

Grażyna Vetulani

Université A. Mickiewicz
Poznań

Répercussions des activités humaines dans les collocations verbo-nominales

Abstract

In the paper we focus on the relationship between knowledge and the world. We address this issue through the study of the class of *verb + noun collocations*. We notice that valid classifications of human behavior may be based on properties of the verb constituents of such structures. Verb + noun collocations appear to be a useful tool for the design of linguistically motivated ontologies.

Keywords

Ontological systems, verb + noun collocations

0. Introduction

Dans ce travail, nous proposons d'étudier le sujet du colloque (*Les images du monde dans les langues*) dans la direction de la représentation au niveau linguistique des comportements et des activités de l'homme. Plus particulièrement, nous nous concentrons sur le rapport entre les connaissances sur les activités et les constructions à structure: **verbe support + nom abstrait** appelées aussi **collocations verbo-nominales**, **verbes analytiques** ou **syntagmes conventionnels**.

1. État de recherches

Les collocations verbo-nominales ont déjà été l'objet de nombreuses analyses linguistiques. Initialement, on les comparait seulement avec les verbes, en examinant les rapports de ressemblance et de différence entre les unités synthétiques (les verbes comme : *voyager, lire, gifler*, etc.) et les unités analytiques (du type : *être en voyage, faire de la lecture, donner une gifle à qqn* et ainsi de suite).

Plus tard, on décrivait avec plus d'attention la nature et le rôle du verbe entrant dans la structure des collocations verbo-nominales. Par exemple, dans les travaux linguistiques de la première période des recherches sur les prédicats nominaux (les noms abstraits) du français, on décrivait le verbe apparaissant avec ce type de prédicats comme un verbe dépourvu (dans ce contexte) de son sens plein, c'est-à-dire comme un élément de fonction (*Funktionsverben* en allemand) qui actualise (« conjugue ») le prédicat dans le discours (M. Gross, 1981). On parlait alors du verbe prédicativement vide, d'un outil grammatical à la manière d'un verbe auxiliaire et on l'appelait le plus souvent **verbe support** (plus loin : *VSup*).

Dans les travaux qui suivaient (pour le français : G. Gross, R. Vivès, 1986; Ibrahim, dir., 1996; pour le polonais : G. Vetulani, 1994), on analysait le *VSup* plus dans les détails, en reconnaissant que son apport est important pour l'identification du sens de la collocation entière. En effet, dans le cas des collocations verbo-nominales, c'est le *VSup* qui fournit les informations sur le registre employé (l'apport stylistique) et / ou sur les catégories de : temps, aspect, personne, nombre, etc. (l'apport grammatical).

Les collocations verbo-nominales s'avèrent un bon champ d'observation non seulement dans l'analyse des propriétés grammaticales et fonctionnelles des éléments cooccurrents, mais elles constituent aussi un bon exemple de répercussions au niveau linguistique des phénomènes sociaux.

Dans la suite, nous basons nos considérations sur ce que nous avons pu observer lors de l'étude systématique des prédicats nominaux du polonais. L'examen des emplois de ces unités a permis notamment de les présenter dans leurs environnements syntaxiques selon le modèle du fonctionnement, à savoir : *N0 VSup (Déf, Mod) Npréd (Prép, Déf) N1 (Prép, Déf) N2...* où, à côté du verbe et du nom (*VSup + Npréd*), nous retrouvons d'autres éléments importants pour la définition du sens, en l'occurrence le sujet et les compléments (respectivement : *N0, N1, N2...* représentant les arguments, introduits parfois par une préposition *Prép*), les déterminants (*Déf*) et tout type de modificateurs (*Mod*). En fin de compte, nous avons dénombré non seulement un ensemble d'unités simples (prédicats nominaux), mais des modèles grammaticaux illustrant différents emplois de ces unités. Les modèles obtenus correspondent

aux discours possibles du type : *wymierzyć komuś jakiś mandat za coś* (en fr. : infliger une amende à qqn), *mieć wątpliwości co do czegoś* (en fr. avoir des doutes sur qqch.), *udzielić komuś zgody na coś* (en fr. donner à qqn l'autorisation de faire qqch.), *wykazywać zainteresowanie czymś* (en fr. avoir de l'intérêt pour qqch.) etc. (G. Vetulani, 2000).

2. Tentatives de classifications

Dans la langue française, l'analyse des structures d'emploi des prédicats nominaux a conduit les spécialistes à les classer en fonction des verbes sélectionnés par ces prédicats (voir les travaux effectués au L.A.D.L. à Paris 7, sous la direction de M. Gross). Puisque le plus souvent, à la place du *VSup*, on rencontre une forme verbale neutre (c'est-à-dire peu recherchée au point de vue sémantique), dite *VSup standard* (ou *verbe générique*), comme *faire* (J. Giry-Schneider, 1978), *avoir*, *prendre*, *perdre* (R. Vivès, 1983), *être* + *Prép* (L. Danlos, 1980) ou une paire de verbes converses du type *donner*, *recevoir* (G. Gross, 1987), c'est autour de ces formes-là que l'on a regroupé (en classes disjointes) les prédicats nominaux de la langue française. Dans le cas des *VSup* plus recherchés que les standards, on parlait des *variantes stylistiques* et des *variantes aspectuelles* (à noter pourtant que ces emplois ont été décrits avec moins de rigueur).

Pour ce qui concerne comportement syntaxique des prédicats nominaux du polonais, nous pouvons constater qu'il n'y a pas d'équivalence directe avec les classes du français. Premièrement, les verbes polonais qui peuvent être acceptés par les prédicats nominaux sont très nombreux. E. Jędrzejko (1998 : 50) observe à ce sujet que chaque verbe peut jouer le rôle du *VSup* (pour un de ses emplois). Souvent, les formes employées sont en plus très recherchées stylistiquement. Enfin, par le jeu de préfixes et de suffixes, elles expriment à la fois l'aspect grammatical (perfectif / imperfectif) et le mode d'action (aussi bien au niveau des infinitifs qu'au niveau des formes personnelles). Tous ces facteurs rendent la classification des prédiats nominaux du polonais compliquée (pourtant non impossible)¹.

Les difficultés classificatoires ne devraient pas pourtant décourager de rechercher une organisation dans l'ensemble du lexique (dans notre cas, dans

¹ Au lieu de tenter une classification de tous les prédicats nominaux autour des *VSup*, il semble qu'il serait plus utile de traiter chaque collocation verbo-nominale en tant que signe linguistique indépendant, une unité autonome (bien que composée) ayant son domaine d'emploi, le type de discours dans lequel elle s'adapte le mieux etc. C'est aussi une raison pour laquelle un tel syntagme devrait constituer une entrée de dictionnaire à part (voir aussi G. Gross, 1989).

l'ensemble des collocations verbo-nominales). Si l'on arrive à obtenir des résultats concrets dans la description, à classer les unités selon des critères bien fondés, ceci peut s'avérer ensuite utile (voire nécessaire) pour certaines applications. On pourrait penser, par exemple, à la recherche automatique des données linguistiques qui ne peut être efficace que si la description du lexique est précise et transparente.

Dans le cas de l'étude systématique sur les prédicats nominaux du polonais, dès le début, s'imposait un classement des unités analysées, mais selon d'autres critères qu'en français (G. Vetulani, 2000, 2003, à paraître). Ici, observons seulement que ces prédicats ne se regroupent pas en grands ensembles autour d'une forme verbale (comme p.ex. autour du *V_{Sup} faire* en français), mais qu'ils acceptent le plus souvent une forme propre de toute une gamme de formes sémantiquement apparentées comportant dans leur sémantisme l'idée de 'faire'; comparons ainsi: *robić* – *zrobić* (*pranie, porządek, psikusa...*), *wykonać* – *wykonywać* (*ćwiczenie, zadanie, wyrok...*), *dokonać* – *dokonywać* (*aktu jakiegoś, kradzieży...*), *uprawiać* (*sport, sztukę, malarstwo...*), *przeprowadzić* – *przeprowadzać* (*analizę, badania, pomiary...*), etc. Dans les exemples mentionnés en dessus, nous avons encore affaire aux verbes relativement neutres (peu recherchés par rapport au verbe le plus neutre *robić / zrobić*), mais le plus souvent, on observe à côté du prédicat donné un verbe spécifique (inattendu et difficilement prévisible) qui lui est propre, comme *uciać* pour le prédicat *drzemka*, cf. *uciać sobie drzemkę* et non pas: **zrobić sobie drzemkę*. Par conséquent, si l'on avait gardé pour le polonais les mêmes principes méthodologiques de classement que pour le français, on aurait obtenu beaucoup plus de classes comportant parfois très peu d'éléments-prédicats et un certain nombre de classes à un prédicat seulement. Un tel travail aurait exigé par ailleurs beaucoup plus de temps et un corpus d'analyse plus important et présentant des emplois de tous les registres linguistiques².

En fin de compte, nous avons désigné 5 classes de prédicats selon les critères syntactico-sémantiques, en reconnaissant qu'une description plus détaillée est tout à fait possible (et même recommandée). La Classe I, la plus grande et la plus variée, contient les substantifs désignant divers types d'activités et de comportements (*abordaż, bluźnierstwo, fiasko, wpływ, zaszczyt, zysk...*). Les autres classes sont plus homogènes et représentent respectivement: Classe II – noms désignant des traits de caractère (*agresja, bezmyślność, chępliwość...*), Classe III – noms désignant des maladies (*anoreksja, bezsenność, grypa...*), Classe IV – noms désignant des professions (*akwizycja, dendrologia, domokrąstwo...*), Classe V – noms qui entrent dans

² Nous avons surtout travaillé sur les données comprises dans un dictionnaire de la langue polonaise de taille moyenne (environ: 80 000 entrées, 40 000 noms de tout type, 8 000 noms abstraits) illustrant principalement des emplois normalisés.

les constructions avec un *VSup événementiel* du type: *avoir lieu, il y a*, etc. (*napływ, narodziny, poszum...*) (G. V e t u l a n i, 2000). Notons ici qu'une même forme prédicative peut être classifiée dans plus d'un ensemble parce qu'elle peut avoir plusieurs sens. Par exemple, le prédicat *dowcip* reçoit une fois le sens de 'plaisanterie' (dans l'emploi avec le *VSup zrobić*, cf. *zrobić komuś dowcip*) et se retrouve dans ce cas dans la Classe I ou, encore, il renvoie à un trait de caractère et se retrouve alors dans la Classe II (cf. *mieć cięty (ostrzy, ciężki) dowcip*).

3. Possibilités de classement plus détaillé

L'analyse des collocations à l'intérieur des grandes classes conduit à la découverte des classes plus spécifiques, entre autres celles qui correspondent à divers comportements et activités de l'homme (opérations, occupations, métiers, actions, pratiques, techniques, méthodes, états, idées, etc.). La reconnaissance de ces sous-classes s'effectue en premier lieu, grâce au *VSup* (ou, dans certains cas, toute une structure jouant le rôle du *VSup*) accompagnant le nom qui désigne l'activité. En ce qui concerne les noms, il est possible de distinguer des prédicats-classificateurs étant des qualificatifs (appellations) des ensembles et des prédicats-éléments des ensembles étant des noms spécifiques, c'est-à-dire des représentants des classes. Ainsi, on pourrait désigner en polonais, par exemple, une classe du genre <STANY UCZUCIOWE> (<SENTIMENTS>) qui comporterait les noms des états (ou des sentiments) particuliers comme: *złość* (colère), *wściekłość* (rage), *strach* (peur), etc., ou une classe <RELIGIE> (<RELIGIONS>) dans laquelle seraient regroupés les noms de renvoyant à des religions spécifiques comme: *religia chrześcijańska*, *religia mojżeszowa*, *judaizm*, etc., ou encore une classe <SPORT> comportant entre autres: *narciarstwo* (ski), *szermierka* (escrime), *jeździectwo* (équitation), et ainsi de suite. Beaucoup d'autres classes sont à envisager; p.ex.: <CECHY CHARAKTERU> (<TRAITS DE CARACTERE>), <GŁOSY> (<VOIX>), <KIERUNKI FILOZOFICZNE> (<COURANTS PHILOSOPHIQUES>), <ZAWODY> (<METIERS>), <CHOROBY> (<MALADIES>).

À l'intérieur des classes ainsi désignées, d'autres regroupements sont encore à envisager selon, toujours, les critères de nature lexico-syntaxique, principalement selon le choix de l'élément cooccurent dans la structure d'emploi du prédicat, l'élément qui lui est le plus proche, à savoir le *VSup*. À titre d'exemple, analysons encore une fois, le nom prédictif *dowcip*, en laissant cette fois-ci de côté la signification de 'trait de caractère' (qui fait que le

prédicat se retrouve dans la Classe II) et prenons en considération d'autres acceptions seulement. À côté du sens 'plaisanterie' (quand le nom est employé avec le *VSup* *zrobić*), la forme *dowcip* peut recevoir aussi le sens de 'blague', mais elle doit apparaître alors avec l'un des *VSup* de parole: *powiedzieć* ou *opowiedzieć* (en fr. *dire*), cf. *powiedzieć komuś dowcip*, *opowiedzieć komuś dowcip*. À cause de cette polysémie (deux sens différents du prédicat *dowcip* au niveau de la même classe), nous avons prévu deux schémas représentatifs dans le dictionnaire (G. Vetulani, 2000: 156)³. En règle générale, au nombre de sens différents d'un prédicat correspond le même nombre de modèles de sa réalisation dans le discours. À titre d'exemple, comparons en bas 4 emplois différents du prédicat *imię* (en fr. *nom*) de la Classe I pour lequel, dans le dictionnaire (ibidem: 167), nous donnons les informations concernant le nombre de sens de base (il y en a 2) et les variantes stylistiques (pour chaque sens de base) en même temps (*nadać* pour *dać* et *nosić* pour *mieć*):

- imię*: 1. *dać* (*imię*) / *N1* (*komuś*) / *MOD* (*jakieś*)
 nadać (*imię*) / *N1* (*komuś*) / *MOD* (*jakieś*)
 2. *mieć* (*imię*) / *MOD* (*jakieś*)
 nosić (*imię*) / *MOD* (*jakieś*)

Bien que, à l'étape actuelle des études, nous rendions compte de la polysémie des prédicats nominaux, nous devons reconnaître que des descriptions plus poussées et rigoureuses restent encore à faire. Des classements plus détaillés pourraient en effet contribuer à une meilleure compréhension de l'organisation du lexique autour des prédiats.

4. Vers un classement des activités selon les principes linguistiques

La classification des prédicats se rapportant aux activités (comportements, occupations, pratiques, actions, etc.) de l'homme pose beaucoup de problèmes pour plusieurs raisons. Premièrement, les difficultés sont liées au fait que les contours des activités ne sont pas toujours précis. Nous ne pouvons pas savoir d'emblée, s'il s'agit, par exemple, d'une activité professionnelle (métier exercé) ou d'une activité non professionnelle, d'un travail pour lequel on est rémunéré ou, seulement, d'un travail (occupation) à un moment donné, d'une pratique

³ En réalité, il y en a plus (mais deux sens de base) parce que le prédicat *dowcip* peut apparaître avec certaines variantes stylistiques et aspectuelles; cf. *zrobić dowcip*, *robić dowcipy*, *pozwolić sobie na dowcip*, *powiedzieć dowcip*, *sypać dowcipami* (G. Vetulani, 2000: 156).

habituelle ou temporaire. Pour beaucoup de cas, il est aussi difficile de décider s'il s'agit d'une idée (générale, passagère) de l'homme ou, encore, s'il s'agit d'un courant philosophique, scientifique, artistique, etc. qu'il pratique. Deuxièmement, la langue évolue avec la société qui l'emploie. Il faut admettre alors que les classes constituées une fois vont subir une évolution. La réalité sociale change vite, l'homme invente de nouvelles activités qui grâce à la structure dynamique de la langue s'y répercutent parallèlement. Toutes ces raisons font qu'il est souvent difficile de décider comment ranger une unité analysée.

Sans proposer ici de classements définitifs, remarquons pourtant que, en ce qui concerne les rapports entre les collocations verbo-nominales et les activités exercées par l'homme, en polonais, il est possible de prévoir un regroupement selon des verbes spécifiques. Analysons la liste suivante (entre <...> nous mettons les noms des classes) d'activités :

1. <religie>, <poglądy> *ktoś wyznaje*
<religions>, <idées>
 2. <urzędy>, <stanowiska> *ktoś pełni / piastuje / otrzymuje / przyjmuje*
<postes>
 3. <głosy> *ktoś wydaje*
<voix>
 4. <choroby> *ktoś ma / choruje na / cierpi na*
<maladies>
 5. <stany> *ktoś popada w / pogrąża się w / jest w*
<états>
 6. <zawody> *ktoś się trudni / zajmuje*
<professions>
 7. <sporty> *ktoś uprawia*
<sports>
 8. <analizy> *ktoś przeprowadza*
<analyses>
 9. <zabiegi> *ktoś stosuje / ktoś się poddaje*
<opérations>
 10. <imprezy>, <gry>, <zabawy> *ktoś bierze udział w*
<fêtes>, <jeux>
- etc.

Il est à noter que nous pouvons observer plusieurs verbes dans une classe. Ceci dit, il est encore souhaitable de réfléchir sur les rapports entre un verbe donné et les éléments particuliers de la classe en question. Par exemple dans la classe des <maladies>, nous avons mis 3 verbes différents, mais il s'avère qu'autour des *VSup mieć* et *chorować na* peuvent se regrouper les mêmes éléments (les mêmes noms de maladies), cf. *mieć grypę / anginę /*

białaczkę, koklusz... et *chorować na grypę / anginę / białaczkę / koklusz...*), comme s'il était indifférent quel verbe accompagne ces noms. Par contre, le *VSup* *cierpieć na* (moins neutre parce qu'il marque l'intensité du procès) apparaît avec d'autres prédicats (noms des maladies plus graves ou durables), cf. *cierpieć na schizofrenię / bezsenność*, etc. Parfois, un même *VSup* est employé avec les prédicats des classes différentes (cf. les exemples 10 et 1 de la liste).

L'observation faite, il paraît légitime de se poser en général des questions sur le phénomène de corrélation entre les *VSup* et les classes ayant une motivation sémantique, de même que sur la façon d'en tirer des profits. En ce qui concerne le premier problème, la réponse est facile: les *VSup* utilisables dans les collocations constituent la base d'un classement formel. Ils peuvent être considérés comme des indices formels des phénomènes sémantiques. Quel usage peut-on en faire?

5. Vers une conclusion: ontologies linguistiques

Les considérations qui nous ont amenée ici s'inscrivent bien dans le cadre des recherches sur les systèmes de représentation des connaissances dits *ontologies*. Le terme-même est connu depuis Aristote et il signifie la 'connaissance de ce qui est', 'la façon d'exister'. Dans son usage contemporain, une ontologie est un système de notions (concepts) et de relations formelles entre elles, souvent sous forme de hiérarchie (N. Guarino, 1997)⁴. Un exemple classique d'ontologie est la systématique du naturaliste suédois Linné qui, au XVIII^e siècle, a donné une classification des plantes. Aujourd'hui, les ontologies relatives aux domaines d'applications informatiques sont exploitées dans des systèmes qui traitent l'information (p.ex., en fournissant des mots clés et des algorithmes de recherches par mots-clés). Parmi les systèmes ontologiques existant déjà, il y en a qui ont des motivations linguistiques explicites tenant compte des propriétés des mots avec, en tête, *WordNet* et *EuroWordNet* (ce dernier étant multilingue). Ces systèmes sont constitués selon des critères concrets (voir pour le français: G. Gross, 2002 et tous les travaux de cet auteur sur les *classes d'objets*, pour le polonais: K. Polański, 1980; Z. Vetulani, 2003, 2004).

Le classement des prédicats nominaux par rapport aux *VSup* est un exemple d'application d'un procédé linguistique pour une ontologie des

⁴ Cf. „[...] *ontologies*, i.e. theories of various kinds expressing the meaning of shared vocabularies, in the specifying field of information retrieval and extraction as well as in the more general area of knowledge and language engineering” (N. Guarino, 1997: 140).

prédicats. Le fait que les phénomènes sociaux se reflètent au niveau formel de la langue peut être exploité dans la technologie du langage.

Références

- Danlos L., 1980: *Représentation d'informations linguistiques: constructions N être Prép X*. Thèse de 3 cycle, L.A.D.L., Université Paris VII.
- Giry-Schneider J., 1978: *Les nominalisations en français: l'opérateur «faire» dans le lexique*. Genève-Paris: Librairie DROZ.
- Gross G., 1987: *Les constructions converses en français*. Genève-Paris: Librairie DROZ.
- Gross G., 1989: «Le dictionnaire et l'enseignement de la langue maternelle». In: *Wörterbücher. Dictionaries. Dictionnaires*. Berlin-New York, 174-180.
- Gross G., 2002: «Recherches théoriques et enseignement des langues». In: *Points communs: linguistique, traductologie, glottodidactique*. Łódź, 88-101.
- Gross G., Vivès R., 1986: «Syntaxe des noms». *Langue Française*, 69.
- Gross M., 1981: «Les bases empiriques de la notion de prédicat sémantique». *Langages*, 63, 7-52.
- Guarino N., 1997: «Semantic Matching: Formal Ontological Distinction for Information Organization, Extraction, and Iteration». In: *Information Extraction. A Multidisciplinary Approach to an Emerging Information technology*. Springer.
- Jędrzejko E., red., 1998: *Słownik polskich zwrotów werbo-nominalnych. Zeszyt próbny*. Warszawa: Energeia.
- Jędrzejko E., 2002: *Problemy predykcji peryfrastycznej. Konstrukcje, znaki, pojęcia*. Katowice: Gnome.
- Polański K., red., 1980: *Syntaktyczno-generatywny słownik czasowników polskich*. Wrocław-Warszawa-Kraków-Gdańsk: Ossolineum.
- Vetulani G., 1994: «Constructions à VSup + NPred et l'aspect. Étude confrontative: français – polonais». In: *Akten des 28. Linguistischen Kolloquiums. Graz 1993*. Tübingen: Max Niemeyer Verlag.
- Vetulani G., 2000: *Rzeczowniki predykatywne języka polskiego. W kierunku syntaktycznego słownika rzeczowników predykatywnych na tle porównawczym*. Poznań: UAM.
- Vetulani G., 2003: «Les collocations verbo-nominales et la traduction». *Studia Romanica Posnaniensia*, 30, 159-168.
- Vetulani G., à paraître: «Le rôle du verbe dans le réseau dérivationnel des prédicats nominaux». *Studia Romanica Posnaniensia*.
- Vetulani Z., 2003: «Linguistically Motivated Ontological Systems». In: N. Callaos, W. Lesso, K.-D. Schewe, E. Atlam, eds.: *Proceedings of the 7th World Multiconference on Systemics, Cybernetics and Informatics, July 27-30, 2003, Orlando, Florida, USA*. Vol. 12: *Information Systems, Technologies and Applications: II*. Int. Inst. of Informatics and Systemics, 395-400.
- Vetulani Z., à paraître: *Komputerowe modelowanie kompetencji językowej*. Warszawa: EXIT.
- Vivès R., 1983: *Avoir, prendre, perdre: constructions à verbe support et extension aspectuelle*. Thèse de 3 cycle, L.A.D.L., Université Paris-VIII.