *une décentralisation et une autonomie descriptives*, puisqu'il est plus probable que les changements éventuels ne concerneront qu'un (petit nombre de) modul(s) et non pas le système entier.

La question de réutilisation, cette capacité du format descriptif des données contenues dans une base électronique de données lexicales à servir dans la construction d'autres applications, est une autre facette – sorte de reflet de miroir de la question de leur extensibilité – de la souplesse du format descriptif choisi. En effet, les données en question devraient avoir un format tel qu'elles puissent être exploitées en dehors d'un système descriptif et logiciel donné et servir de point de départ à d'autres développements éventuels.

Les exigences d'une architecture flexible se ramènent en fait à l'exigence de la construction, conformément, aussi bien du point de vue concep-

tuel que du point de vue applicatif, au second principe de Descartes, d'une architecture modulaire, qui serait constituée de composantes autonomes, reliées par une structure relativement simple.

2. Modularité de l'architecture d'une base de données lexicales

Le second principe de Descartes, celui qui nous ordonne de diviser chacune des difficultés que l'on examine *en autant de parcelles qu'il se pourrait, et qu'il serait requis pour les mieux résoudre*, a en effet des conséquences non seulement purement conceptuelles, mais aussi des conséquences applicatives qui mettent ces premières en pratique lors de l'élaboration d'une base de données lexicales.

En effet, il ne s'agit pas seulement de décomposer conceptuellement le problème complexe en autant *de parcelles qu'il se pourrait* pour le mieux résoudre, mais aussi de pouvoir représenter la description de la solution du problème en question sous forme d'une architecture applicative qui suit de près la façon dont le problème était analysé, ce qui assure la conformité du cheminement du réalisé avec le conceptuel.

Et puisque, en dernière instance, c'est la capacité d'une théorie – réalisée ici dans une application, de type computationnel p.ex. – à fournir des moyens pratiques à imiter et à transformer la réalité qui prouve son adéquation (nous proclamons ainsi la paix entre ceux qui, comme Francis Bacon, croient ne pouvoir comprendre et / ou triompher de la nature qu'en lui obéissant et ceux qui, comme Gaston Bachelard, croient ne pouvoir comprendre la nature qu'en lui résistant), la modularité convenable, que constituent des composantes autonomes organisées dans une structure cohérente et «qui marche» (on ne se prononce pas pour l'instant quelle est la forme précise d'une telle structure, v. ci-dessous), se présente comme une façon quasi naturelle d'analyser et de modéliser les problèmes complexes et de réaliser les solutions qui découlent de ces analyses.

L'application du second principe de Descartes à la modularité elle-même, en appliquant ainsi la modularité à la modularité, nous invite à envisager plusieurs modules d'exigences que les architectures modulaires devraient satisfaire (cf. à ce propos p.ex. J. A. Fodor, 1981, 1983 ; E. Y. Yourdon, L. L. Constantine, 1979 ; B. Meyer, 1997).

Si l'on y ajoute la précepte contenue dans le troisième principe de Descartes, dans son interprétation forte, non seulement comme le contraire de l'analyse effectuée, étant alors juste un mouvement de synthèse de toutes les parcelles relevées, mais aussi comme le contraire de l'analyse quelle qu'elle soit, étant alors un mouvement de synthèse qui n'est pas forcément

fondé sur toutes les parcelles relevées, on retrouvera en fait toutes les exigences qu'on doit formuler par rapport à un système modulaire (cf. aussi p.ex. B. Meyer, 1997).

Une architecture modulaire doit donc se caractériser par :

- la décomposabilité,
- la composabilité,
- la compréhensibilité,
- la continuité.

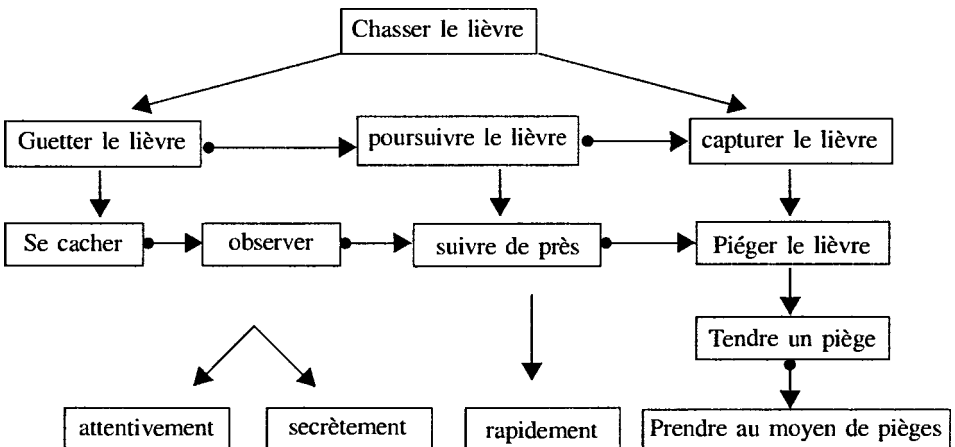
2.1. Décomposabilité modulaire

Une architecture est modulaire si elle contient un certain nombre de composantes autonomes qui sont suffisamment indépendantes pour qu'on puisse les étudier chacune séparément et qui sont liées les uns aux autres par une structure simple. Les composantes autonomes en question peuvent être relativement complexes et admettre, à leur tour, une décomposition en sous-composantes, la décomposition pouvant continuer jusqu'à ce qu'on trouve qu'on a divisé le problème *en autant de parcelles qu'il se pourrait, qu'il serait requis pour les mieux résoudre.*

Une telle décomposition permet d'analyser les modules chacun à son tour et indépendamment les uns des autres dans la mesure où les interdépendances éventuelles entre eux, que l'on doit réperer, sont minimales.

Les approches du type haut – bas, partant d'une description très générale du système / du problème et détaillant ses éléments dans les étapes ultérieures de l'analyse satisfont en général cette exigence.

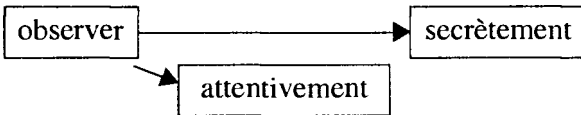
P. ex. le «problème» *comment chasser le lièvre*, pourrait être décomposé, en simplifiant la chose, de ce point de vue, du haut en bas, en modules suivants :



où les flèches du type \longrightarrow indiquent une décomposition logique (sémantique) du problème en modules et les flèches $\bullet\longrightarrow$ indiquent une décomposition logique (sémantique) en modules ordonnés temporellement (les modules en question formant ainsi, conjointement, un script).

2.2. Composabilité modulaire

La composabilité est le contraire de la décomposabilité et admet la possibilité d'employer les modules analysés, qui doivent, par conséquent, être suffisamment autonomes, en dehors de leur contexte de problème d'origine, p. ex. le module contenant les informations morphologiques, faisant partie des modules composant la description de la langue dans une approche théorique déterminée, peut être utilisé dans d'autres types de synthèse linguistiques, aux fondements et buts différents de ceux qui présidaient à l'analyse en question. Dans l'exemple ci-dessus, le module



est tout à fait composable dans d'autres applications, p.ex. du type *guetter l'ennemi*.

La décomposabilité et la composabilité étant des tendances contraires, il peut y avoir des systèmes qui satisfont à l'une, mais non pas à l'autre. En général, la décomposabilité du problème en modules s'accompagne d'une contextualisation relativement forte de modules obtenus : ils sont fonction de différentes facettes du problème et de différente nature éventuelle des éléments qui entrent en jeu et il n'est pas rare qu'ils ne puissent pas être réutilisés tels quels dans la synthèse de la solution d'un autre problème et il faut leur apporter des modifications nécessaires pour que cela puisse se faire (d'où l'insistance ci-dessus sur l'interprétation forte du troisième principe de Descartes pour pouvoir y arriver). Tel p.ex. le module

Prendre au moyen de pièges

ne pourrait pas être réutilisé tel quel dans la solution du problème *chasser un ours* p.ex., puisqu'il comporte l'élément *piège* qui diffère d'une chasse à l'autre (l'architecture de la solution du problème *chasser le lièvre* est donc une architecture modulaire mixte, certains de ces modules admettant une composabilité extérieure, d'autres non) ; telle aussi p.ex. une flèche d'arc, l'un des modules de la solution du problème de la défense / de l'attaque, qui peut être utilisée comme une arme blanche, et non pas comme un projectile, moyennant pourtant quelques adaptations du maniement vu sa fonction télique d'origine.

2.3. Compréhensibilité modulaire

La compréhensibilité modulaire, à son tour, est fonction du premier principe de Descartes qui nous invite à *ne comprendre rien de plus en mes jugements que ce qui se présenterait si clairement et si distinctement à mon esprit que je n'eusse aucune occasion de le mettre en doute*. De ce point de vue, chaque module devrait être compréhensible sans avoir recours à la compréhension d'autres modules. Dans l'exemple ci-dessus, en fait tous les modules sont compréhensibles d'eux-mêmes sans qu'on ait besoin d'aller consulter un autre pour voir ce que signifie celui que l'on est en train d'analyser ; d'autre part, et c'est là l'une des lumières du troisième principe de Descartes, *en commençant par les objets les plus simples et les plus aisés à connaître, pour monter peu à peu comme par degrés jusques à la connaissance des plus composés*, on en vient aux objets (modules) plus composés qui, pour avoir la signification qu'ils ont dans la solution du problème analysé, doivent être pris ensemble avec d'autres modules, ce qui, favorisant une vue synthétique, holistique, du problème, relativise, sans l'éliminer naturellement, la compréhensibilité modulaire ; cf. p.ex. les modules

Se cacher

observer

attentivement

secrètement

qui sont tout à fait compréhensibles pris isolément et peuvent, grâce à cela et leur composabilité faire partie d'autres structures modulaires, p.ex. *Le Conseil de Ministres délibère secrètement à huis clos, le savant lit attentivement un article, Descartes observe le monde ou Un fuyard se cache* ; il est pourtant, d'autre part évident, que, pour pouvoir comprendre ce qui est *plus composé*, dans une structure hiérarchique donc, comme celle-ci, on aura besoin d'intégrer plusieurs modules en même temps : p.ex. pour comprendre le module

Guetter le lièvre

on aura recours aux quatre modules cités ci-dessus ; autant dire que la compréhensibilité modulaire est dans une relation dialectique entre la vision «dépendantiste» des modules par rapport au problème et / ou d'autres modules vus comme un tout et une vision «plate», «ensembliste», des modules qui tous ensemble, sans hiérarchie, représentent le problème en question. Le type de vision dépend, d'une part, de la nature (hiérarchique ou pas) du problème lui-même, et, d'autre part, du niveau de l'analyse auquel on se place : sous-modules d'un module ou du problème, ou modules du problème, ou problème entier. Ce dernier cas est intéressant dans la mesure où, admettant même que les modules

Guetter le lièvre

poursuivre le lièvre

capturer le lièvre

soient compréhensibles indépendamment les uns des autres, leur structure est du type hiérarchique temporel, et, par conséquent, le problème entier n'est résolu / n'est à comprendre que si l'on observe l'ordre de la présentation / de l'application des modules : ainsi, le troisième module ne peut fonctionner correctement que s'il suit la réalisation du second et celui-ci ne peut fonctionner correctement que s'il suit la réalisation du premier.

2.4. Continuité modulaire

Le quatrième critère d'une construction modulaire, celui de la continuité modulaire est une caractéristique intéressante d'une architecture modulaire et est liée d'une manière naturelle, puisque logique, à d'autres paramètres déjà cités, surtout celui d'extensibilité, de décomposabilité et de composabilité.

La continuité modulaire signifie que de petits changements dans la spécification du problème vont effectuer des changements juste dans un module ou dans un nombre restreint des modules individuels et non pas dans le système entier (cf. p.ex. B. Meyer, 1997).

Ainsi p.ex. dans le cas de notre problème *chasser le lièvre* ci-dessus, le changement dans la spécification du problème consistant en ceci que ce n'est pas un lièvre, mais un loup que l'on va chasser entraîne un petit changement dans le module

Tendre un piège

parce que les pièges que l'on tend pour attraper un lièvre et un loup diffèrent ; d'autre part, si c'était un ours que l'on chassait et non pas un lièvre, ce changement dans la spécification du problème entraînerait un changement non seulement dans le module ci-dessus, mais aussi dans celui qui le précède dans la hiérarchie temporelle :

Suivre de près

dans sa partie «manière», à savoir

rapidement

la vitesse dans tous les trois cas envisagés n'étant pas la même.

La question intéressante qui s'impose dans ce contexte est celle de savoir combien de modules et comment peuvent être modifiés pour qu'on

puisse parler encore de la même architecture ou plutôt d'un même type d'architecture du (même type de) problème.

Le problème, *mutatis mutandis*, est du même genre que celui que l'on pose en sémantique cognitive, à savoir quand on dépasse les limites d'une catégorie centrée autour d'un prototype ou les limites entre les représentants moins typiques d'une catégorie, mais non pas encore atypiques.

Aussi bien là qu'ici, il n'y a pas de réponse automatique : dans le cas du problème analysé ci-dessus, on pourrait dire que les changements de spécification du problème consistant à changer d'objet chassé à l'intérieur de la catégorie « gibier » tout d'abord (*lièvre, loup, ours...*), puis à y mettre ce qui n'est pas du gibier, p.ex. *homme, papillon*, etc., ensuite, à changer d'objet qui chasse, p.ex. *lion, ours, loup*, etc. conservent, moyennant les adaptations nécessaires de certains modules, l'architecture de départ.

Par contre, un changement du type d'objet chassé en *vache* (p.ex. *chasser les vaches aux champs*), ou en *antiquités* (*chasser les antiquités*), c'est-à-dire. un changement qui entraîne une modification radicale des modules ou même leur suppression (de tous les modules ou d'une bonne partie d'eux) (cf. l'exemple de *vache*, où tous les modules sont supprimés et remplacés par d'autres), la continuité modulaire est soit entamée soit détruite.

3. Approche orientée objets et description de la langue

Le choix de l'architecture modulaire du système ne fournit pas encore la réponse à la question de savoir quelle est la nature des modules.

En principe, nous avons, dans notre cas, deux types de candidats au statut de modules : unités d'une décomposition fonctionnelle ou unités d'une décomposition en types d'objets.

Dans un langage plus courant, on pourrait se demander si les modules en question seront constitués par les prédicats ou par les arguments (objets).

Quelque soit le choix que l'on va faire, une chose est certaine : le choix en question n'exclut aucunement les uns ou les autres de la représentation complète, achevée, de la langue. C'est seulement leur position et leur statut respectifs qui changent, c'est soit les uns soit les autres qui sont choisis comme le coeur de l'architecture du système, conférant ainsi aux autres le statut, d'ailleurs applicable récursivement, d'humbles et obéissants serveurs.

3.1. Approche orientée objets et types d'objets

L'approche que nous allons présenter ci-dessous, qui est orientée objets, nous impose le choix de (types d') objets comme modules de l'architecture de représentation.

La définition d'un (type d') objet est de type opérationnel et est fournie dans cette optique par les opérations de toutes sortes qu'il peut effectuer ou qui peuvent être effectuées sur lui (du type donc : *dis-moi ce que tu peux faire ou ce que l'on peut faire avec toi, je te dirai qui tu es*) dont les opérations d'«accès» qui débouchent, en dernière instance, sur une caractérisation de sa structure interne représentée par un ensemble (de valeurs) d'attributs qui y sont du type cette fois-ci : *dis-moi comment tu es, je te dirai qui tu es*.

De ce point de vue, un objet donné est caractérisé donc, si l'on se plaçait à un niveau primaire d'observation, par les opérations et d'autres objets qui l'accompagnent dans les situations représentées, ce qui est une conséquence directe de la définition ci-dessus : pour pouvoir opérer sur un objet, il faut être en contact avec lui, ce qui se traduit du point de vue observationnel pur par la co-présence des objets et des opérations dans les situations.

Si on le traduisait dans le contexte de l'analyse de la langue, on dirait que, les situations étant représentées par les phrases, la définition d'un objet linguistique (la définition du sens d'un mot ou de la syntaxe ou de la morphologie d'un mot p.ex.) est fournie par la co-présence des objets et des opérations d'un certain type dans les phrases.

Cette manière de voir le sens des mots à laquelle on vient par un chemin philosophico-opérationnel est proche des travaux p.ex. de l'école anglaise de linguistique (cf. p.ex. le fameux slogan de J. R. Firth : *You shall know a word by the company it keeps*, J. R. Firth, 1951; M. A. K. Halliday, R. Hasan, 1976; J. M. Sinclair, 1987), des recherches philosophico-sémantiques de L. Wittgenstein (1953) (cf. sa *ressemblance de famille*) et, ces derniers temps, des travaux de G. Gross (1992, 1994a, b, 1995, 1997) et de son école, insistant sur la description du sens d'un mot en termes de son emploi.

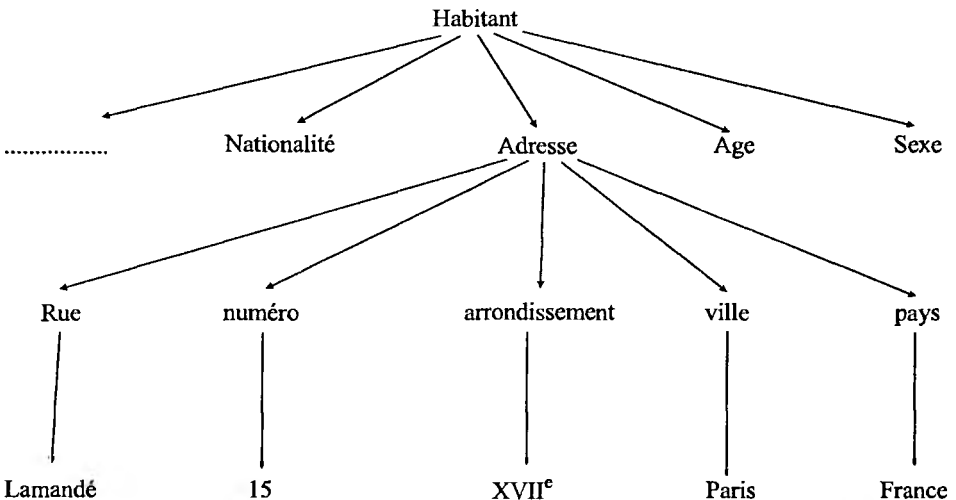
3.2. Objets, fonctions et attributs

Dans une approche orientée objets, le statut d'objet est un statut fonctionnel et n'a rien à voir avec les caractéristiques ontologiques des objets du monde extérieur, bien que les objets concrets, perceptibles par les sens, en soient le prototype. Par contre, le type de représentation linguistique, la façon donc dont la langue considère les entités extralinguistiques, joue un rôle important et nous invite à considérer p.ex. ce qui a la forme d'un substantif comme objet de départ de l'analyse.

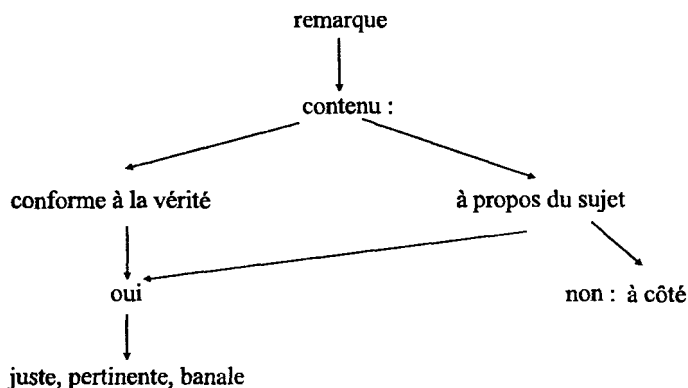
Ainsi, par un mouvement dialectique si l'on se situe au point de vue philosophique, ou par un mouvement récursif, si l'on se situe au point de vue opérationnel, ce qui est un objet à un niveau d'analyse peut devenir attribut ou opération (d'une manière ou une autre : prédicat) à un autre niveau, et, inversement, ce qui est un attribut ou une opération à un niveau d'analyse peut devenir objet à un autre niveau, la stabilité de la différence entre objets et opérations étant assurée par la possibilité de constituer ou non l'objet de départ de l'analyse, caractérisé, comme nous l'avons vu ci-dessus par un ensemble d'opérations et d'attributs qui y sont applicables.

C'est ainsi p.ex. que *adresse* est un des attributs de l'objet *habitant* et prend différentes valeurs selon les cas : *tel numéro, telle rue, tel quartier, telle ville, tel pays*. Ce même attribut deviendra toutefois objet si on le pose comme objet et si l'on essaye de trouver, récursivement, ses attributs à lui : l'objet *adresse* a comme attributs, p.ex., les paramètres tels que *numéro de rue, rue, quartier, ville, pays*, qui sont saturés de différentes façons suivants les cas.

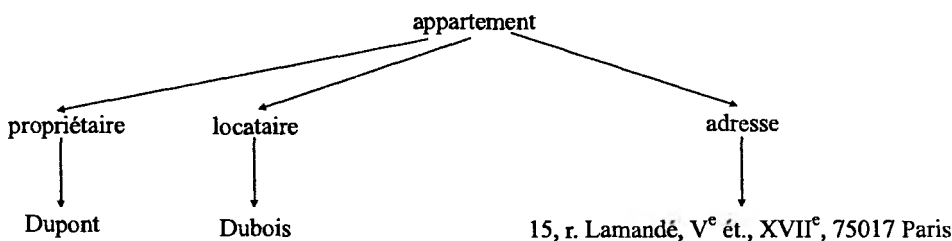
C'est donc cette vision de structuration de la réalité en parcelles-embêtements qui permet, dans le jeu dialectique d'objet et d'opération / d'attribut, d'assigner, par un mouvement d'analyse récursive, différents statuts fonctionnels aux mêmes entités ontologiques en fonction de leur rôle dans l'analyse effectuée ou en fonction de leur représentation linguistique (cf. p.ex. aussi la différence entre *autorisation* \Rightarrow *X donne l'autorisation à Y de faire Z* vs *autoriser (X, Y, à faire Z)*) :



Cf. aussi un même type de relations dialectiques et d'analyse récursive dans le cas des adjectifs descriptifs du type p.ex. *juste* qui est une valeur de l'attribut *contenu* de l'objet *remarque* :



ou encore des objets concrets inanimés :



Le sens en question étant donc considéré comme un objet (qu'il s'agisse donc du mot qui se réfère à un objet concret ou à une abstraction), il est caractérisé par les attributs et leurs valeurs qu'on peut énoncer à son propos et par les opérations qu'il peut effectuer ou que l'on peut effectuer sur lui.

Remarquons à ce propos une différence de comportement d'attributs dans le cas de *adresse* et *appartement* ci-dessus de ce point de vue, le premier étant caractérisé par les attributs et leurs valeurs qui peuvent être, à leur tour, des attributs (*juste* pouvant avoir, à son tour, dans d'autres structures, des valeurs, p.ex. la vérité et la fausseté, cf. p.ex. : le juge *X* : attribut : *juste*, valeur : *oui* ; le juge *Y* : attribut : *juste*, valeur : *non*) et le second étant caractérisé par les attributs et leurs valeurs qui ne peuvent pas être des attributs.

Les attributs et les valeurs en question peuvent être considérés comme déterminant la «structure» interne de l'objet (prédiquant les valeurs typiques, ou s'en éloignant, des paramètres inhérents à l'objet en question), et les opérations comme déterminant ses fonctionnalités, donc aussi, comme on l'a vu, ses capacités d'apparaître dans les contextes déterminés dans les textes. Les deux définissent ensemble les propriétés du sens donné.

Cette manière de voir la description des sens des mots inverse les priorités dans la description des sens des mots que nous avons appliquées dans le cadre des structures prédicats – arguments (v. p.ex. S. Karolak, 1972, 1977, 1984, 1992, 1998a, b ; K. Bogacki, S. Karolak, 1991 ; W. Banyś, 1981, 1983, 1984, 1985, 1988) et se situe du côté du camp des objets dans cette bataille éternelle entre la décomposition primaire du monde en objets ou en fonctions : là, le point de départ est constitué par le sens (prédicatif) d'un mot, autour duquel, grâce à l'opération de l'implication sémantique (nécessaire) (v. p.ex. S. Karolak, 1974, 1984 ; A. Bogusławski, 1973, 1974, 1977 ; W. Banyś, 1989), on établit la structure sémantique du mot (d'où justement les structures **prédicat** – arguments et non pas **arguments** – prédicat). Ainsi, pour prendre un exemple simple, lorsqu'on a p.ex. le prédicat *aboyer*, la question est de savoir, dans cette optique, quels sont les «objets» qui peuvent remplir la position d'argument ouverte par le prédicat et l'on en vient nécessairement aux «objets» animés : chiens.

La procédure peut être résumée donc comme suit :

On part de f (fonction propositionnelle, prédicat) pour arriver à ses arguments («objets»), soit :

$$f \Rightarrow x.$$

La procédure orientée objet renverse la direction d'étude et nous invite à se poser la question d'un autre type, à savoir, étant donné qu'on a un objet x , quels sont les prédicats qui lui peuvent être assignés, les prédicats se répartissant de manière supplémentaire en attributs (adjectifs et toutes sortes de compositions N (*Prép*) N) et en opérations (verbes).

Elle se résume donc par la formule suivante:

$$x \Rightarrow f.$$

Dans le cas ci-dessus, la question serait de savoir, étant donné qu'on a un objet du type «chien», quelles sont ses attributs typiques et les opérations qu'il peut effectuer et qui peuvent être effectuées sur lui ; la réponse qu'on donne est une caractérisation de l'objet en question comportant, entre autres, p. ex. les attributs suivants *savant*, *dressé*, *méchant*, *hargneux*, *docile*, et les opérations du type p.ex. *mordre*, *aboyer*, *dresser*, *siffler*, *clabauder*, *clatir*, *glapir*, *gronder*, *hurler*, *japper*.

D'une manière générale, les adjectifs «descriptifs» (et non pas «relationnels», du type *une visite ministérielle* = *la visite d'un ministre*) qualifient, soit du point de vue de la transgression de la norme ou de la conformité à la norme soit du point de vue de son évaluation : *bon / mauvais*, en leur assignant ainsi une valeur, selon le schéma : *attribut – f de l'objet – x = valeur – h*,

- l'attribut que l'objet en question possède objectivement, cf. p.ex. :
 {plus grand que la norme/la moyenne} :
*({attr – poids (objet – valise)} = valeur – lourde (le poids de cette valise est grande = cette valise est lourde),
 (longueur (route) = longue (la longueur de cette route est grande = cette route est longue),
 (évaluation : mauvais (comportement (chien)) = méchant (ce chien se comporte mal = ce chien est méchant),
 (évaluation : bon (comportement (chien))) = docile (ce chien se comporte bien = ce chien est docile),*
- l'une des (ou la) fonction(s) pour lesquels / laquelle l'objet a été créé, cf. p.ex. :
 {plus grand que la norme/la moyenne} :
*(vitesse (conduite (conducteur))) = rapide (la vitesse avec laquelle ce conducteur conduit la voiture est grande = c'est un conducteur rapide),
 {conforme à la norme}{la fonction téléique en faisant partie):
 (pesée (balance)) = juste (une balance est destinée à peser exactement, cette balance pèse comme elle doit peser = c'est une balance juste, précise, exacte),
 (contenu (remarque)) = juste, pertinent (une remarque est destinée à souligner une particularité d'une situation, le contenu de cette remarque le fait = c'est une remarque juste, pertinente),*
- l'une des (ou la) fonction typique d'une action, cf. p.ex. :
 {conforme à la norme} :
*(conforme à la norme ({rendu {jugement}})) = juste (un jugement doit être équitable, celui-ci est conforme à la norme, c'est un jugement juste, équitable),
 (dressage (chien) = dressé, savant) (le dressage d'un chien a pour but de le rendre savant, ce chien a été dressé, c'est chien dressé, savant).*

La nécessité d'avoir recours aux attributs et fonctions de ce type dans la description linguistique est évidente ; on ne pourrait pas s'en passer p.ex. si l'on essaye d'analyser les formules où apparaissent les adjectifs qui évaluent certains attributs des objets (tels *bon / mauvais*) et la fonction typique de l'objet mis en corrélation n'admet pas de caractérisation comparable, cf. p.ex. :

Cette pierre est un bon marteau.

Il serait impossible de comprendre à quoi se rapporte *bon* dans sa relation avec *pierre*, dans cet emploi de type métaphorique, qui s'effectue par l'intermédiaire de *bon (marteau)* si *marteau* n'avait pas de fonction téléique qui lui est assignée : *(enfoncer des clous (marteau) = bon)* (cf., pour ce qui est des attributs et des adjectifs en général, p.ex. les discussions dans E. Sapir, 1944 ; K. Dunker, 1945 ; G. A. Miller, R. Beckwith,

Chr. Fellbaum, D. Gross, K. Miller, R. Teng, 1990 ; J. Pustejovski, 1998 ; cf. aussi W. Banyś, à par. 1).

3.3. Objets et classes d'objets

On obtient ainsi une description du monde, composé (de classes) d'objets du monde, tel qu'il est vu par la langue et dans la langue, ce qui fait que la base de données lexicales de ce type, compilant des informations sur le comportement linguistique des éléments de la langue en fonction aussi des connaissances sur le monde, devient en même temps un thésaurus (cf. p.ex. N. Calzolari, 1988 ; K. Sparck-Jones, 1986; Y. Wilks, 1978).

Si ce n'était la possibilité qu'on voit se réaliser de temps à autre de ne pas distinguer la classe d'objets d'une instance d'objet d'une classe, ce ne serait pas la peine d'y insister, mais, pour la clarté et la précision de l'exposé, on s'arrêtera une seconde à la distinction entre *classe d'objets* et (*instance d'*) *objet*.

Quand nous parlons de la classe d'objets, nous parlons d'un ensemble d'objets qui partagent un certain nombre d'opérations et d'attributs. Un objet est une instance d'une classe d'objets, ainsi p.ex. une instance, un objet concret, de la classe de propriétaires est p.ex. M. Dupont. Ce qui est intéressant pour une description linguistique, c'est naturellement une description de classes et non pas de leurs instances particulières : et si l'on décrit une instance de classe, elle n'est intéressante que comme instance d'une classe justement.

Les confondre, ce serait, dans le cas du pire des scénarios possibles, se retrouver à la place de ce gars de la blague ci-dessous :

Un gars ouvre son journal et il lit : «A Chicago, c'est vingt fois par jour qu'un homme se fait descendre».

Le gars branle la tête et il dit :

– Pauvre type !

H. Nègre (1973 : n° 6)

où l'expression *un homme* a été considérée comme activant le cadre *_homme particulier* et non pas le cadre *_homme en général*, et dans le cas du meilleur des scénarios possibles, se retrouver à la place de quelqu'un qui prendrait au pied de la lettre la boutade de R. G. Ingersoll à propos de Lincoln : *Lincoln was not a type. He stands alone — no ancestors, no fellows, no successors.*

Il est naturellement possible, c'est d'ailleurs souvent le cas, qu'une classe ait des sous-classes comme éléments et non seulement des instances d'objets.

C'est le cas p.ex. ci-dessus de l'attribut *vitesse* qui a comme élément la classe p.ex. *conduite* ou *course*.

Dans certains langages de programmation orientés objets, on appelle ce type de classes qui ont d'autres classes comme instances méta-classes, c'est p.ex. le cas de Smalltalk ; nous n'allons pas le faire vu, d'une part, le caractère très naturel et fréquent du phénomène en langue naturelle et, d'autre part, le manque de confusion possible si l'on veille à préciser, si besoin se présente, d'une manière supplémentaire, s'il est question de classe d'objets et non pas d'instance d'objet.

Cette approche s'apparente, sans y être identique, à l'approche de «classes d'objets» dans le sens de G. Gross (1992, 1994a, b, 1997). En effet, aussi bien pour Gaston Gross que pour nous, les ensembles des caractéristiques attributives et opérationnelles des sens donnés permettent d'établir des classes d'objets linguistiques (déterminés par le comportement linguistique et non pas «ontologique») auxquelles appartiendront les objets caractérisés par les ensembles, ou leurs intersections, de caractéristiques en question (p.ex. la classe d'objets *moyens de transports, professions, titres, dates*, etc.).

3.4. Caractère opérationnel – «implicite» de la caractérisation de la classe d'objets

La caractérisation d'un type d'objets par l'intermédiaire des opérations qui s'y appliquent est de type ouvert et est donc, de ce point de vue, «implicite», puisqu'elle ne doit pas fournir un *listing* explicite exhaustif de toutes les opérations possibles qui s'y appliquent ; autant dire, la caractérisation en question n'implique pas que les opérations citées soient les opérations uniques qui s'appliquent au type d'objets en question ; en général, il est tout à fait évident que d'autres opérations, ne serait-ce que par le mécanisme d'héritage sémantique, peuvent s'y appliquer aussi (cf. p.ex. *peser* par rapport à un *marteau* qui est tout à fait possible, mais qui n'est pas listé parmi les opérations qui s'appliquent à un marteau : cela est, entre autres, fonction de l'économie de la description : de la possibilité de déduire, grâce au mécanisme d'héritage, les opérations à appliquer venant des super-type(s) de la classe d'objets étudiée et de la distinction entre les opérations / les attributs «appropriés» (cf. p.ex. G. Gross, 1992, 1994a, b) et les opérateurs généraux (les raisons d'économie descriptive nous ont amené aussi à supprimer des éléments descriptifs présentés ci-dessus le module où le type d'objet en question peut être attribut d'un autre type d'objets (cf. à ce propos p.ex. W. Banyś, à par. 2). La distinction entre un opérateur approprié et un opérateur général – un opérateur approprié étant un opérateur s'appliquant à la

classe en question ou à quelques classes en question et un opérateur général étant un opérateur qui s'applique à beaucoup, sinon à la plupart, de classes d'objets – détermine, couplé avec les mécanismes d'héritage sémantique, le type d'opérations à lister dans le cas d'une classe d'objets particulière.

Les deux distinctions : *opérateur approprié* / *opérateur hérité* et *opérateur approprié* / *opérateur général*, bien que très proches l'une de l'autre et allant, fonctionnellement dans la même direction, ne se recouvrent forcément pas, ni du côté *approprié* ni du côté *hérité* / *général*.

Si l'on prend p.ex. le cas de la classe d'objets *marteau* et *tomate* qui font, respectivement, partie de la super-classe¹ *instruments* et *aliments* qui, à leur tour, font partie de la super-classe² *inanimés* qui, à son tour, appartient à la super-classe³ *concrets*, on voit que nous pouvons appliquer à *marteau* p.ex., dans le cadre de la première distinction, d'une part, l'opération de *enfoncer les clous avec* (opérateur approprié représentant la fonction téléique), et, d'autre part, *prendre dans la main / couper en morceaux* (opérateur hérité représentant l'un des éléments du script *enfoncer les clous avec / préparer comme (partie d'un) plat*, hérité de la classe d'objets *instruments [manuels] / aliments*) ce qui ne serait sans doute pas considéré du point de vue de la seconde distinction comme approprié, puisque l'on peut prendre tout concret dans la main ; cela est parfaitement vrai, mais toujours est-il qu'il y a peu de concrets qu'il faut prendre dans la main pour réaliser leur fonction téléique, je peux d'une part, p.ex. prendre une chaise, une pierre, une chèvrete dans la main, et, d'autre part, couper en morceaux une étoffe, un journal sans que cela soit une opération appropriée, même par héritage.

3.5. Objets et fonctions : modularité et abstraction

D'une certaine manière, cela pourrait paraître bizarre que, après avoir discuté la question des meilleurs candidats au statut de modules du système descriptif, après avoir déclaré la primauté des objets dans la description linguistique et en être venu à la caractérisation des objets, on ait recours aux fonctions pour les définir.

Il n'est pourtant pas étonnant que, après avoir choisi la modularité comme l'un des principes de la description et après avoir choisi les objets comme modules de départ, on en vienne aussi aux fonctions. Tout d'abord, parce que les objets et les fonctions sont comme le champignon et l'algue dans un lichen : celui-ci n'existerait pas sans l'une et l'autre en même temps ; on pourrait le dire aussi à la manière d'Andrzej Bogusławski (1973) qui disait que les prédicats et les arguments font partie d'un tuyau : les prédicats forment le conduit et les arguments – les ouvertures, le tuyau n'existerait ni sans conduit ni sans ouvertures.

Mais ce qui distingue fondamentalement l'approche orientée objets des approches «classiques», fonctionnelles, c'est la distribution des rôles : ce sont les (classes d') objets qui constituent les modules de départ et ce sont les fonctions qui deviennent leurs serviteurs. Les fonctions, déterminant le type de classe d'objets, sont encapsulées, comme le disent les informaticiens, dans les modules de l'architecture descriptive. Aucune fonction n'existe dans une approche orientée objets de par elle-même, elle n'existe, encapsulée, qu'à l'attache à une classe d'objets.

C'est davantage une approche pragmatique du type *Dis-moi ce que tu as, pour que je sache ce que je peux, et comment, obtenir de toi* que celle du type *dis-moi qui tu es*. Si je suis un mécanicien, c'est toutes sortes d'instruments, dont un marteau, dont j'ai besoin ; si je suis un linguiste, c'est toutes sortes de textes, de sons, de mots, etc. dont j'ai besoin ; si je suis un chasseur, c'est toutes sortes d'instruments de chasse dont j'ai besoin ; les objets ne sont intéressants pour moi que si j'en ai besoin ; je peux ne pas avoir, certes temporairement dans le cas de la langue, besoin de certaines classes d'objets en construisant une architecture modulaire donnée, alors ils n'existent pas pour moi, ils n'entrent pas dans l'architecture que je suis en train de construire.

Le monde des objets est, de ce point de vue, une société d'agents qui sont en contact et qui se communiquent (cf. aussi, dans ce contexte, les idées de M. Minsky, 1986) et la modularité sous cette forme va la main dans la main avec l'abstraction et l'élimination des éléments qui ne sont pas pertinents dans le cadre d'une analyse donnée.

Dans W. Banyś (2002), nous esquissons la manière dont on peut organiser ce monde d'objets linguistiques.

Références

- Banyś W., 1981: «Descriptions indéfinies : arguments ou prédicats en position d'argument?» *Linguistica Silesiana*, 1981, 4.
- Banyś W., 1983: *Ambiguïté référentielle des phrases à descriptions indéfinies en français*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Banyś W., 1984: «Sémantique, structure, syntaxe et lexique». *Cahiers de Lexicologie*, 45.
- Banyś W., 1985: «Structure thème-rhème dans une grammaire à base sémantique». *Linguistica Silesiana*, 6.
- Banyś W., 1988: «Sur le dictum thématique : articulations secondaires du rhème». In : Banyś W., Karolak S., eds.
- Banyś W., 1989: *Théorie sémantique et 'si ...alors* ». *Aspects sémantico-logiques de la proposition conditionnelle*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Banyś W., 2002: «Bases de données lexicales électroniques – une approche orientée objets. Partie II : Questions de description. *Neophilologica*, 15, 207–250.

- Banyś W., à par. 1: «Approche orientée objets et fonctions : attributs, valeurs d'attributs, classes d'objets et fonctions lexicales». *Neophilologica*, 16.
- Banyś W., à par. 2: «Format descriptif modifié – modifieur du lexique de spécialité». *Neophilologica*, 16.
- Banyś W., Karolak S., eds, 1988: *Structure thème-rhème dans les langues romanes et slaves*. Wrocław: Ossolineum.
- Baudoin C., Hollowell G., 1996: *Realizing the Object-Oriented Lifecycle*. Upper Saddle River (N.J.): Prentice Hall Object-Oriented Series.
- Biber D., Conrad S., Reppen R., 1998: *Corpus Linguistics. Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Bogacki K., Karolak S., 1991: «Fondements d'une grammaire à base sémantique». *Lingua e Stile*, 26, 3.
- Bogusławski A., 1973: «O analizie semantycznej». *Studia Semiotyczne*, 4.
- Bogusławski A., 1974: «Preliminaries for Semantic-Syntactic Description of Basic Predicative Expressions with Special Reference to Polish Verbs». In: Orzechowska A., Laskowski R., eds.
- Bogusławski A., 1977: *Problems of Thematic-Rhematic Structure of Sentences*. Warszawa: PWN.
- Calzolari N., 1988: «The Dictionary and the Thesaurus can be Combined». In: Evens M., ed.
- Cox B., 1986: *Object-Oriented Programming. An Evolutionary Approach*. N.Y.: Addison-Wesley.
- Delobel C., Lécuse Chr., Richard Ph., 1991: *Bases de données : des systèmes relationnels aux systèmes à objets*. Paris: InterEditions.
- Dunker K., 1945: «On Problem Solving». *Psychological Monographs*, 58 : Whole No. 270.
- Evens M., ed., 1988: *Relational Models of the Lexicon*. Cambridge: Cambridge University Press.
- Fellbaum Chr., ed., 1998: *Wordnet : An Electronic Lexical Database*. Cambridge, Mass., London: The MIT Press.
- Firesmith D. G., Eyckholt E. M., 1995: *Dictionary of Object Technology, the Definitive Desk Reference*. New York: SIGS Books.
- Firth J. R., 1951: «Modes of Meaning». In: *Papers in Linguistics 1934–1951*. London.
- Fodor J. A., 1981: «Methodological Solipsism considered as a Research Strategy in Cognitive Psychology». In: Haugeland J.
- Fodor J. A., 1983: *The Modularity of Mind : An Essay on Faculty Psychology*. Cambridge, Mass.: MIT Press.
- Fontenelle T., Vanandroye J., 1989: «Retrieving Ergative Verbs from a Lexical Database». *Dictionaries : Journal of the Dictionary Society of North America*, 11, 11–39.
- Fontenelle T., 1997a: *Turning a Bilingual Dictionary into a Lexical-Semantic Database*. Tübingen: Max Niemeyer Verlag.
- Fontenelle T., 1997b: «Using a Bilingual Dictionary to Create Semantic Networks». *International Journal of Lexicography*, 10, 4, 275–303.
- Forsgren M., Jonasson K., Kronning H., eds, 1998: *Prédication, assertion, information. Actes du Colloque d'Uppsala en linguistique française, 6–8 juin, 1996*. Uppsala, Ed. de l'Univ. d'Uppsala, 243–254.
- Garside R., Leech G., McEnery A., eds, 1997: *Corpus Annotation. Linguistic Information from Computer Text Corpora*. London and New York: Longman.
- Grochowski M., Karolak S., Topolińska Z., eds, 1984: *Gramatyka współczesnego języka polskiego*. Warszawa: PWN.
- Gross G., 1992: «Forme d'un dictionnaire électronique». In: *La station de traduction de l'an 2000*. Éd. Presses de l'Université du Québec.

- Gross, G., 1994a: «Classes d'objets et description des verbes». *Langages*, 115 [Paris: Larousse], 15–31.
- Gross G., 1994b: «Classes d'objets et synonymie». In: *Annales Littéraires de l'Université de Besançon*. Série Linguistique et Sémiotique. Vol. 23. Besançon, 93–103.
- Gross G., 1995: «A propos de la notion d'humain». In: *Linguisticae Investigationes Supplementa*. N° 17: *Lexiques-grammaires comparés en français*. Actes du Colloque International de Montréal (3–5 juin 1992). Montréal, 71–80.
- Gross G., 1997: 'La grammaire, les dictionnaires et l'informatique». *Les dictionnaires de langue française et l'informatique*. Actes du Colloque «La Journée des dictionnaires». Université de Cergy-Pontoise, Centre de Recherche Texte/Histoire, 55–65.
- Guo C. M., ed., 1992: *Machine Tractable Dictionaries : Design and Construction*. Norwood, N.J.: Ablex.
- Habert B., Nazarenko A., Salem A., 1997: *Les linguistiques de corpus*. Paris: Armand Colin.
- Halliday M. A. K., Hasan R., 1976: *Cohesion in English*. London: Longman.
- Haugeland J., 1981: *Mind Design*. Cambridge Mass.: MIT Press.
- Karolak S., 1972: *Zagadnienia składni ogólnej*. Warszawa: PWN.
- Karolak S., 1974: «O programie składni wyrażen predykatywnych w gramatyce języka polskiego». In: Orzechowska A., Laskowski R., red.
- Karolak S., 1977: „Z problematyki opisu wyrażen predykatowo-argumentowych”. *Studia Gramatyczne*, 1.
- Karolak S., 1984: „Składnia wyrażen predykatywnych”. In: Grochowski M., Karolak S., Topolińska Z., red., 11–210.
- Karolak S., 1992: „Gramatyka a kształtowanie świadomości poznawczej”. In: *Prace Filologiczne*. T. 27. Warszawa, 166–174.
- Karolak S., 1998a: «Sur la structure sémantico-combinatoire des concepts». In : *XXIF Congrès International de Philologie Romane*. Bruxelles, juin.
- Karolak S., 1998b: 'La structure conceptuelle et la forme syntaxique de la phrase». In : Forsgren M., Jansson K., Kronning H., eds, 243–254.
- Lausen G., Vossen G., 1997: *Models and Languages of Object-Oriented Databases*. N. Y.: Addison-Wesley Longman Limited.
- Masini G., Napoli A., Colnet D., Léonard D., Tombre K., 1989: *Les langages à objets*. Paris: InterEditions.
- Meyer B., 1997: *Object-Oriented Software Construction*. Upper Saddle River (N.J.): Prentice Hall Object-Oriented Series, 2d ed.
- McEnery A., Wilson T., 1996: *Corpus Linguistics*. Edinburgh: Edinburgh Univer. Press.
- Miller G. A. ed., 1990: *WordNet : An on-line Lexical Database*. Special issue of *International Journal of Lexicography*, 3, 4.
- Miller G. A., Beckwith R., Fellbaum Chr., Gross D., Miller K., Teng R., 1990: «Five Papers on WordNet». *International Journal of Lexicography*, 3, 4.
- Minsky M., 1986: *The Society of Mind*. New York, Simon & Schuster (trad.fr. : *La Société de l'esprit*. Paris: InterEditions 1988).
- Nègre H., 1973: *Dictionnaire des histoires drôles*. Paris: Fayard.
- Ooi V. B. Y., 1998: *Computer Corpus Lexicography*. Edinburgh: Edinburgh University Press.
- Orzechowska A., Laskowski R., red., 1974: *O predykcji*. Wrocław: Ossolineum.
- Page-Jones Meilir, 1995: *What Every Programmer Should Know about Object-Oriented Design*. New York: Dorset House.
- Pustejovsky J., 1998: *The Generative Lexicon*. Cambridge Mass., London: MIT Press.
- Sapir E., 1944: «Grading: A Study in Semantics». *Philosophy of Science*, 11, 83–116.
- Sinclair J. M., ed., 1987: *Looking up*. Glasgow, Collins.

- Sparck-Jones K., 1986: *Synonymy and Semantic Classification*. Edinburgh: Edinburgh University Press.
- Stroustrup B., 1989: *Le langage C++*. Paris: InterEditions.
- Yourdon E. N., Constantine L. L., 1979: *Structured Design : Fundamentals of a Discipline of Computer Program and Systems Design*. Englewood Cliffs (N.J.): Prentice Hall.
- Wilks Y., 1978: «Making Preferences More Active». *Artificial Intelligence*, **11**, 75–97.
- Wilks Y., Slator B. M., Guthrie L. M., 1996: *Electric Words. Dictionaries, Computers, and Meanings*. Cambridge, Mass., London: The MIT Press.
- Wittgenstein L., 1953: *Philosophical Investigations*. Oxford: Blackwell.

Wiesław Banyś

Elektroniczne bazy danych leksykalnych – ujęcie zorientowane obiektowo Część I: Zagadnienia modularności

Streszczenie

To, że informacje leksykalno-semantyczne – obok informacji fonetycznych, morfologicznych, składniowych – muszą być w sposób konieczny wzięte pod uwagę w badaniach, które mają być skutecznie w automatycznej analizie i syntezy, jest generalnie przyjęte jako pewnik.

W tej kwestii jawi się jednak wiele wątpliwości, z których kilka poruszono w prezentowanym tekście.

Pierwsza dotyczy samej konieczności tworzenia słowników elektronicznych: nie chodzi przecież bowiem tylko o przedstawienie w postaci elektronicznej informacji zawartej w formie papierowej; to, co jest istotne, to możliwość zebrania razem nie tylko informacji, które są zawarte *explicite* w słownikach tradycyjnych, ale także wydobyć z nich i zaprezentować informacji zawartych w nich *implicite* (np. hierarchie zależności semantycznych), dorzucenie do nich nowych informacji, tak by utworzyć rzeczywistą elektroniczną poliwalentną bazę danych leksykalnych.

Problem jest zatem być może bardziej typu „strukturalnego” niż „faktualnego”, związany bardziej z organizacją i prezentacją informacji leksykalnej niż z samą informacją.

Ten stan rzeczy każe nam postawić pytanie o typ opisu informacji w takiej bazie danych z punktu widzenia zarówno zawartości, jak i formatu, który byłby w stanie zaspokoić wymagania prezentacji danych, która byłaby w stanie zaspokoić wymagania na modyfikacje i ewentualne wzbogacenia, a jednocześnie nie stawiała pod znakiem zapytania integralności samej bazy.

Te założenia wymuszają taki typ opisu danych, który byłby wystarczająco elastyczny i ogólny, by mógł być zintegrowany ewentualnie z różnymi ujęciami teoretycznymi. Okazuje się jednocześnie, iż wymagania rozszerzalności i używalności do wielu celów narzucają, wprawdzie nie w sposób absolutnie konieczny, ale jednak, pewne ujęcie teoretyczne samego opisu – ujęcie opisowe, które jest zorientowane obiektowo.

Artykuł przedstawia założenia takiego opisu zorientowanego obiektowo, zgodne z czterema zasadami dobrej metody Kartezjusza.

Wiesław Banyś

**Electronic lexical databases: the object-oriented approach
Part One: The problems of modularity**

Summary

It is taken for granted that apart from phonetic, morphological and syntactic information, lexical and semantic information must be taken into account in the studies whose result will be a successful automatic synthesis and analysis. However, some doubts arise here and they are discussed in this article.

The first issue is the necessity of making electronic dictionaries that should not only copy the contents of standard dictionaries with all their explicit information, but should also include some of the implicit information contained in traditional dictionaries, e.g. hierarchies of semantic dependencies. Some extra information should be added so that we should have polyvalent electronic lexical databases.

The major problem here is the adequate organization and presentation of the lexical information rather than the information itself. One should consider the best ways of describing the information in terms of contents and format that would meet the requirements of adequate presentation. It should be expandable, modifiable and multipurpose, but the integral character of the database should be preserved. The data description should be general, flexible and compatible with different theoretical approaches.

The article presents the major assumptions of such object-oriented description, made according to the four principles of a good method put forward by Descartes.