

Uniwersytet Mikołaja Kopernika

**STUDIA
ROMANICA
ET
LINGUISTICA
THORUNENSIA**

(publiés par Witold Stefański)

Toruń 1998

2

COMITÉ DE RÉDACTION

Leszek Bednarczuk (Kraków)
Witold Mańczak (Kraków)
Karl Horst Schmidt (Bonn)
Christian Schmitt (Bonn)
Józef Sypnicki (Poznań)

RECENZENCI

Leszek Bednarczuk
Wiesław Malinowski

PROJEKT OKŁADKI

Alicja Majewska

REDAKTOR

Jadwiga Gosieniecka

KOREKTOR

Iwona Wakarecy

ISSN 0860–1232

ISSN 1234–8945

Printed in Poland

© Copyright by Wydawnictwo Uniwersytetu Mikołaja Kopernika
Toruń 1998

Wydanie I. Nakład 250 egz. Ark. wyd. 13,5 Ark. druk. 13.

Skład: Dział Wydawnictw UMK

Druk: Zakład Poligraficzno-Wydawniczy POZKAL, Inowrocław, ul. Ceglana 10/12

ALBERTO GIL, NICO WEBER
Köln

Courants récents en théorie linguistique et leur application à la sémantique lexicale et phrasale

Avec une étude empirique sur un corpus du portugais contemporain

Point de départ

La sémantique comme étude du sens des signes linguistiques a été traditionnellement subdivisée en deux domaines majeurs:

- 1° La sémantique „lexicale” ayant pour objet la structure interne des mots en tant qu'unités signifiantes et les relations de sens établies entre eux en vertu de leurs propriétés "internes", ainsi que de principes d'organisation "externes" du lexique.
- 2° La sémantique "phrasale" portant sur l'analyse des propositions, de leur structure interne ou composition, de leur enchaînement et de leurs modifications (interrogation, négation, modalités diverses).

L'opposition méthodique entre ces deux points de vue est formulée, pour ne citer que deux exemples, par Johnson-Laird 1982 et Kay 1992. Selon Johnson-Laird 1982 la première, qui serait du domaine psychologique, doit livrer une interprétation des unités de base d'une langue en rapport avec des "modèles mentaux" (*mental models*) du monde extérieur. La seconde, qu'il appelle sémantique formelle ou structurale, et qui serait tributaire surtout de la logique intensionnelle, modale et temporelle, porte sur l'interprétation compositionnelle d'expressions complexes formées, entre autres, à l'aide d'unités lexicales de base. Le but du travail cité est la réconciliation des deux points de vue. De la même façon Kay 1992 distingue deux "traditions majeures" au sein de la sémantique. D'un côté, l'étude des champs lexicaux ou, dans une approche

plus générale, la sémantique lexicale, surtout orientées vers les relations paradigmatiques ou "de contraste" entre unités du lexique (sur la base de leur structure interne), ainsi que vers leurs rapports avec les "objets non-linguistiques" qu'elles symbolisent¹. De l'autre, la sémantique formelle (d'orientation logique) et la "sémantique combinatoire" (d'orientation structurale) étudiant les relations syntagmatiques entre lexèmes considérés comme unités de base non-analysables, en vue d'un calcul compositionnel². L'opposition suggérée par les deux auteurs, comme par beaucoup d'autres, est donc la suivante:

Sémantique lexicale	↔	Sémantique phrasale
≈ Paradigmatique		≈ Syntagmatique
≈ Contrastive		≈ Combinatoire
≈ Psychologique		≈ Logique
≈ Modèles mentaux		≈ Modèles formels

La conséquence en est que la "sémantique linguistique" (par opposition à "logique") traite surtout des relations paradigmatiques entre unités lexicales. Mais en consultant les traités correspondants (Cruse 1986³ par exemple), il est difficile, voire impossible, d'évincer l'impression que les quelques relations et régularités de forme ou de sens plus ou moins systématiques qu'on peut dégager de l'étude du lexique ne sont pas à la mesure de la multiplicité et de la complexité des phénomènes conditionnant le fonctionnement sémantique des langues⁴. Étant donné l'universalité et l'importance linguistique primordiale des phénomènes syntaxiques, (considérés, dans une perspective typologique, comme un

¹ "Students concerned with lexical fields and lexical domains ("lexical semanticists") have interested themselves in the paradigmatic relations of contrast that obtain among related lexical items and the substantive detail of how particular lexical items map to the nonlinguistic objects they stand for" (Kay 1992, 309).

² "Formal semanticists" (those who study the combinatorial properties of word meanings) have been mostly unconcerned with these issues, concentrating rather on how the meanings of individual words, whatever their internal structure may be and however they may be paradigmatically related to one another, combine into the meanings of phrases and sentences (and recently, to some extent, texts). Combinatorial semanticists have naturally been more concerned with syntax, especially as the leading idea of formal semantics has been the specific combinatorial hypothesis of Fregean compositionality" (loc. cit.).

³ Cet ouvrage sur la sémantique lexicale (avant tout mais pas exclusivement) paradigmatique fait fonction d'autorité dans le domaine. Outre un traitement profond et souvent original des phénomènes en question on y trouve un grand nombre de créations terminologiques valables.

⁴ "Linguistics has tended in the past to overemphasize substitution (paradigmatic) relations, and to undervalue cooccurrence (syntagmatic) relations. Paradigmatic relations

continuum entre la "syntaxe des mots", ou morphologie, et la construction des phrases et de la cohésion transphrastique (Morris/Hirst 1991)), il semble peu naturel d'écarter de ces réflexions les relations consécutives, ou de les remettre au second plan. L'encodage et le décodage syntaxiques sont des faits tout à fait essentiels dans le processus de génération et de compréhension, puisque, dans des circonstances normales, personne ne parle en mots isolés ou construit son discours consciemment et principalement sur des rapports paradigmatiques entre unités lexicales.

En fait le biais, après avoir été reconnu, est en train d'être redressé sous l'influence simultanée de plusieurs courants récents en théorie linguistique. L'essor du traitement automatique des langues (*computational linguistics*) a conduit au développement de ce qu'on appelle la "linguistique des corpus" (*corpus linguistics*). En théorie de la grammaire (*theory of grammar*) récente il est question de la "grammaire de construction" (*construction grammar* / = CG), que nous nous proposons d'esquisser dans la première partie de cet article, suite à quelques remarques sur l'analyse assistée par ordinateurs des grands corpus de textes et sur l'approche traditionnelle à la sémantique lexicale. Dans la deuxième partie nous discuterons quelques résultats de l'évaluation d'un corpus électronique du portugais contemporain. Il s'agira de concrétiser, pour quelques verbes choisis comme exemples et leurs valences, la notion de "construction", en fonction des paramètres déterminant les structures actanciellles et leurs sérialisations superficielles tels que: dépendances relationnelles (régularités de valence), déterminants pragmatiques (projection iconique de l'enchaînement événementiel en réalité) et aspects fonctionnels de la textualisation (facteurs communicatifs, discursifs et situationnels).

Principes d'analyses sémantiques empiriques

L'exploitation de larges collections de textes, couramment appelées "corpus (de textes)", assistée par ordinateurs, a connu un essor rapide depuis une dizaine d'années à peu près (Church/Mercer 1993). Cela est dû à trois facteurs combinés: 1° Le développement des outils informatiques, mémoire et systèmes à haute capacité du côté *hardware*, logiciels de support et de gestion intelligents du côté *software*; 2° la chute, parallèle, des prix, surtout pour le matériel; 3° la compilation et la mise à la disposition de la communauté

are of course important for acquiring word class, but are a somewhat blunt-edged tool for further discrimination" (Aitchison/Lewis 1995, 17).

scientifique de collections de plus en plus larges et de mieux en mieux réfléchies de textes représentant un nombre croissant de langues⁵. Malheureusement il n'est pas trop sûr que le développement des capacités informatiques ait entraîné un accroissement équivalent des ressources linguistiques. L'objectif général de ces efforts est l'acquisition d'informations linguistiques, notamment grammaticales et sémantiques, sur les unités lexicales et supralexicales, à travers l'identification et l'évaluation d'un ensemble vaste et représentatif de *cooccurrences* (terme neutre), *collocations* (accentuation syntaxique) ou *associations* (accentuation sémantique)⁶. Pour citer quelques exemples: La reconnaissance de collocations non-aléatoires, récurrentes parmi les cooccurrences d'un mot peut servir à en distinguer les sens divers d'une façon précise (Biber 1993b). L'analyse des cooccurrences d'un verbe dans un corpus de textes permet d'identifier ses collocations syntaxiques régulières et non-erronées qui, à leur tour, peuvent être sémantiquement interprétées comme arguments associés au prédicat représenté par le verbe (Brent 1993). Les collocations d'expressions identifiantes ou relatives peuvent être interprétées comme associations d'expressions explicatives (*explicatum* et *explicans*) dans le cadre d'une sémantique définitoire (Büchel/Weber 1995). Nous concluons ces remarques en disant qu'il reste à établir un fondement théorique plus élaboré pour l'analyse linguistique et l'interprétation sémantique des collocations et associations de mots et de phrases repérés dans les corpus et catalogués moyennant les outils informatiques spécialisés que les chercheurs ont maintenant à leur disposition. Il est concevable que la CG puisse fournir ce fondement.

Sur un plan très général, on peut distinguer deux classes de relations lexicales paradigmatiques, celle des relations *hiérarchiques* et celle des relations *symétriques*. Les relations hiérarchiques sont asymétriques et orientées, l'orientation étant partiellement marquée sur le plan terminologique. Elle est bidirectionnelle dans le cas de l'*hyperonymie* ou *superordination* et de l'*hyponymie* ou *subordination* marquant la relation entre le générique et le spécifique, sur

⁵ "Over the past few years, improvements in computer technology, coupled with a drop in hardware prices, have made it possible to handle ever-larger text corpora. Accompanying the increase in scale of the corpora, the operations which can be performed on them have become more and more complex. These changes have, for example, enabled corpus researchers to progress from looking at a simple frequency-order word list of a corpus to examining the collocational patterns of each type in the list" (Collier 1993, 292).

⁶ Cf., par exemple: Martin/Al/van Sterkenburg 1983; Kjellmer 1987, 1990; Basili et al. 1992; Biber 1993a, 1993b; Brent 1993; Collier 1993; Smajda 1993; Bindi et al. 1994; Cowie et al. 1994; Sinclair (ed.) 1987; Souter/Atwell (eds.) 1993; Oostdijk/de Haan (eds.) 1994.

laquelle sont fondés les systèmes *taxinomiques*. La *méronymie* ou *partonymie* indique la relation entre un tout et ses parties sans que l'orientation soit clairement marquée (en anglais: *part-whole relation*). Les relations symétriques „classiques” sont la *synonymie* et l'*antonymie*. On peut admettre qu'en principe les deux termes marquent une seule relation de base accentuée de façons différentes. La synonymie peut ainsi être caractérisée comme mise en évidence de similarités ou absence de mise en évidence de différences de sens entre unités lexicales⁷, l'antonymie, inversement, comme mise en évidence d'un contraste sémantique sur la base de similarités plus ou moins prononcées⁸. Il est intéressant de noter un parallélisme entre les classes de relations paradigmatiques hiérarchique et symétrique d'une part et les classes de relations syntagmatiques hypotaxique et parataxique de l'autre. Une autre remarque à faire est que les relations lexicales hiérarchiques sont plutôt fondées sur un ordre extra-linguistique des choses tandis que les relations symétriques reflètent des rapports linguistiques entre unités lexicales.

Les relations paradigmatiques aussi bien que syntagmatiques qu'une unité lexicale peut entamer avec d'autres déterminent son „comportement” sémantique et donc la manière dont sa signification peut être acquise, aussi bien que la façon dont elle peut être expliquée ou, si l'on préfère, sa définition. En étudiant les types d'explications ou de définitions qui peuvent être reconnus et distingués (Weber 1996) on constate que chaque type se prête particulièrement à l'explication de certaines classes d'unités lexicales, à l'exclusion d'autres types. Les concepts taxinomiques, par exemple, sont invariablement expliqués à l'aide de la définition dite Aristotélicienne *per genus proximum et differentiam specificam*. La façon dont le sens d'une unité lexicale peut être, ou est le plus souvent décrit, découle évidemment des environnements et configurations par lesquels l'unité est déterminée au sein du système linguistique (à l'inclusion des déterminants et implications extralinguistiques). Il serait donc intéressant d'essayer d'établir une typologie lexicale à partir des explications de mots (définitions) dans les dictionnaires et autres textes „explicatifs”. Ceci dit en passant pour indiquer que les textes lexicographiques représentent un type de

⁷ "Synonyms must not only manifest a high degree of semantic overlap, they must also have a low degree of implicit contrastiveness" (Cruse 1986, 266). "Synonyms, then, are lexical items whose senses are identical in respect of 'central' semantic traits, but differ, if at all, only in respect of what we may provisionally describe as 'minor' or 'peripheral' traits" (op. cit. 267).

⁸ "The paradox of simultaneous difference and similarity is partly resolved by the fact that opposites typically differ along one dimension of meaning: in respect of all other features they are identical, hence their semantic closeness" (op. cit. 197).

corpus particulièrement intéressant, ce qui d'ailleurs n'a pas échappé à la recherche en lexicologie, sémantique et traitement automatique des langues (Weber 1995 ms). D'ailleurs les dictionnaires depuis toujours offrent, avec les exemples d'emploi, les citations, les locutions, les indications d'usage (registre, style, spécialisation) un nombre d'informations qui, en CG, seraient à qualifier comme éléments de la „construction” du lexème. Nous essayerons maintenant d'esquisser cette idée.

Rudiments de la "Construction Grammar"

La CG se propose de fournir le cadre théorique et technique pour une intégration des deux approches à la sémantique — analyse contrastive du lexique et étude des combinaisons syntaxiques⁹. Il s'agit d'établir, d'une façon empirique et méthodique, les relations systématiques entre formes et significations linguistiques. Ces relations sont manifestées par les "associations conventionnelles" ("conventional associations", Kay 1992) d'informations lexicales, syntaxiques, sémantiques et pragmatiques. Une construction grammaticale, selon Fillmore 1988, est un schéma (*pattern*) syntaxique remplissant une (ou plusieurs) „fonction(s) conventionnelle(s)” dans une langue. Ce qui signifie qu'il s'agit d'une unité identifiable du code linguistique réunissant diverses propriétés typiques pour le code en question, c'est-à-dire des propriétés combinatoires (syntaxiques) mais aussi symboliques et conventionnelles (sémantiques et pragmatiques)¹⁰. Pour ce qui est des propriétés syntaxiques on peut distinguer les propriétés exocentriques (que Fillmore appelle "external") et endocentriques ("internal"). Les propriétés "internes" sont les caractéristiques structurales exhibées par la construction elle-même, tandis que les propriétés "externes" déterminent ce qui lui permet de s'associer à ou s'intégrer dans d'autres constructions, coordonnées ou superordonnées. La grammaire est l'ensemble des constructions d'une langue ainsi que des principes généraux

⁹ Nous suivons ici les réflexions dans Fillmore 1988, Fillmore/Kay/O'Connor 1988, Kay 1992. Nous n'avons pas l'intention de discuter les détails techniques de la CG (cf. Fillmore 1988, Fillmore/Kay 1995). Notre objectif est plutôt d'esquisser ses principes généraux et d'estimer leur applicabilité en référence aux approches empiriques décrites dans la seconde partie de cet article.

¹⁰ "By *grammatical construction* we mean any syntactic pattern which is assigned one or more conventional functions in a language, together with whatever is linguistically conventionalized about its contribution to the meaning or the use of structures containing it" (Fillmore 1988, 36).

gouvernant leur enchaînement, leur imbrication et leur "superimposition" (Fillmore 1988, 37). Les constructions grammaticales sont décrites en termes de rôles (positions "à remplir") et de catégories (éléments ou classes d'éléments "remplissants"). Les catégories fonctionnelles peuvent être "minimales" (lexicales) ou "maximales" (phrasales complètes), des manifestations intermédiaires ou indéterminées étant prévues. Les unités lexicales de leur côté sont caractérisées par les constructions dans lesquelles elles peuvent entrer. Du point de vue grammatical (donc morpho-syntaxique) c'est de cette façon qu'est déterminée leur appartenance à certaines catégories (classes lexicales) qui à leur tour peuvent se qualifier pour certains rôles. C'est là que se trouve la jointure entre la grammaire et le lexique. Il ne faut pas oublier toutefois qu'aux propriétés "catégorielles" on doit ajouter les propriétés sémantiques (et pragmatiques) d'une unité pour parvenir à une description adéquate de ses propriétés "constructionnelles". Les généralisations syntaxo-sémantiques peuvent être exprimées en termes de "valences"¹¹.

En fait, les *constructions* de la CG sont semblables aux *règles* de la grammaire traditionnelle. La grammaire traditionnelle décrit des classes de formes linguistiques et attribue certaines catégories sémantiques à ces classes¹². Mais, ces catégories sont aprioriques et "absolues", comme le remarquait déjà Gross 1975, plaidant pour une "approche empirique à la sémantique largement basée sur la syntaxe"¹³. En outre, les règles grammaticales traditionnelles s'appliquent à des catégories lexicales très générales, comme les classes de mots¹⁴, ou non sémantiquement motivées (paradigmes flexionnels, "slots" syntaxiques). Pour la CG les associations formelles et sémantiques d'une unité lexicale ou d'un groupe de lexèmes ayant des associations semblables sont des constructions qui doivent être apprises séparément et reconnues en com-

¹¹ "The lexicon, which in important ways is not distinct from the repertory of constructions, associates with each lexical items, explicitly or implicitly, information about the grammatical constructions in which the item can participate. To the extent that a given lexical item is closely tied to one or more specific grammatical constructions, describing that item is equivalent to describing the constructions in which it participates. [...] In those cases in which generalizations about lexical items can be made without reference to particular constructions, the combinatorial properties of lexical items can be stated as their *valence descriptions*" (op. cit. 42).

¹² "Traditional grammars classify forms into families, and attribute to these families absolute categories of meanings" (Gross 1975, 389).

¹³ "Thus, we seem to be advocating here an empirical approach to semantics that is largely based on syntax" (op. cit. 403).

¹⁴ Pour une discussion concise et critique, du point de vue typologique, de la catégorisation lexicale traditionnelle, cf. Sasse 1993.

munication par les usagers d'une langue. Cela implique — autre différence importante — que les expressions non-compositionnelles, notamment idiomatiques ne sont pas considérées comme des exceptions mais comme constructions de plein droit. D'après Kay 1992 une construction en CG peut être un lexème ou un idiome singulier, une collocation déterminée par une règle "mineure", un "modèle lexico-grammatical partiellement productif" ou une règle (grammaticale) "générale"¹⁵. Déclarative en général, la CG se sert de l'héritage (*inheritance*) comme élément procédural pour rendre compte (et tirer avantage) des relations d'inclusion de traits et de propriétés découlant de l'appartenance simultanée d'unités linguistiques à des constructions plus générales (catégories ou règles grammaticales) et plus particulières (sous-classes syntaxiques, sémantiques, pragmatiques).

Ce qui distingue la CG des grammaires transformationnelles est surtout le fait qu'elle ne prévoit pas de transformations. Dans le paradigme générativiste la CG a adopté certaines conceptions, en abandonnant toutefois la notation, de la théorie X-barre. La notation en CG (Fillmore 1988, Fillmore/Kay 1995) s'inspire de celle de la grammaire syntagmatique généralisée (*generalized phrase structure grammar / GPSG*). En outre, la CG fait usage de diverses conceptions et techniques bien connues comme la représentation en structure de traits (*feature structure representation*), les paires attribut-valeur (*attribute-value pairs*), structures à prédicat-argument (*predicate-argument structures*), rôles sémantiques et relations actancielles, approche *slot-and-filler*, unification et héritage¹⁶. La double fonction des unités lexicales comme "objets morpho-syntaxiques" et représentants de "schémas conceptuels" assure la connexion

¹⁵ "Every such conventional association that must be learned or recognized separately by the speaker of a language is a construction. This includes all idioms and partially productive lexico-grammatical patterns" (Kay 1992, 310); "although sometimes the conventions that tie a certain semantic form to a corresponding syntactic form are general, these conventions can also at times be highly specific. Insofar as the conventions are general we tend to think of them as productive rules, while the narrower the conventions associating meaning to form become the more likely we are to think of them as minor rules, collocations, idioms, or the like" (op. cit. 314, note 8).

¹⁶ "Summarizing, we treat grammatical constructions as syntactic patterns which can fit into each other, impose conditions on each other, and inherit properties from each other. Grammatical constructions define positions which require or welcome fillers with certain properties, and fillers of those positions can introduce constructions of their own and can impose requirements of their own on positions within the constructions which contain them. At least some aspects of the grammar operate on simple principles of unification, augmented by principles of inheritance and principles for checking the presence of obligatory elements" (op. cit. 54).

entre la grammaire et le lexique ou, si l'on veut, entre l'expression et le contenu du langage¹⁷.

La CG se veut une grammaire, c'est-à-dire une théorie du langage, psychologiquement réelle. En effet, il ya lieu d'admettre, dans une plus large mesure qu'on n'avait tendance à le faire traditionnellement, que l'acquisition lexicale, à l'inclusion des relations paradigmatiques, chez les enfants aussi bien que chez les adultes se fait à travers l'apprentissage de cooccurrences (Aitchison/Lewis 1995). Les usagers d'une langue auraient alors à leur disposition des *généralisations individuelles* émanant de la représentation internalisée des données observées (Kay 1992). Ces généralisations sont déductibles de la récurrence des phénomènes consécutifs de surface, ou structures lexicales syntagmatiques; par comparaison on arrive à l'observation de contrastes également récurrents, donc aux structures lexicales classées comme paradigmatiques. Elles constituent ce que Wierzbicka 1992 appelle la "structure sémantique" d'un mot. Celle ci peut être plus ou moins consciente aux usagers d'une langue qui, par conséquent, peuvent en rendre compte plus ou moins explicitement. Mais en tout cas elle est reflétée par les collocations et environnements et, en général, les contextes linguistiques et pragmatiques dans lesquels le mot en question peut être utilisé¹⁸. L'objectif d'une grammaire serait d'élaborer une *généralisation commune* des constructions dans une langue sur la base de l'étude des textes¹⁹. Il n'est pas difficile, à ce point, de reconnaître la convergence du programme de la CG et des principes de l'analyse et du traitement assistés par ordinateurs des corpus. Nous dirons que la CG représente, ou plus précisément, peut être élaborée de sorte à représenter le fondement théorique des efforts relatifs dans le domaine du traitement automatique des langues.

¹⁷ "The semantic part of a lexical entry in a construction grammar connects the morpho-syntactic object in question to a schema or frame, which represents a conceptual analysis of (the linguistically encoded aspects of) some area of extralinguistic experience – not excluding the experience of using language" (Kay 1992, 309).

¹⁸ "(...) the word's semantic structure. Native speakers may or may not have access to this structure on a conscious level, but they control it on a subconscious level, often they can bring it to the surface of their consciousness by methodological introspection; and in any case, it is reflected in common collocations, metaphors, set phrases, and other direct linguistic and ethnographic evidence" (Wierzbicka 1992, 221).

¹⁹ "A reasonable goal for the grammarian would appear to be to describe every generalization that a speaker *could* abstract from the data of his or her language, leaving to the psycholinguist the hard task of determining which generalizations actually are abstracted by which speakers. We could think of this as the minimal, psychologically real grammar of the language" (Kay 1992, 329).

Valences et constructions

L'analyse des valences verbales est essentiellement liée à la notion de "construction" (Fillmore 1988, 42 — cf. note 11) étant donné que les structures syntaxiques et sémantiques se déterminent mutuellement. Dans notre étude empirique nous allons par conséquent nous concentrer sur une analyse de valences concrète en prêtant particulièrement attention à l'ordre linéaire des éléments valenciels dans la phrase — un point qui, dans les études sur la valence présentées jusqu'ici, a été plutôt négligé, malgré son importance pour la dimension communicative de la valence. Nous voulons ainsi arriver à une perspective plus intégrale de la langue qui, d'une part, est représentée par la tradition de la grammaire fonctionnelle (Daneš 1964) mais qui, d'autre part, est aussi de plus en plus exigée dans les études récentes sur la valence (Feuillet 1996).

Les études sur la valence verbale révèlent en plus des caractéristiques importantes de langues particulières. Ce type d'analyses se prête donc particulièrement bien à l'étude comparée des langues. Dans le cadre présent, où nous ne présentons qu'une première étape qui devra être suivie par d'autres travaux concrets, nous allons nous borner à appliquer la théorie à une seule langue. Plusieurs raisons nous ont amenés à choisir le portugais.

1° Le fait que, grâce à l'obligeance du *Centro de Linguística da Universidade de Lisboa* nous avons à notre disposition un corpus électronique de cette langue. 2° Le fait qu'il existe déjà plusieurs études contrastives entre le portugais et d'autres langues (à savoir l'allemand, le français et l'espagnol — Busse/Vilela 1986, Busse 1994, Vilela 1994). 3° Le fait que les travaux les plus récents relatifs au portugais ont fait usage des résultats actuels de la théorie des valences. (Vilela 1995, Andrade Peres/Móia 1995).

Dépendances et ordre des mots

Alors que les structures actantielles comme telles ont toujours été l'objet d'études spécialisées sur la valence, peu d'attention fut prêtée aux structures configurationnelles, c'est à dire l'ordre linéaire des éléments à la surface de la phrase. L'ordre des mots était considéré comme étant déterminé par des paramètres communicatifs et situatifs, bien qu'il n'y ait aucun consensus général sur ces questions. Ainsi, selon Greinadus 1990, 47, „l'ordre des constituants dans la phrase ne dépend pas du verbe constructeur, mais de facteurs discursifs”. Vilela 1995, 225 de son côté est d'avis que la détermination

de l'ordre des mots est un attribut essentiel de la valence. Ceci rapproche le linguiste portugais à la tradition grammaticographique allemande, surtout en ex-RDA (Bondzio 1980, 140f., Welke/Meinhard 1980: 150). Dans la même tradition, les travaux se multiplient qui étudient l'ordre superficiel des mots en partant de structures valenciennes sous-jacentes. Dans ces travaux sont analysés les degrés de cohésion profonde entre les éléments valenciens (compléments et circonstanciels) et les conséquences pour l'ordre linéaire superficiel (Korzen 1996, Gil 1996).

Du point de vue de la CG les questions suivantes se posent:

- 1° Dans quelle mesure la répartition des poids informationnels sur l'ordre linéaire superficiel est-elle déterminée par la valence?
- 2° Dans quelle mesure le verbe est-il le „centre de cristallisation” d'un ordre non marqué d'éléments représentant des configurations informationnelles conventionnalisées?
- 3° Quelle est la relation entre la valence du verbe et les paramètres discursifs respectifs permettant de prédire un certain ordre des mots?

Pour les langues romanes les travaux de P. Koch (1981, 1994, 1995, 1996) ont fourni des impulsions importantes pour la réponse à ces questions. Il part de l'hypothèse que les dimensions syntaxiques, sémantiques et pragmatiques des structures valenciennes sont étroitement imbriquées les unes dans les autres. Si par exemple un actant facultatif devient explicite, son poids informationnel s'accroît et il est prédestiné à devenir plus rhématique (Koch 1981, 94). Nous pouvons ajouter que dans le cas d'un ordre non marqué des mots cet actant occupera une position dans le segment final de la phrase.

Dans ce contexte il est question de la „mise en perspective” comme phénomène relatif à la „valence pragmatique” (Welke 1988/1989, 49). Cela veut dire qu'une signification particulière d'un verbe est chargée d'une perspective particulière de la réalité. Helbig 1992, 47 ff. de son côté distingue différentes perspectives ou orientations parmi les verbes de mouvement, à savoir: SOURCE (ex. *surgir*), PATH (ex. *ambuler*) et GOAL (ex. *arriver*). On peut penser ici à un ordre des mots correspondant à une sérialisation iconique ancrée sur le verbe comme centre de cristallisation.

La multiplicité des facteurs déterminant l'ordre des mots — discursifs *ad hoc* d'un côté et véritablement ou quasi grammaticalisés de l'autre — mène à une vue probabiliste sur la grammaire (Delbecque 1995, 152f.) suite à laquelle les phénomènes linguistiques sont considérés comme paramètres variables dont l'importance peut être calculée par des méthodes quantitatives et notamment statistiques sur la base de corpus de textes.

Taxinomie et méthodes d'analyse

Nous nous proposons maintenant de présenter une analyse empirique et d'en discuter quelques résultats. Nous partons des éléments dont les positions dans la phrase ont été comparées. Évidemment nous avons choisi les segments les plus „mobiles”, i.e. NP, PP et adverbies. Les pronoms et les compléments phrasals à infinitif ou subordonnés ont été exclus, leur position dans la phrase étant ou bien grammaticalisée (ce qui est vrai surtout pour les pronoms) ou bien fondamentalement déterminée par des paramètres situationnels ou discursifs.

Au moins deux éléments sont comparés à chaque fois: soit un complément actantiel et un circonstanciel libre²⁰ soit deux compléments actantiels²¹. Les exemples proviennent du segment „Textes de journaux” du „Corpus of Contemporary Portuguese” (19 600 722 mots courants) au *Centro de Linguística da Universidade de Lisboa*. Dans certains cas ils furent complétés par des exemples émanant du segment „Textes littéraires” (7 644 942 mots courants) puisque, parfois il était difficile de rassembler un nombre significatif d'exemples en ne considérant que les cas des éléments nominaux.

Puisqu'il s'agit d'une étude sur l'ordre des mots, il n'est pas nécessaire de résoudre dans tous les cas le problème difficile de la distinction entre compléments actantiels adverbiaux et circonstanciels libres (cf. Helbig 1992, 72 ff.). De même, la question de la distinction entre valence obligatoire et facultative n'est pas abordée ici. Comme l'ont montré Busse 1974: 128 f. et plus récemment Pittner 1996, 96 ff. le statut obligatoire d'un complément dépend de la situationnalité donnée et de l'aspect focalisé avec elle.

Nous avons adopté la classification des compléments actantiels présentée par Busse/Vilela 1986, 35 ff., résumée ici (les mises en évidence dans les exemples sont les nôtres):

- „Sujeito tradicional, ex. *Uma empresa portuense apresenta computadores novos*.
- Complemento direito tradicional, ex. *Uma empresa portuense apresenta computadores novos*.
- Complemento indireito tradicional, ex. *Uma empresa portuense apresenta computadores novos ao público*.

²⁰ Par exemple, pour le verbe „regressar”, un complément directionnel et un adverbial libre: „regressam ao trabalho na segunda feira” vs. „regressou domingo à noite da Arábia Saudita”.

²¹ Par exemple, pour le verbe „introduzir”, un complément d'objet direct et un directionnel: „introduzi a genética molecular em Portugal” vs. „introduzimos na rede o dicionário 'Internês'-Português”.

– Complément préposicional (com preposição fixa), ex. *Precisamos de ajuda.*

– Actante locativo, ex. *O João vive em Lisboa.*

– Actante directionnel, ex. *Irei para Lisboa.*

– Actante temporel, ex. *O documento data de então.*

– Actante de medida, durée, prix, ex. *A sessão durou três horas.*

– Actante modal, ex. *A moça cheira a rosas.*

– Actante predicatif, ex. *Ela considerou o João chefe.*”

Le cadre présent nous oblige à nous limiter à un des types énoncés. Nous avons choisi les compléments directionnels parce qu’avec ce type nous avons obtenu les résultats les plus clairs, ce qui les qualifiait à servir d’exemples.

Pour répondre aux questions posées concernant la relation entre les configurations valenciennes et les paramètres discursifs déterminant l’ordre des éléments dans la phrase, il convient de se référer aux critères élaborés par des approches actuelles de l’analyse des textes à l’étude de l’ordre des mots. Siewierska 1993, 826 résume les facteurs déterminant les configurations et sérialisations d’éléments de la façon suivante:

- a) "grouping relations (e.g. dependency and/or constituency relations),
- b) grammatical relations (e.g. subject, object, indirect object, etc.),
- c) thematic relations (e.g. topic, focus, theme, etc.),
- d) semantic roles (e.g. agent, patient, recipient, etc.),
- e) syntactic features (e.g. categorial status, internal categorial structure, tense, aspect, modality, mood, finiteness, etc.),
- f) semantic features (e.g. animacy, humanness, definiteness, referentiality, etc.),
- g) pragmatic factors (e.g. perceptions of salience or dominance, familiarity, iconicity, relative identifiability, etc.).

Kindt 1994, 51 ff. souligne le facteur de "l’iconicité", mettant en évidence l’homomorphie existant entre le cours d’une action et sa symbolisation linguistique. Il observe qu’il existe des "ordres standard" tels que: "le proche précède l’éloigné"; "le connu précède l’inconnu"; "le défini précède l’indéfini"; "l’universel précède le particulier" et "l’important précède l’insignifiant". Becker/Gutfleisch-Rieck 1994, 249 f. de leur côté complètent ces sérialisations par des règles de préférences basées sur les principes: "topique précède focus" et "défini précède indéfini". A ces régularités discursives il faut ajouter, selon Korzen 1996, 68 ff., la relation étroite qui existe entre le choix d’un verbe et l’ordre des éléments de la phrase, ce qui nous amène au centre de l’analyse présentée ici. Korzen utilise l’exemple des *adverbiaux scéniques*, qui peuvent être

disloqués en présence de certains verbes seulement. Ainsi, avec *se garer* il est possible de disloquer l'adverbial temporel mais non pas l'adverbial local:

- Avant hier la voiture de mon père était garée devant la gare.
- Devant la gare, la voiture de mon père était garée avant hier.

Dans le cas de *fermer* nous observons le contraire, l'adverbial local seulement pouvant être disloqué:

- En France, les magasins ferment à huit heures.
- ?À huit heures, les magasins ferment en France.

A notre avis le contraste s'explique par le fait que l'adverbial local associé à *se garer* est un complément actantiel et l'adverbial temporel un circonstanciel libre, tandis que pour *fermer* c'est le contraire.

En résumé nous pouvons dire que les études actuelles sur l'ordre des mots montrent clairement la nécessité d'une analyse intégrale tenant compte de facteurs valenciels aussi bien que sémantiques et discursifs. Dans ce qui suit nous allons présenter une analyse détaillée de trois verbes construits avec des compléments directionnels et nous allons essayer de répondre, dans la perspective intégrale exigée, aux questions concernant la relation entre des configurations valencielles déterminées, considérées comme "constructions", et les paramètres discursifs qu'on peut qualifier de "conventionnels".

Analyse de trois structures de valence directionnelles

Les trois verbes (V) choisis pour cette étude sont: *introduzir* et *separar* avec deux compléments actantiels chacun (sans compter le sujet), un complément directionnel et un complément d'objet direct, ainsi que *regressar* avec un actant directionnel seulement. Nous symboliserons l'actant directionnel par "AD" et le second élément, donc le complément d'objet direct dans le cas des deux premiers verbes, ou le circonstanciel libre dans le cas de *regressar*, par "X". Les structures représentées dans le corpus avec leurs fréquences relatives sont les suivantes:

<i>introduzir:</i>	V—X—AD (31)
	V—AD—X (15)
<i>separar:</i>	V—X—AD (51)
	V—AD—X (2)
<i>regressar:</i>	V—X—AD (68)
	V—AD—X (51)

On observe une tendance nette à séparer le directionnel du verbe, quoique la différence entre cette structure et son contraire soit beaucoup plus évidente pour les verbes à complément d'objet direct que pour *regressar*. Une première approche à l'analyse de ces verbes permet de découvrir la cause principale de cette divergence: dans les deux premiers cas nous avons une séquence iconique telle: *introduire/séparer — quelque chose — dans/de quelque chose*. Par conséquent nous trouvons beaucoup d'exemples pour l'ordre V—X—AD, sans qu'interviennent des raisons de textualité, par exemple:

- (1) "*Para obter o seu apoio, o Presidente vai ter de introduzir mudanças substanciais no seu Governo*". (Público, 17.06.96)
- (2) "*pois os 146km que separam Vila Real de Vila Pouca de Aguiar não apresentam obstáculos muito difíceis*". (Público, 03.06.97)

On rencontre même des formules stéréotypiques présentant cet ordre:

- (3) "*...para que se separe o trigo do joio*". (Jornal de Notícias, 25.05.96)
- (4) "*se esquece aquilo que nos separa uns dos outros*". (Diário de Notícias, 18.06.96)

La stabilité de la construction V—X—AD se montre clairement dans le cas de propriétés textuelles concurrentes qui, en principe, exigeraient que X se trouve à droite de AD. Tel est le cas si X a plus de syllabes que AD et devrait donc, selon la loi des membres croissants, se trouver en position finale, Ainsi, dans:

- (5) "*Berisha foi o homem que introduziu as reformas económicas na Albânia e que chegou ao poder com o rótulo de primeiro dirigente democrático do país*". (Público, 26.05.96)

l'AD "*na Albânia*", quoique inférieur en longueur à "*as reformas económicas*" ne se trouve pas à gauche de ce syntagme. Aussi avons nous rencontré le cas (6) où AD ("*do edifício do Congresso*") était séparé de V par deux X ("*em Washington*" et "*a Casa Branca*"):

- (6) "*Apesar da audiência global, o Crossfire da CNN é orientado claramente para o público norte-americano, mais precisamente para a fauna que tem como "habitat" natural as poucas milhas que separam em Washington a Casa Branca do edifício do Congresso*". (Público, 04.06.96)

La construction résiste même dans le cas où X ("*uma norma*") et une proposition relative ("*que nos permite...*") sont textuellement entrelacés, ce qui exigerait un changement de position en combinaison avec le pronom:

- (7) "*e por isso também introduzimos uma norma no regulamento que nos permite, sempre que a estabilidade e a isenção da arbitragem forem postas em causa, punir o árbitro ao ponto de o retirarmos do quadro de nomeações*". (Jornal de Notícias, 11.06.96)

Il n'y a pas de doute que la cause principale de cette stabilité structurelle est l'iconicité de l'ordre des mots. Il s'en suit qu'il sera particulièrement intéressant d'analyser en détail les cas de V-AD-X pour voir s'ils vont de pair avec des régularités discursives. La plus grande régularité est observée dans les cas où AD est anaphorique et X est soit particulièrement rhématique soit focus du segment. Ainsi, dans:

- (8) "*...porque outros países poderiam introduzir nas suas rações os componentes ingleses e vendê-las sem problemas*". (Diário de Notícias, 19.06.96)

le possessif "*suas*" fait référence au syntagme "*outros países*" antérieurement mentionné et se place le plus près possible de lui.

La dislocation de X comme particulièrement rhématique est appuyée par d'autres moyens. Ainsi, dans (9) X ("*ingredientes explosivos*") est typographiquement marqué entre guillemets:

- (9) "*O desalinhado deputado do Partido Popular criticou Rui Rio por ter introduzido no debate 'ingredientes explosivos'*". (Jornal de Notícias, 26.06.96)

Dans (10), la rhématicité particulière de X est mise en évidence par des moyens lexicaux:

- (10) "*E sublinha «o privilégio» que lhe coube em sorte ao ter introduzido no vocabulário, nas escolas e nos laboratórios portugueses o admirável mundo da genética molecular*". (Diário de Notícias, 10.07.96)

Le "*privilégio*" entre guillemets au début du segment anticipe l'importance communicative du complément d'objet direct "*mundo da genética molecular*", appuyé lexicalement par l'adjectif "*admirável*", ce qui lui permet de se positionner de façon atypique à droite d'un AD aussi étendu que "*no vocabulário, nas escolas e nos laboratórios portugueses*".

La stabilité supérieure de V-AD-X peut néanmoins être constatée dans les cas où il existe une coopération de facteurs comme dans (11), où AD représente une information donnée et X est rhématique:

- (11) "*Neste último caso, é simplesmente impensável que qualquer governo se atreva a mexer nas débeis garantias da população em matéria de Segurança Social, sem primeiro introduzir na sociedade portuguesa um mínimo de justiça fiscal*". (Diário de Notícias, 17.06.96)

L'AD "*na sociedade portuguesa*" a déjà été anticipé par "*população*" et le focus du paragraphe (X), hyperbolisé par "*mínimo*", se trouve dans le "*um mínimo de justiça fiscal*" final.

La même coopération de facteurs contribuant à la stabilisation de la construction antiiconique V-AD-X se retrouve également dans le segment

littéraire du corpus qui renferme d'autres exemples avec *separar*. Dans (12), ni AD ni X ne représentent une information nouvelle, mais X est focus:

- (12) "*Minha Luisinha, como tu estás! Pareces um marfim velho, minha querida! — Leve-a para a cama — ordena, seco, Gil Dinis. Olha-o Vina, espantada. Balbucia: — Mas... a casa dela não é aqui. É preciso que ela... — Vina! — exclama um Berto impaciente, vagueando no espaço que separa de Gil Dinis uma Vina irresoluta e desalinhada*". (Bothelho, Fernanda, Xerazade e os Outros, 1926)

Aussi bien "*Gil Dinis*" que "*Vina*" sont des informations données, mais dans ce cas il s'agit d'avancer au centre de l'intérêt la protagoniste Vina en son état d'insécurité absolue: "*uma Vina irresoluta e desalinhada*". Un autre cas de coopération de facteurs se présente dans (13) où AD est l'information donnée et X est textuellement entrelacé:

- (13) „*Animados uns pela presença dos outros, encaminham-se para o grande tûmulo e dali olham para o lugar de onde haviam soado os cânticos. Eis o temeroso espectáculo que têm diante de si: Grossos e altos cancelos de roble separam do resto do templo um extenso recinto sem sepulcros, imediato ao altar principal*". (Herculano, Alexandre, Eurico o Presbítero, Lisboa, ³⁶1944)

L'AD "*do resto do templo*" fait référence au texte immédiatement précédent — une visite de l'enceinte d'un temple — et le X: „*um extenso recinto sem sepulcros*" est concaténé à l'apposition "*imediato ao altar principal*".

En ce qui concerne le verbe *regressar* qui, comme nous l'avons vu, présente une relation plus équilibrée entre les structures V—AD—X (68) et V—X—AD (51), la question se pose s'il s'agit d'un ordre des mots dicté par la structure de valence ou de stéréotypes liés à l'usage d'une certaine configuration adverbiale. En d'autres mots, le problème linguistique à résoudre est celui de décider si nous sommes en présence d'une norme (dans le sens de Coseriu) de valence. Pour trouver la solution, le chemin à suivre — dans le sens de la CG — est l'analyse des différents actes de parole donnés dans lesquels apparaît cette structure afin de découvrir des régularités de récurrence significatives.

Une analyse détaillée des 119 occurrences dans notre corpus montre que les régularités se rapportent à l'élément X qui se combine avec AD. Dans les 68 exemples avec V—AD—X nous avons surtout rencontré les X suivants:

- complément prédicatif: 4 occurrences
- adverbial modal d'action: 9 occurrences
- adverbial temporel déterminé: 48 occurrences²²

²² Les termes "modal d'action" et "temporel déterminé" sont empruntés au travail de L. Lonzi (1991). Le premier désigne un sous-type de l'adverbe modal lié tant au verbe

Exemples caractéristiques:

- (14) „*ontem foi Basler a regressar a casa lesionado*”. (Público, 17.06.96)
 (15) “*os especialistas estrangeiros entretanto chegados a Portugal vão regressar a casa de mãos vazias*”. (Público, 24.07.96)
 (16) “*Carlos regressou a Portugal em 1985*”: (Diário de Notícias, 26.07.96)

Ces résultats montrent qu’au sein de cette structure de valence nous avons une concurrence rhématique des segments. Même si l’information nouvelle de *regressar* est transportée par le directionnel correspondant on peut dire que le prédicatif et le modal d’action qui, au fond, est son homologue au niveau adverbial, portent une charge majeure de l’importance informationnelle parce qu’ils représentent une extension du prédicat, ce qui parfois se montre clairement dans l’entrelacement textuel du prédicat, si on ajoute une information majeure:

- (17) „*Quando, meia hora depois, regressei ao carro, desiludido por não ter conseguido um lugar decente no concerto, o referido ermo tinha a paz e o terror dos cemitérios*”. (Jornal de Notícias, 30.05.96)

Dans cet exemple, l’information de „*desiludido*” est renforcée par la subordonnée infinitive „*por não ter conseguido...*” et le prédicatif, séparé par une virgule, est isolé de l’élément précédent et se trouve ainsi mis en focus.

Les résultats de l’analyse des 51 occurrences de V—X—AD confirment les observations faites à propos de la configuration précédente. Deux X seulement apparaissent dans cette position comme modaux d’action et ils se retrouvent au sein d’une structure où AD est segmenté, montrant ainsi, étant focalisé, sa plus grande importance communicative. Ceci peut être observé dans (18) où l’AD “*as modalidades...*” se trouve séparé de “*em desespero*” par une virgule:

- (18) “*E assim como as pessoas regressam em desespero, às modalidades de sobrevivência tradicionais, não é impossível o «regresso» do doutor Cavaco, se e quando a situação política se vergar sob o peso da velha e justa aspiração ao pleno emprego...*” (Diário de Notícias, 24.06.96)

Les autres X dans la structure V—X—AD sont des adverbiaux aspectuels (11 occurrences) ou bien des connecteurs (5 occurrences) qui évidemment ont peu d’importance communicative. Exemples typiques:

- (19) “*desembarcou no Funchal, mas, como confessou aos seus raros amigos íntimos, regressou de imediato a Lisboa*”. (Público, 31.05.96)

qu’au sujet, par ex.: “*Piero lo ha curato amorevolmente*”; et le second fait référence à un espace de temps défini, comme un fait concret, ou un point défini, par ex: “*Mi occuperò di loro domani*”.

(20) "*Inclinou-se durante algum tempo para o cristianismo protestante, regressando depois, ao hinduísmo*". (Jornal de Notícias, 29.05.96)

Il convient de mentionner spécialement les X en tant qu'adverbiaux temporels déterminés qui se retrouvent au nombre de 48 occurrences dans la structure V-AD-X, comme nous avons vu, et de 21 pour l'arrangement V-X-AD. Une analyse plus détaillée que ces X se placent à droite de l'AD quand il s'agit réellement d'un circonstanciel de temps concret, comme le montrent clairement les trois exemples suivants:

(21) "*regressa à competição após uma greve de fome*". (Público, 26.07.96)

(22) "*...regressada à competição após três meses de paragem*". (Público, 25.07.96)

(23) "*regressou à competição em 1989*". (Diário de Notícias, 29.07.96).

Les circonstanciels ponctuels qui apparaissent directement à droite du verbe sont, au contraire, normalement brefs, à un ou deux mots (*agora, amanhã, antontem, esta noite*, etc.), et servant en général d'arrière-plan:

(24) "*Bielski regressa esta noite ao Palco Grande do Festival com um espectáculo baseado num intrigante texto dramático de Miguel de Unamuno*". (Público, 09.07.96)

Finalement nous voulons attirer l'attention sur la construction causative *fazer regressar*, qui en effet présente un alignement iconique des éléments, comme le montrent les exemples suivants:

(25) "*e tem como principal objectivo para a próxima época fazer regressar a equipa de futebol senior à I Divisão*". (Diário de Notícias, 27.06.96)

(26) "*Netanyahu pretende impor ao Libano um acordo derrotista, sabotar o processo de reconciliação libanês e fazer regressar a guerra civil a este país*". (Publico, 30.07.96)

Il correspond à l'ordre naturel donné par la réalité qu'en parlant de retour, celui qu'on fait retourner soit nommé d'abord et le lieu vers lequel il retourne ensuite. Jusqu'à quel degré l'ordre des mots est figé ne peut être déterminé qu'en analysant un nombre plus important d'exemples, ce qui toutefois ne permettait pas notre corpus réduit.

Conclusions

L'analyse des exemples supporte clairement l'hypothèse qu'il existe en portugais des configurations d'ordre des mots basées sur les structures actanciellles. Dans le cas des adverbiaux directionnels la structure sémantique de certains verbes (*introduzir, separar*) porte en soi sa structure pragmatique,

superficiellement concrétisée par un ordre iconique des éléments actanciels. Seuls des paramètres textuels de poids, comme la valeur rhématique ou l'entrelacement textuel sont capables d'altérer la configuration normative. Dans d'autres cas (*regressar*), ce sont les éléments libres qui déterminent la position de l'actant directionnel. Les éléments les plus rhématiques: prédicatifs, modaux d'action et de temps déterminés placent le directionnel à gauche, immédiatement derrière le verbe, tandis que les éléments aspectuels, les connecteurs et les éléments temporels d'arrière-plan déplacent le directionnel à droite d'eux-mêmes, étant donné que, dans ce cas, il est plus rhématique.

Références

- Aitchison, J., Lewis, D. 1995, *How to Handle Wimps: Incorporating New Lexical Items as an Adult*, *Folia Linguistica* 29/1–2 (7–20).
- Andrade Peres, J., Mória T. 1995, *Áreas Críticas da Língua Portuguesa*. Lisboa.
- Basili, R., Pazienza, M. T., Velardi, P. 1992: *A Shallow Syntactic Analyser to Extract Word Associations from Corpora*, *Literary and Linguistic Computing* 7/2 (113–123).
- Becker, A., Gutfleisch-Rieck, I. 1994, *Freie Wortstellung im Text: Kontextrelationale Reihungsbedingungen*, in: Hafitka, B. (ed.): *Was determiniert Wortstellungsvariation? Studien zu einem Interaktionsfeld von Grammatik, Pragmatik und Sprachtypologie*, Opladen (247–258).
- Biber, D. 1993a, *Using Register-Diversified Corpora for General Language Studies*, *Computational Linguistics* 19/2 (219–241).
- Biber, D. 1993b, *Co-occurrence Patterns among Collocations: A Tool for Corpus-Based Lexical Knowledge Acquisition*, *Computational Linguistics* 19/3 (533–538).
- Bondzio W. 1974, *Die Valenz zweiter Stufe als Grundlage der Adverbialsyntax*, *Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin* XXIII–3/4 (245–257).
- Brent M. R. 1993, *From Grammar to Lexicon: Unsupervised Learning of Lexical Syntax*, *Computational Linguistics* 19/2 (243–262).
- Büchel G., Weber N. 1995, *Semantische Relationen in Definitionsstrukturen*. In: L. Hitzemberger (ed.): *Angewandte Computerlinguistik. Jahrestagung 1995 der GLDV*, Regensburg, 30.–31.3.1995. Hildesheim/Zürich/New York: (127–140).
- Busse W. 1974, *Klasse, Transitivität, Valenz. Transitive Klassen des Verbs im Französischen*, München.
- Busse W. (Coordenador) 1994, *Dicionário sintático de verbos portugueses*. Coimbra.
- Busse W., Vilela M. 1986, *Gramática de Valências*. Coimbra.
- Church K. W., Mercer R. L. 1993, *Introduction to the Special Issue on Computational Linguistics Using Large Corpora*, *Computational Linguistics* 19/1 (1–24).
- Collier A. 1993, *Issues of Large-Scale Collocational Analysis*. In: Aarts J., de Haan P., Oostdijk N. (eds.), *English Language Corpora. Design, Analysis and Exploitation*. 13th International Conference on English Language Research on Computerized Corpora, Nijmegen 1992, Amsterdam–Atlanta, GA: (289–298).
- Cowie J., Dunning T., Guthrie L., Wilks Y. 1994, *Text Processing Using Multilingual Resources at the Computing Research Laboratory*, *Literary and Linguistic Computing* 9/1 (65–78).

- Cruse D. A. 1986, *Lexical Semantics*. Cambridge.
- Daneš F. 1964, *A three-level approach to syntax*, Travaux de Linguistique de Prague I (225–240).
- Delbecq N. 1995, *Un caso de libertad condicionada. La variabilidad del orden de los sintagmas*. In: de Kock J., [et al.] (eds.), *Gramática española. Enseñanza e investigación. Apuntes metodológicos*. Salamanca (149–209).
- Feuillet J. 1966, *Les types de fonctions*. In: Gréciano, Schumacher (129–136).
- Fillmore Ch. J. 1988, *The Mechanisms of „Construction Grammar”*, In: Fourteenth Annual Meeting of the Berkeley Linguistics Society. Berkeley, CA (35–55).
- Fillmore Ch. J., Kay P. 1995, *Construction Grammar*. Ms Stanford: CSLI (CSLI Lecture Notes).
- Fillmore Ch. J., Kay P., O'Connor M. C. 1988, *Regularity and Idiomaticity in Grammatical Constructions: the Case of Let Alone*, Language 64/3 (501–538).
- Gil A. 1996, *Zur Klammerkonstruktion 'Verb-Adverbial-Akkusativobjekt' im Französischen*. In: Gil A., Schmitt Ch. (eds.), *Kohäsion, Kohärenz, Modalität in Texten romanischer Sprachen*. Bonn (70–90).
- Gréciano G., Schumacher H. (eds.), *Lucien Tesnière — Syntaxe structurale et opérations mentales*. Tübingen.
- Gréinadus T. 1990, *Les constructions verbales en français parlé Étude quantitative et descriptive de la syntaxe des 250 verbes les plus fréquents*. Tübingen.
- Gross M. 1975, *On the relations between syntax and semantics*, In: Keenan, Edward L. (ed.): *Formal Semantics of Natural Language*. Cambridge (389–450).
- Helbig G. 1992, *Probleme der Valenz- und Kasus-theorie*. Tübingen.
- Hens G. 1996, *(jm) (einen) Briefschreiben: Zur Valenz in der Konstruktionsgrammatik*, Linguistische Berichte 164 (334–356).
- Johnson-Laird Ph. N. 1982, *Formal Semantics and the Psychology of Meaning*, In: Peters S., Saarinen E. (eds.), *Processes, Beliefs, and Questions*. Essays on Formal Semantics of Natural Language and Natural Language Planning. Dordrecht/Boston/London (1–68).
- Kindt W. 1994, *Wortstellung als Problem einer dynamischen Grammatik*, In: Haftka B., (ed.), *Was determiniert Wortstellungsvariation? Studien zu einem Interaktionsfeld von Grammatik, Pragmatik und Sprachtypologie*. Opladen (49–62).
- Kay P. 1992, *At Least*, In: Lehrer, Kittay (eds.) (309–331).
- Kjellmer G. 1987, *Aspects of English Collocations*, In: Meijs W. (ed.), *Corpus Linguistics and Beyond*. 7th International Conference on English Language Research on Computerized Corpora, Amsterdam, 8.–11.6.1986. Amsterdam/Atlanta, GA (133–140).
- Kjellmer G. 1990, *Patterns of Collocability*, In: Aarts J. Meijs W. (eds.), *Theory and Practice in Corpus Linguistics*. Amsterdam/Atlanta, GA (163–178).
- Koch P. 1996, *La métataxe actancielle. De Tesnière à Busse/Dubost*, In: Gréciano/Schumacher (211–224).
- Korzen H. 1996, *L'unité prédictive et la place du sujet dans les constructions inversées*, Langue française 111 (59–82).
- Lehrer A., Kittay E. (eds.), 1992, *Frames, Fields, and Contrasts*. New Essays in Semantic and Lexical Organization. Hillsdale, NJ.
- Levy F. (ed.) 1994, *Approches Sémantiques*. Traitement Automatique des Langues 35/1.
- Lonzi L. 1991, *Il sintagma avverbale*, In: Renzi L., Salvi G. (eds.), *Grande grammatica italiana di consultazione*. Vol. 2: *I sintagmi verbale, aggettivale, avverbale, la subordinazione*. Bologna (341–414).

- Martin W. J. R., Al B. P. F., van Sterkenburg P. J. G. 1983, *On the processing of a text corpus*. From textual data to lexicographical information. In: Hartmann, R. R. K. (ed.), *Lexicography: Principles and Practice*. London [etc] (77–87).
- Morris J., Hirst G. 1991, *Lexical Cohesion Computed by Thesaural Relations as an Indicator of the Structure of Text*, *Computational Linguistics* 17/1 (21–48).
- Oostdijk N., de Haan P. (eds.) 1994, *Corpus-Based Research into Language*. In honour of Jan Aarts. Amsterdam/Atlanta, GA.
- Pittner K. 1996, *Valenz und Relevanz – eine informationsstrukturelle Erklärung für 'obligatorische' Adverbiale*. In: Pittner R. J., Pittner K. (eds.), *Vorträge der 4. Münchner Linguistik-Tage*. München (95–106).
- Sasse H. -J. 1993, *Das Nomen – eine universale Kategorie?*, *Sprachtypologie und Universalienforschung* 46/3 (187–221).
- Siewierska A. 1993, *On the Interplay of Factors in the Determination of Word Order*. In: Jacobs J. [et al.] (eds.), *Syntax*. Ein internationales Handbuch zeitgenössischer Forschung. 1. Halbband (HSK 9.1). Berlin/New York (826–846).
- Sinclair J. M. (ed.) 1987, *Looking Up*. An Account of the COBUILD Project in Lexical Computing. London/Glasgow.
- Smajda F. 1993, *Retrieving Collocations from Text: Xtract*, *Computational Linguistics* 19/1 (143–177).
- Souter C., Atwell E. (eds.), 1993, *Corpus-Based Computational Linguistics*. Amsterdam.
- Vilela M. 1994, *Tradução e Análise Contrastiva: Teoria e Aplicação*. Lisboa.
- Vilela M. 1995, *Gramática da Língua Portuguesa*. Coimbra.
- Weber N. 1995, *Die Semantik von Bedeutungsbeschreibungen*. ms, Bonn.
- Weber N. 1996, *Formen und Inhalte der Bedeutungsbeschreibung: Definition, Explikation, Repräsentation, Simulation*. In: Weber, Nico (ed.): *Semantik, Lexikographie und Computeranwendungen*. Tübingen (1–46).
- Welke K. M., 1988/89: *Pragmatische Valenz*. In: *Grammatisches Jahrbuch DDR-UdSSR* 4. Moskau (46–56).
- Welke K., Meinhard H. J. 1980, *Prinzipien einer operativen Valenzgrammatik*, *Zeitschrift für Germanistik* 2 (146–156).
- Wierzbicka A. 1992, *Semantic Primitives and Semantic Fields*, In: Lehrer, Kittay (eds.) (209–227).

Summary

Semantic theory and methodology have been traditionally divided into different currents such as psychological vs. logical approaches, operating with mental vs. formal models and roughly corresponding to so-called lexical (paradigmatic and contrastive) vs. phrasal (syntagmatic and combinatorial) semantics. Computational linguistics came up with quantitative, empirically and text-oriented approaches to linguistic meaning, the main test field for these being a recently emerged subfield dubbed corpus linguistics. Arising from practical concerns about how to organise and exploit the information represented by large amounts of stored text "corpus linguistics", unlike many other attribute pairings with "linguistics", does not stand for any theoretical, nor even principled approach. A theoretical foundation possibly could be provided by "Construction Grammar" (CG) which not only proposes to integrate different semantic perspectives but also to take into account pragmatic and textuality parameters.

Some characteristics of CG are sketched here, but no attempt is made to expose its formal apparatus. Some of the relevant publications by its initiators are referred to instead. The focus is then directed to the relation between CG and Valency Grammar (VG). The view taken (not exclusively) here is that CG may in fact be interpreted as a kind of up-to-date VG enhanced by pragma- and text-linguistic aspects. As a consequence the idea arose to try to relate ongoing studies on dependency structures and surface serialisation in Romance languages to the theoretical concepts of CG in order to see if they were basically compatible. This short study of course can only represent a first step towards a real assessment of the relation between CG and VG.

The second part is dedicated to an empirical analysis of sample sentences from the newspaper and literary segments of the "Corpus of Contemporary Portuguese" at the *Centro de Linguística da Universidade de Lisboa*. These illustrate the use of three verbs with a direct and/or a directional complement as well as free circumstantial elements. It is shown that the different actancial structures corresponding to these verbs are represented each by a conventionalized order of elements at the surface of sentences (depending on verb meaning) and that this conventional order may only occasionally be supplanted by important textual (theme-rheme) and discursive (functional) factors. Word order considerations have only recently advanced into the center of attention in valency and dependency studies, and some of the newest publications related to Romance languages, in which this has been the case are used here. The major part of the discussion however is about the evidence presented by the Portuguese corpus examples and the conclusions it leads to.

Streszczenie

Na materiale współczesnego języka portugalskiego autorzy omawiają ważniejsze zagadnienia semantyki leksykalnej i frazowej. Terminologię gramatyki zależności zastosowano do analizy i klasyfikacji przykładów zaczerpniętych z prasy portugalskiej.