

Functional Grammar Series 23

Editors

A. Machtelt Bolkestein
Casper de Groot
J. Lachlan Mackenzie

Mouton de Gruyter
Berlin · New York

Constructing a
Lexicon of English Verbs

1749
Walter de Gruyter
250
1999
Berlin · New York

Constructing a Lexicon of English Verbs

by

Pamela B. Faber
Ricardo Mairal Usón



Mouton de Gruyter
Berlin · New York 1999

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

⊗ Printed on acid-free paper which falls within the guidelines of the
ANSI to ensure permanence and durability.

Library of Congress Cataloging-in-Publication-Data

Faber, Pamela B., 1950–
Constructing a lexicon of English verbs / by Pamela B.
Faber, Ricardo Mairal Usón.
p. cm. – (Functional grammar series ; 23)
Includes bibliographical references (p.) and indexes.
ISBN 3-11-016416-7 (cloth : alk. paper)
1. English language – Verb. 2. English language – Lexi-
cology) I. Mairal Usón, Ricardo, 1965– . II. Title.
III. Series.
PE1271.F3 1999
425–dc21 99-33327
CIP

Die Deutsche Bibliothek – Cataloging-in-Publication-Data

Faber, Pamela B.:
Constructing a lexicon of English verbs / by Pamela B. Faber ;
Ricardo Mairal Usón. – Berlin ; New York : Mouton de
Gruyter, 1999
(Functional grammar series ; 23)
ISBN 3-11-016416-7

© Copyright 1999 by Walter de Gruyter GmbH & Co. KG, D-10785 Berlin
All rights reserved, including those of translation into foreign languages. No part of this
book may be reproduced in any form or by any means, electronic or mechanical, including
photocopy, recording or any information storage and retrieval system, without permission
in writing from the publisher.
Printing: Werner Hildebrand, Berlin. – Binding: Lüderitz & Bauer-GmbH, Berlin.
Printed in Germany.

To the memory of Leocadio Martín Mingorance

Preface

by Christopher Butler

In July 1995, the functional linguistics community was shocked and deeply saddened to hear of the sudden and untimely death of Professor Leocadio Martín Mingorance, of the University of Córdoba, Spain. Martín Mingorance's work, combining the Functional Grammar of Simon Dik with the lexematics of Eugene Coseriu into the lexically-based Functional Lexematic Model, began the process of developing the Functional Grammar conception of the lexicon into a model which integrates semantic, syntactic and pragmatic aspects of lexemes within a framework in which both paradigmatic and syntagmatic patterning find their place.

Prominent among Martín Mingorance's collaborators were Pamela Faber and Ricardo Mairal Usón, whose determination to carry on and develop the line of research pioneered by their friend and mentor has resulted in the present volume. Their aim in this book is impressively ambitious: to give an account of the English verbal lexicon which not only systematises the meanings of lexemes within a hierarchical framework, but also demonstrates the principled connections between meaning and, on the one hand, the syntactic complementation patterns of verbs, and on the other hand, patterns of conceptualization in the human mind.

Such an endeavor is entirely compatible with the tendency towards lexically-based approaches in modern grammatical theory. This shift in paradigm is explored in the first part of Chapter 1 of the book, where developments in lexicology and lexicography are reviewed in relation to their impact on linguistic theorising. Matters of psychological adequacy and computational implementation are also discussed.

Chapter 2 provides an overview of Functional Grammar, concentrating on the role of the lexicon, as a background to the integrated onomasiological model of lexical description represented by the Functional Lexematic Model, within which Faber and Mairal's own proposals are situated.

Chapter 3 provides a more detailed picture of the Functional Lexematic Model and begins to outline Faber and Mairal's own proposals. The scene is set by discussion of the often-criticised concept of 'semantic field' (relabelled as 'domain' in the Functional Lexematic Model), and its relationship with cognitive and generative linguistics, and with a number of approaches to lexical semantics. The aim in the Functional Lexematic Model is to find a kind of onomasiological organization of the lexicon which will be consonant with what is known of the organization of the mental lexicon. Faber and Mairal argue that a hierarchical paradigmatic model comes closest to this goal, but also that the paradigmatic organization of the lexicon into domains and subdomains is the basis for predicting the syntagmatic behaviour of lexemes, insofar as this is represented in syntactic complementation patterns. Hierarchies are developed through detailed, bottom-up analysis of entries in a range of monolingual dictionaries, the guiding principles being those of definitional analysis and lexical decomposition. Faber and Mairal's aim is to develop hierarchies, and their associated complementation patterns, for the whole of the verbal lexicon, in order to achieve a global characterisation of the design of the English vocabulary. Here, as throughout the book, methods are explained, and arguments made, through the analysis of an impressively wide range of semantic domains and their attendant syntax.

In Chapter 4, the relationships between the semantic and syntactic behaviour of verbs are explored in much greater depth, through discussion of three types of lexically-realized parameters which play a part in the generation of clause structures. Grammatical parameters are those which have a direct effect on complementation, and include duration, temporal sequence, iteration, inception, achievement, cessation, causation, conation and factivity: optional parameters are those which are semantically present, but not necessarily syntactically realized; while contextual parameters act as clues for contextual setting. An important claim is that the greater the semantic scope of a lexeme, the greater is its variation in complementation behaviour.

Chapter 5 completes Faber and Mairal's extension of the Functional Lexematic Model, by providing an account of the relationship between lexical structure and cognition. A key concept

here is that of the predicate schema, which integrates paradigmatic and syntagmatic information about an individual lexeme, a lexical subdomain, or a whole domain. Predicate schemas at the subdomain level are accorded particular importance in the model, as they represent the links between the lexical macrostructure and individual lexemes. The authors demonstrate the power of the predicate schema approach in explicating the systematic relationships which underlie many metaphorical extensions of meaning. Chapter 6 concludes the book and presents an overview of the main claims.

Faber and Mairal's book is most welcome at this stage in the development of functional linguistics, in presenting a thought-provoking, challenging and at times controversial account of lexical patterning and its relationship with meaning, syntax and cognition. It is sure to give rise to discussion which will take us even further forward on the road to a functional account of language which responds to the criteria of pragmatic, psychological and typological adequacy so strongly advocated by Dik. It also represents a fitting tribute to the pioneering work of a scholar who was able to take an aspect of Dik's theory and develop its potential: Leocadio Martín Mingorance.

January 1999

Contents

List of tables and figures.....	xv
List of abbreviations.....	xix
1. The lexicon and linguistic theory.....	1
1.1. Introduction.....	1
1.2. Lexical revival.....	4
1.2.1. Linguistic theory.....	6
1.2.2. Lexicology and lexicography.....	8
1.2.3. Psychology.....	15
1.2.4. Computational linguistics.....	18
1.3. Lexicon and grammar.....	22
1.3.1. Event structure and lexical representation.....	23
1.4. Summary.....	34
2. Simon Dik's Functional Grammar.....	37
2.1. An overview.....	37
2.2. FG lexical representation.....	46
2.2.1. Representational parameters.....	47
2.3. FLM lexical description.....	56
2.3.1. The FG lexicon and onomasiological structure.....	58
2.4. Summary.....	65
3. Lexical organization and the FLM.....	67
3.1. Introduction: the concept of semantic field.....	67
3.1.1. Cognitive Linguistics.....	69
3.1.2. Generative Linguistics.....	71
3.1.3. Pustejovsky's Generative Lexicon.....	73
3.1.4. Fillmore's Frame Semantics.....	73
3.1.5. WordNet.....	75
3.1.6. Natural Semantic Metalanguage (NSM).....	76
3.1.7. Meaning Text Theory and the Explanatory Combinatorial Dictionary.....	77
3.2. The FLM: paradigmatic axis.....	79
3.2.1. Definitional analysis and lexical decomposition.....	89
3.2.2. Lexical hierarchies.....	101

3.2.3. Differentiae: semantic parameters	107
3.3. The FLM: syntagmatic axis	114
3.3.1. The semantic characterization of complement selection	115
3.3.2. Complementation and FG	118
3.3.3. Complementation and the FLM	121
3.4. Summary	140
4. Towards a semantic syntax	143
4.1. Introduction	143
4.2. Meaning definitions and syntactic projections	144
4.2.1. Lexically-realized grammatical parameters	145
4.2.2. Lexically-realized optional parameters	176
4.2.3. Lexically-realized contextual parameters	177
4.2.4. Position	179
4.3. Lexical iconicity	186
4.3.1. The genus and its scope	187
4.3.2. The interaction of semantic and syntactic hierarchies	193
4.4. Summary	200
5. The cognitive axis	203
5.1. Introduction	203
5.2. Language and brain	204
5.3. Language and memory	207
5.4. Lexical structure and conceptual structure	210
5.5. Predicate schemas in the FLM lexicon	212
5.5.1. Towards a typology of predicate schemas in FG	217
5.6. Form and representation	248
5.7. A semantic macronet	251
5.7.1. Connections and dependencies	251
5.7.2. The nature and type of lexical interconnections	254
5.8. Summary	271
6. Conclusion	275
Appendix 1: Lexical domains	279
Appendix 2: Semantic network	294
Appendix 3: <i>Possession</i>	298
Appendix 4: Synsem interface of <i>to come to have</i>	301

Notes.....	305
References.....	308
Author index.....	336
Subject index.....	339

List of tables and figures

Tables

Table 1.	Vocabulary matrix.....	21
Table 2.	SoAs in RRG.....	29
Table 3.	Typology of RRG lexical representations.....	30
Table 4.	The structure of the clause in FG.....	39
Table 5.	Symbols used in FG representations.....	41
Table 6.	Overview of operators.....	42
Table 7.	Overview of satellites.....	43
Table 8.	First argument semantic functions.....	51
Table 9.	Second and third argument semantic functions.....	52
Table 10.	Segmentation of the dictionary definitions of <i>embezzle</i>	93
Table 11.	Meaning components of <i>steal</i>	96
Table 12.	Acquisition phase of <i>possession</i>	102
Table 13.	<i>Possession</i> : inventory of complementation patterns of <i>get</i> verbs.....	104
Table 14.	Interaction of semantic distinctions and syntactic complementation.....	108
Table 15.	Semantic and perceptual parameters in manner-of-walking verbs.....	113
Table 16.	Relations between complements and matrix predicates.....	120
Table 17.	<i>Shatter</i> : inventory of selection restrictions.....	127
Table 18.	Semantic and morphosyntactic typology of complements of <i>to think carefully about something</i>	138
Table 19.	Typology of syntactic frames: <i>plan</i> and troponyms.....	139
Table 20.	Semantic and morphosyntactic typology of complements of <i>understand</i>	140
Table 21.	Typology of syntactic frames: <i>to complain</i> <i>continuously</i>	156
Table 22.	Affected arguments in resultative constructions.....	170
Table 23.	Movement in/on liquid.....	171

Table 24. Verbs of <i>contact</i> : lexically-realized contextual parameters	185
Table 25. Typology of syntactic frames: <i>liquid consumption</i>	189
Table 26. Typology of syntactic frames: <i>visual perception</i>	191
Table 27. Typology of syntactic frames: <i>to think something is true</i>	192
Table 28. Typology of syntactic frames: <i>to say that something is certain</i>	194
Table 29. Domain of cognition: interrelations with other domains	214
Table 30. Domain schema.....	218
Table 31. Sense components of <i>bawl</i>	224
Table 32. Definitional constituents of <i>bawl</i> and <i>warble</i>	227
Table 33. Typology of syntactic frames: <i>to think carefully about something</i>	230
Table 34. Typology of arguments: <i>to think carefully about something</i>	231
Table 35. Contemplate: argument structure and meaning correlations.....	232
Table 36. Lexical distinctions in <i>perception</i>	237
Table 37. Phasal distinctions in <i>existence</i>	238
Table 38. Phasal distinctions in <i>cognition</i>	239
Table 39. Variational features	240
Table 40. Institutionalized cultural models	241
Table 41. First-level axiological polarizations in <i>sound</i>	244
Table 42. Second-level axiological polarizations in <i>feeling</i>	247
Table 43. Third-level axiological polarizations in <i>cognition</i> and <i>speech</i>	248
Table 44. Scale of intensity in verbs of <i>damage</i>	257
Table 45. Intensity scale of <i>light</i>	261
Table 46. Verbs of stable <i>light</i>	261
Table 47. Verbs of unstable <i>light</i>	261
Table 48. <i>Glare</i> : argument structure and meaning correlations	263
Table 49. Vision as an event and process.....	266

Figures

Figure 1.	Paradigmatic lexical functions	13
Figure 2.	Outline of FG modules	38
Figure 3.	Macrostructural organization of a lexical domain	64
Figure 4.	<i>To come to have</i>	103
Figure 5.	Synsem interface of <i>promise</i> and its troponyms	200
Figure 6.	Functional divisions of memory	209
Figure 7.	Representation of the domain of <i>existence</i>	219
Figure 8.	Discrimination net for verbs of <i>contact</i>	221
Figure 9.	Image schema: visual image of light sources	259
Figure 10.	Container schema and <i>feeling</i>	260
Figure 11.	Scientific model of vision	264
Figure 12.	<i>Visual perception</i> (physical)	265
Figure 13.	Manner-of-staring verbs	267
Figure 14.	<i>Visual perception</i> with a parallel mental process	268
Figure 15.	<i>Mental perception</i>	269
Figure 16.	<i>To see something in one's mind</i>	270

List of abbreviations

Theories

CG	Construction Grammar
FG	Functional Grammar
FLM	Functional-Lexematic Model
GB	Government and Binding Theory
LFG	Lexical Functional Grammar
MTT	Meaning Text Theory
NSM	Natural Semantic Metalanguage
RRG	Role and Reference Grammar
TG	Transformational Grammar

Theoretical concepts

CA	Componential Analysis
ICM	Idealized Cognitive Model
LCS	Lexical Conceptual Structure
LDB	Lexical Database
LF	Lexical Function
LKB	Lexical Knowledge Base
LS	Logical Structure
LTM	Long Term Memory
MRD	Machine-Readable Dictionary
MTM	Meaning Text Model
NLP	Natural Language Processing
NP	Noun Phrase
SLD	Stepwise Lexical Decomposition
SoA	State of Affairs
SoW	State of the World
STM	Short Term Memory
TAM	Tense-Aspect-Modality
VP	Verb Phrase

Dictionaries

AHD	<i>The American Heritage Dictionary of the English Language</i>
CC	<i>Collins Cobuild English Language Dictionary</i>
ECD	<i>Explanatory-Combinatorial Dictionary</i>
LDCE	<i>Longman Dictionary of Contemporary English</i>
LLCE	<i>Longman Lexicon of Contemporary English</i>
OALD	<i>Oxford Advanced Learners' Dictionary</i>
RHD	<i>The Random House Dictionary of the English Language</i>
WDYR	<i>Webster's New World Dictionary for Young Readers</i>

Abbreviations used in FG-representations

Word classes		Semantic functions	
A	Adjective	∅	zero
N	Noun	Ag	agent
V	Verb	Ben	beneficiary
		Dir	direction
		Exp	experiencer
Syntactic functions		Fo	force
Obj	Object	Go	goal (patient)
Subj	subject	Instr	instrument
		Loc	location
Pragmatic functions		Man	Manner
Foc	focus	Po	Positioner
Top	topic	Proc	Processed
		Rec	Recipient
Layers		Ref	Reference
f	predicate	So	Source
x	term	Temp	Time
e	predication		
X	proposition		
E	clause		

Satellites

- σ_1 any predicate satellite
- σ_2 any predication satellite
- σ_3 any proposition satellite
- σ_4 any illocutionary satellite

Π -operators

- Π_1 any predicate operator
- Π_2 any predication operator
- Π_3 any proposition operator
- Π_4 any illocutionary operator
- Poss Possibility
- Post Posterior
- Sim Simultaneous

Term operators

- Ω any term operator
- \emptyset zero quantifier
- 1 singular
- d definite
- I Indefinite

1. The lexicon and linguistic theory

1.1. Introduction

One of the major points of agreement in contemporary grammatical theory is the centrality of the lexicon as an important source of information about sentence structure. In recent years, many linguistic theories have undergone a gradual change in paradigm, discarding syntactically oriented postulates in favor of more lexically-based ones. The fact that the lexicon is presently in the spotlight signifies that lexemes (and the different types of information they convey) are the building blocks of language in all senses. According to Fellbaum (1998b: 3), the perception of the lexicon as a central component of grammar is due largely to “the discovery that the lexicon is a highly structured repository of rules and principles that give it status and prominence previously accorded only to syntax”. Lexical knowledge has a central role because it is now assumed that much of the structural information of a sentence is best encoded from a lexical perspective.

This panlexicalist perspective has logically brought the message home that semantics cannot be divorced from syntax in any sense. Langacker (1991b: 275) affirms that it is ultimately as pointless to analyze grammatical units without reference to their semantic value as to write a dictionary which omits the meanings of lexical items. Semantic properties must be accounted for, regardless of where the semantics is encoded or what one’s assumptions about the lexicon and syntax are. As to the place where semantics is encoded, most linguistic models agree that it is in predicate-argument structure and that the structure of the clause is predictable from the semantics of the predicates. This premise is now implicit in a wide range of models, such as Role and Reference Grammar (RRG), Government and Binding Theory (GB), and Simon Dik’s Functional Grammar (FG). Other theoretical frameworks, such as Construction Grammar (CG), argue that lexicon and grammar form a continuum, and encode semantics in *constructions*, or pairings of form-meaning

2 *The lexicon and linguistic theory*

correspondences. Syntactic-semantic mappings are thus explained by the association of verb meanings and constructions (Goldberg 1995, 1996; Kay 1997).

As the convergence point for syntax and semantics, the lexicon also becomes the interface of the grammatical and the conceptual component. According to Pustejovsky (1995: 6), the meaning of words should somehow reflect the deeper conceptual structures in the cognitive system and the domain it operates in. In consonance with this, cognitive approaches study meaning in terms of a speaker's construal of situations and how that meaning is embedded in a particular background, scene or frame. As shall be seen, there are different views on the extent to which the lexicon is a map of conceptual structure, but no one can deny that nonlinguistic organizing principles are reflected in the semantics of natural language.

Our study of the English verbal lexicon has been carried out over the last decade within the framework of the Functional-Lexematic Model elaborated by Martín Mingorance (1984, 1990, 1995). We have chosen to examine the organization of the verbal lexicon because of the decisive role that verbs play in sentence structure. Since, to a great extent, verbs determine the number as well as the semantic characteristics of the other elements in the linguistic expression, we believe that their role is just as decisive in the paradigmatic structure of the lexicon. We have found that the analysis of verbs in terms of their semantic and conceptual components can reveal many of their syntactic properties, and that such regularities also are present in other European languages. If a way can be found to organize the verbal lexicon in terms of both semantic and syntactic properties, that would provide the basis for a viable model of lexical representation.

Although there are many lexical semantic analyses of one or two verbs, considerably less has been written about the lexicon as a whole. Those studies with a wider scope invariably seek to explain semantic organization through syntactic behavior. However, in lexical semantic analysis, semantic considerations are all-important because syntactic distinctions in themselves are not sufficient as a basis for the establishment of an inventory of semantic categories.

The evident importance of semantic considerations in any study of the lexicon led us to choose Simon Dik's Functional Grammar (FG) as the framework for our analysis. FG is a dictionary-based grammar in

which the lexicon has always been the most important component (Dik 1997ab). Within this framework, each predicate is described in terms of its form, syntactic category, quantitative valence, qualitative valence, and meaning definition. Proof of the relevance of these parameters is the fact that they also appear in many other grammatical models, both formal and functional. By themselves, however, they are insufficient to account for the full scope of information to be included in a lexical entry. Although the structure of the information within a lexical entry is crucial, provision must also be made to include information about how lexemes are related to others in the lexicon. More specifically, it is necessary to come to grips with meaning on both the microstructural and macrostructural level. Microstructurally, this means examining the role meaning definitions play in the development of an interface between syntax and semantics. Macrostructurally, this signifies situating lexemes within the larger context of their lexical domain, and specifying their relations with lexemes in other areas of meaning. In this way, lexemes are not conceived as a frozen list of items, but rather as dynamic representations within a conceptual network.

Nevertheless, the specification of the macrostructural design of the lexicon, or how lexical items are related through the meaning they convey, cannot be done by using one's intuition to first create a set of *ad hoc* meaning areas, and then trying to distribute lexemes *a posteriori* in pre-ordained semantic classes. An inventory of lexical domains is best arrived at by working upward from word meaning through the systematic application of theoretical premises.

In the 1980s, Martín Mingorance (1984, 1985ab, 1987, 1990, 1995, 1998ab) advanced the theoretical claim that lexical representations are the key to the determination and explanation of the syntactic representations of predicates. The result was the elaboration of the Functional-Lexematic Model (FLM), which integrates Dik's FG and an adapted version of Coseriu's Lexematic Theory. The chapters that follow describe how this model develops the FG lexicon component by configuring it in lexical domains, according to principles of definitional analysis implicit in *Stepwise Lexical Decomposition* (Dik 1978b). The lexical organization obtained is the result of the convergence of both paradigmatic and syntagmatic information. As result, both the micro- and macrostructure of these domains are a reflection of the interface of syntax and semantics. In a parallel way, it

4 *The lexicon and linguistic theory*

also provides insights into the encoding of lexical meaning as a base for knowledge representation.

This means that an accurate linguistic theory should be able to account for the interrelationships between syntax and semantics. One of the research goals in lexical semantics is the development of semantic representations that could serve as a basis for the formulation of a set of linking rules, which will ultimately determine the syntactic realization of arguments. As a result, the design of the lexicon is all-important because it has become the core of most grammatical models. Jackendoff (1975: 639–640) states that a theory of the lexicon must meet the following three levels of adequacy:

- (i) *Observational adequacy*, by means of which a theory of the lexicon supplies each lexical entry with sufficient information to account for the full potentiality of the speaker's lexical competence.
- (ii) *Descriptive adequacy*, which maps out the type of relations, regularities and generalizations which hold between lexical items;
- (iii) *Explanatory adequacy*, which shows how and why the set of regularities and particular relationships in the lexicon are chosen, and moreover, how these relations affect grammar.

Since it is beyond the scope of this book to discuss all of these levels in detail, we shall focus principally on the levels of explanatory and descriptive adequacy. However, first we shall examine how different models of lexical analysis have tried to explain lexical regularities and relations, and how they affect grammar.

1.2. Lexical revival

The lexicon did not always hold the privileged position in linguistic theory that it now enjoys. For many years, it was relegated to the background, and regarded as a wastebasket for peripheral information and irregularities which syntax could not explain. Giving clear priority to formal grammatical description, Bloomfield (1933: 274) considered the lexicon merely to be “an appendix of grammar”. For this reason, he

is often blamed for shoving meaning into the closet, but as Wierzbicka (1996: 4) points out, the banishment of meaning from linguistics can primarily be attributed to Bloomfield's followers. Although Bloomfield himself wished to eliminate semantic considerations from linguistic analysis, he did not reject meaning in the sense of avoiding any mention of it in linguistic description. As syntax was easier to study objectively, linguists seemed to think that it should be dealt with first, and that meaning could be put on the back-burner until all the mysteries of syntax had been resolved.

This is one of the basic premises of the first version of Transformational Grammar. Wierzbicka (1996: 5–6) mentions the fact that although Chomsky was one of the main actors in the “cognitive revolution” of the late fifties and sixties, he is still Bloomfieldian in the sense that he is also reticent about dealing with meaning. This was most evident in *Syntactic Structures*, in which Chomsky (1957) makes no provision for either a lexicon component, or specific principles to determine the contextual restrictions that a language imposes on lexical insertion. The majority of linguistic schools at that time were not very interested in lexical items, and chose to disregard the fact that the lexicon contains regularities that are meaningful for syntax. There was a marked tendency to minimize lexical information to the benefit of a more algorithmic syntactic component. Even semantically oriented linguists seemed to be more concerned with explaining how to obtain the meaning of sentences recursively from syntactic structure, as well as from the semantic value of those constituents with an autonomous meaning.

Nevertheless, it soon became evident that the emphasis on autonomous syntax was not especially conducive to understanding language, and that certain syntax-related problems could only be resolved by taking a closer look at meaning. Important factors in the process were the evolution of Chomskyan linguistics and the increasingly important role of the lexicon within its framework. In fact, the latest versions of Chomskyan linguistics now include references to meaning though statements in this direction have been largely programmatic.

Another contributing factor was the appearance of other competing frameworks, which were even more lexically centered. Relevant examples are Hudson's (1976: 1–26) *Word Grammar*, which

dispenses with transformations altogether; the grammar proposed by Montague (1970ab, 1973) with its semantic component, which formalizes the relations that transformations signal at the level of semantics; and Lexical-Functional Grammar (Bresnan 1979, 1982), which treats syntactic phenomena as operations on lexical forms.

The specific reasons why linguists began to look towards semantics again are various, but it is no coincidence that this “lexical reconciliation” (Grimshaw 1994) is closely linked to major advances in linguistic theory, lexicology and lexicography, psychology, and computational linguistics.

1.2.1. Linguistic theory

One of the things that contributed greatly to this lexical revival was the progressive theoretical reorganization that took place in Transformational Grammar (TG). In *Aspects of the Theory of Syntax* (Chomsky 1965), a lexicon containing context-sensitive lexical-insertion rules is introduced as a subcomponent of the base. Phrase structure rules, together with other components and principles, generate a deep structure, which is a sort of skeleton with preterminal nodes. The function of the lexicon is to insert *a posteriori* lexical items into the nodes provided by the syntactic rules of the categorial component (Katz—Fodor 1963).

In the TG lexicon, each lexical entry has a subcategorization frame, specifying the type of complements a lexeme can co-occur with. The semantic nature of the complements is encoded by means of a series of selection restrictions (e.g. animate, inanimate, abstract, concrete). In this sense, *Aspects* paved the way for the more lexically-based approaches that followed. At the end of the sixties, serious attempts were made to give TG a more important semantic component. Generative Semanticists formulated underlying semantic representations on the basis of what they believed to be an inventory of semantic primitives (McCawley 1968).

Of special interest here is the use of semantic decomposition, a process by means of which lexical items are analyzed in terms of a finite set of semantic components. A case in point is McCawley's (1968) decomposition of the verb *kill*, which Fodor (1970) criticizes.

Despite the fact that this conception of semantic decomposition came to be considered inadequate for semantic representation, it was never completely abandoned, and has now been incorporated in various semantic theories of the lexicon (Jackendoff 1983, 1987; Talmy 1985; Ravin 1992; Van Valin 1993b).

Generative Semantics was harshly criticized by Interpretative Semantics, a rival theory which extended the format proposed in *Aspects* (Chomsky 1970; Jackendoff 1972). The new proposal for the subcategorization frames of lexical entries included information, such as category, syntactic pattern, and thematic roles, introduced as variable positions in the semantic representations of predicates (Jackendoff 1972: 38–43). The debate between Generative and Interpretative Semantics had the positive effect of causing linguists to reconsider the lexicon and of including linguistic phenomena previously regarded as syntactic in lexical entries. For example, Chomsky's *Lexicalist Hypothesis* postulates that more and more syntactic phenomena (including transformations) need semantic explanations (Kiparsky—Kiparsky 1970; Katz—Postal 1964).

However, after the fall of Generative Semantics, linguistic models became extremely complex and abstract. As a result, they had difficulty in coping with morphology, word-formation, lexical semantics, and pragmatic factors such as presupposition and discourse analysis (Aronoff in Nowakowski 1990: 4; Sánchez de Závala 1976: 95–103). In an effort to make generative models more explanatory, transformations were virtually eliminated by reducing all of them to *move α* (Chomsky 1973, 1981b; Newmeyer 1980). The rise in importance of the lexicon brought with it a corresponding simplification of algorithmic rules, and in some cases, the categorial component (e.g. phrase structure rules) was eliminated altogether.

Chomsky (1986: 86–87) questions the validity of the classical theory, which postulates that both categorial selection (c-selection) and semantic selection (s-selection) should be included in each lexical entry. In his opinion, the classical analysis contains redundant information, and he comes to the conclusion that U(niversal G(rammar) does not need an independent theory of subcategorization since these properties are largely predictable from other principles such as Lexical Projection and Case Theory. He suggests that the internal make-up of the lexicon should be restricted to s-selection since subcategorization

can be derived from s-selection, but not vice versa. An important consequence of this radical change of perspective in Chomskyan linguistics was the formulation of lexical theories in which the lexicon is conceived as a repository of the syntactic properties of argument-taking predicates.

1.2.2. Lexicology and lexicography

As meaning began to come out of the closet in linguistic theory, more importance was logically given to lexicology and lexicography. This led to the development of lexicological frameworks with a clear semantic focus, one of the aims of which was the elaboration of enriched lexical representations. Two of these frameworks, which have certain aspects in common with the FLM, are Wierzbicka's Natural Semantic Metalanguage and Mel'cuk and Zholkovskij's Meaning Text Theory.

1.2.2.1. Natural Semantic Metalanguage

The *Reductive Paraphrase* or Natural Semantic Metalanguage (NSM) approach elaborated by Wierzbicka (1987, 1988, 1996) combines the philosophical and logical tradition in the study of meaning with a typological approach to the study of language, and with empirical crosslinguistic investigations. Wierzbicka's research has focused on the identification of the shared core of all natural language in order to formulate the NSM, an intermediary language of semantic description, which uses natural language expressions, thus avoiding abstract features, terms or logical symbols (Goddard 1998: 56–64).

The NSM is the result of semantic analysis, the underlying principle of which is explication through reductive paraphrase. Wierzbicka (1995: 149) compares the meanings of words to “objects constructed out of various Lego blocks”. Her approach to lexical semantics analyzes word meaning in order to find its most basic components. With this inventory of semantic primitives, she claims that it is theoretically possible to define all of the other words in the

language. In this way, the meaning of a word is formulated in the simplest possible terms, and circularity is avoided. Nevertheless, in order for such a description to be considered descriptively adequate, it must be able to replace the original expression without change of meaning.

For example, the NSM definition of *mother* is the following:

- (1) X is Y's mother
 at some time before now, Y was very small
 at this time Y was inside X's body, Y was like part of X
 because of this, people can think something like this:
 X wants to do good things for Y
 X doesn't want bad things to happen to Y
 (Goddard 1998: 62).

One of the main objectives of this semantic analysis is the elaboration of an inventory of a semantically minimal core, consisting of primitives that cannot be defined in simpler terms. According to Wierzbicka, such primitives are not directly observable, but can only be established by trial and error. Over the years, her original list of fourteen semantic primitives has grown to nearly sixty. These include predicates from various semantic categories: mental [*think, want, know, feel, see, hear*], speech [*say*], action and event [*do, happen, move*], existence and life [*be, live*], relational [*like, part of*], space [*in, under, above, on, side, inside, far, near*], evaluation [*good, bad*], and description [*big, small*].

Such semantic primitives are accorded the status of lexical universals, which are the symbolic representation of universal conceptual primitives. The sets of semantic primitives identified in this way across languages are regarded as language-specific manifestations of a universal set of fundamental human concepts, and are the result of a limited isomorphism in the lexicon as well as grammar. Wierzbicka (1995: 154) affirms that these concepts are the basis of the *tertium comparationis* that necessarily underlies successful interlinguistic communication.

Within a particular language, every element belongs to a unique network of elements, and occupies a particular

10 *The lexicon and linguistic theory*

place in a unique network of relationships. When we compare two or more languages, we cannot expect to find identical networks of relationships. We can, none the less, expect to find corresponding sets of indefinables (Wierzbicka 1996: 15).

The set of semantic primitives is supposed to be a complete lexicon for semantic analysis. As a kind of “mini-language” or subset of a full natural language, it naturally has its own syntax as well. Wierzbicka (1996: 144) characterizes the syntax of the NSM as analogous in some respects to, but much simpler than, the syntax of natural languages with a rudimentary parts-of-speech system. Goddard (1998: 329) describes it as primarily combinatorial in nature. In fact, the latest research in this area is focused on discovering universal patterns of co-occurrence among primitives, which constitute *canonical context* sentences or sentence fragments.

Another concept within this theory is that of valency options in relation to semantic primitives. For example, in the NSM *say* has four valency options:

- (2) SOMEONE SAID SOMETHING
SOMEONE SAID SOMETHING TO SOMEONE [optional addressee slot]
SOMEONE SAID SOMETHING ABOUT SOMETHING [optional topic slot]
SOMEONE SAID SOMETHING TO SOMEONE ABOUT SOMETHING
[both optional slots filled]

Since universal semantic primitives cannot be defined because more basic words than themselves do not exist, they are identified through universal syntactic patterns. This is extremely important, given that such primitives are frequently polysemous, and their different senses vary from one language to another.

Goddard (1998: 335) affirms that a universal set of participant roles can be formulated by tying them to the argument slots of semantic primitives like DO, HAPPEN, MOVE, KNOW, THINK, FEEL, and WANT. In this sense, Wierzbicka’s proposal runs parallel to Jackendoff (1990), who also posits that a complex or non-primitive predicate can impose

multiple participant roles upon an argument because of its compositional nature.

Since the NSM is intended to be a model of the innate and universal *lingua mentalis*, the NSM grammar styles itself as a hypothesis about the grammar of human cognition. The basic claim is that there exists a set of correspondences between semantic primitives of all languages and also between certain combinations of primitives, so that essentially anything that can be said in one NSM can be accurately translated into another.

1.2.2.2. Meaning Text Theory

Meaning Text Theory (MTT), first proposed by Mel'cuk—Zholkovskij (1970), is one of the most influential meaning-based theories in contemporary linguistics. Though it is somewhat different from the FLM approach, there are also many areas of convergence. MTT is based on the fact that any act of linguistic communication involves content (meaning), an oral or written signal (text), and a mapping (a set of correspondences between meanings and texts). One of its postulates is that a natural language L is a finite set of many-to-many correspondences between a specific set of meanings and a specific set of texts. Though meanings and texts are accessible to the linguist, the correspondence between them is not. As a result, a natural language can only be described as a functional model, or a system of formal rules which simulates the linguistic behavior of L's speakers.

MTT describes a set of Meaning Text Models (MTMs) or possible models, which relate texts to their meaning representations through the description of mappings between meanings and natural language texts. It specifies the following seven levels of description:

- (i) Semantic representations (SemR)
- (ii) Deep syntactic representations (DsyntR)
- (iii) Surface syntactic representations (SsyntR)
- (iv) Deep morphological representations (DmorphR)
- (v) Surface morphological representations (SmorphR)
- (vi) Deep phonetic representations (DphonR)

(vii) Surface phonetic representations (SphonR)

An utterance is simultaneously characterized by representations at all levels (Wanner 1996b: 23).

Lexical information is encoded in the *Explanatory Combinatorial Dictionary* (ECD) (Mel'cuk 1981, 1988, 1989, 1996; Mel'cuk—Zholkovskij 1970; Mel'cuk et al. 1984). The purpose of the ECD is to cover all linguistic knowledge and capture lexical regularities at all levels. Lexical entries are divided into three zones: the semantic zone, the syntactic zone and the lexical combinatorial zone. The semantic zone specifies a semantic network, which defines the meaning of the lexical entry in terms of simpler meaningful elements. The syntactic zone of an entry contains the government pattern, which specifies for each semantic actant of the headword, the corresponding deep-syntactic actant, and all surface representations of the latter in the text. The lexical combinatorial zone specifies semantically related lexemes as the values of lexical functions.

MTT also proposes an inventory of lexical functions (LFs), which codify different types of semantic and syntactic relations. According to Mel'cuk (1996: 39), a lexical function f is a correspondence that associates a given lexical expression L with a set of lexical items L_1 which express a specific meaning associated with f . This can be represented by the following formula: $f(L) = L_1$.

Mel'cuk defines close to sixty lexical functions and using Saussure's dichotomy, divides them into two types: paradigmatic LF relations and syntagmatic LF relations. Paradigmatic LFs associate with a keyword a set of lexical items that share a non-trivial semantic component, and include all contrast and substitution relations between lexical items in specific contexts. A few examples of such functions, some of which coincide with sense relations, can be seen below:

- (3) a. *Gener (republic) = state*
 b. *Anti (joy) = grief*
 c. *Contr (fire) = ice*

The lexical function *Gener* (3a) not only covers hyponymic institutionalized lexical relations, but also the relation between a lexeme and its closest generic concept, which can be used as an

attributive modifier. Mel'cuk specifies different types of antonymy, but generally this relation in MTT is the result of the semantic decomposition of lexical units, as seen in their respective definitions. In (3b), *joy* is considered antonymous to *grief* because the former is defined as a pleasant emotion and the latter, as an unpleasant emotion. The opposition between *pleasant* and *unpleasant* is thus the basis for antonymy. The contrastiveness relation is different from traditional antonymy in that it does not necessarily entail a logical opposition between the lexical units opposed. Example (3c) shows a contrastiveness relation in the sense of Apresjan's modal antonyms (Wanner 1996b: 9).

Although various classifications for LFs have been proposed, one of the most coherent proposals is that of Alonso Ramos, Tutin, and Lapalme (1995: 353), who have summarized paradigmatic LFs in Figure 1:

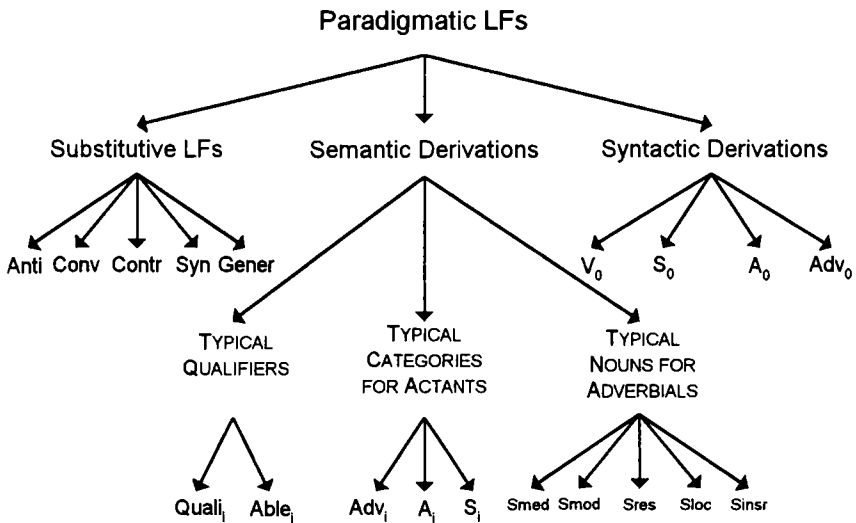


Figure 1. Paradigmatic lexical functions

Syntagmatic LFs formalize a semantic relation between two lexemes L_1 and L_2 , which is instantiated in the textual string in a nonpredictable way. Such a relation is nonpredictable when the co-occurrence of one cannot be derived from the semantic selection restrictions of the other, but rather has to be learned as an instantiation

of a specific syntagmatic relation. Syntagmatic LFs may only be semantically constrained, or they may have a lexical bias. Examples of such functions can be seen in the following:

- (4) a. *Magn* (*naked*) = *stark*
 b. *Instr* (*satellite*) = *via*
 c. *Func*₀ (*snow*_N) = *falls*

For example, the lexical function *Magn* (4a) is an attributive relation between two lexemes, and codifies the highest degree of the semantic content in question. *Nakedness* always means not having any clothes on, but being *stark naked* is the total absence of any covering/clothing whatsoever, probably within a context that makes this state seem even more surprising than usual. *Instr* (4b) signifies *by means of*. In the case of *satellite*, the instrument relation is indicated by *via* since television programs are more often described as being received *via* satellite, rather than *by/by means of* satellite. Another example of a syntagmatic lexical function is that of *Func* (4c), which codifies a type of noun-verb collocation indicating performance. In this sense, the verb that generally collocates with *snow* is *fall* because this is the action usually associated with it.

Such relations are extremely widespread across language and languages, and have proved to be extremely useful in computational linguistics. LFs are a means for a systematic description of what Wanner (1996b: 1) describes as “institutionalized” lexical relations, and can be used in the representation of lexical semantic relations in computer applications, such as text generation (Alonso Ramos—Tutin—Lapalme 1995).

1.2.2.3. The formalization of lexical knowledge

Both the NSM and the MTT, each in its own way, underline the fact that conceptual knowledge can be derived from fully specified lexical entries, and that such entries constitute mini-knowledge representations. The models described are similar to the FLM in that both of them are based on meaning and lexical analysis. Another similarity is that they initially arose from a general dissatisfaction

with the traditional concept of a dictionary and the desire to produce a better one by enriching it with more information, structured according to a set of underlying theoretical principles. This new type of dictionary endeavors to meet various standards of adequacy. It is envisioned as being a lexicological as well as a lexicographic product, whose dynamic character accounts for the actual lexical competence of speakers. This dynamism is manifest in an enriched entry, which formalizes, at least to a certain degree, different types of relations between words and aspects of meaning previously not accounted for. Such lexically-based approaches to linguistic theory point to a more mentalistic approach, in which language is conceived as a blueprint of the mind of the Natural Language User.

1.2.3. Psychology

Enhanced lexical representations go hand in hand with psychological adequacy. The objective of including more information about words and their interrelations in lexical entries is to make their configuration more similar to the organization within our mental lexicon. In line with this search for psychological (and neurological) adequacy, developments in psychology have contributed to a better understanding of meaning, as well as the structure of the mental lexicon.

The word *lexicon* is used in psycholinguistics to refer to the mental lexicon of the natural language user. Emmorey—Fromkin (1988: 12) define the mental lexicon as a component in the grammar in which information about individual words and/or morphemes is entered. Consequently, the mental lexicon contains what a speaker/hearer of a language knows about the form of the entry (its phonology), its meaning (its semantic representation), and its combinatorial properties (its syntax). The lexicon is thus conceived as lying at the interface between those conceptual structures generated by our cognitive systems and their syntax (González Escribano 1991: 213).

Two central issues in psycholinguistics are lexical acquisition and the configuration of lexical knowledge in memory. The fact that a speaker can mentally find the word that he/she wants in less than 200

milliseconds, and in certain cases, even before it is heard, is proof that the mental lexicon is organized in such a way as to facilitate access and retrieval. In order to arrive at the meaning of an utterance, the hearer must, among other things, determine the semantic relationships between rapidly arriving words. This is done by analyzing the acoustic information received in order to access the lexicon for both semantic and syntactic information in lexical representations. The ultimate goal is to develop the propositional representation of the utterance or the relationship among the entities, processes and events described (Wingfield—Titone 1998: 241–242). Studies of sentence processing suggest that normally the hearer/text receiver tries to understand the meaning as quickly as possible and then discards the surface structure (Bransford—Franks 1971; Bransford—Barclay—Franks 1972). Wingfield—Titone (1998: 243) point out that if this is true, then it is the conceptual representation of the utterance that is the primary focus of the memory system. The logical conclusion is that surface form is subordinate to meaning and not the other way around.

In a parallel way, the complexity of lexical memory has fascinated many psycholinguists, who have proposed different ways of exploring and analyzing the cognitive processes involved in acquisition and retrieval, as well as the configuration of concepts. It is not easy to gain access to the “black box” of the mind, and for this reason, methods have been principally analogical, or according to Miller—Fellbaum (1991: 198), only centered on a tiny part of the lexicon. Researchers often have used language structure to obtain a better understanding of conceptual organization, despite the fact that the exact relationship between lexical structure and conceptualization is not known.

Pederson—Nuyts (1997: 4–6) affirm that linguistic behavior (language comprehension and production) has a special status because it provides the richest possibilities for investigating conceptualization. Furthermore, theories which consider conceptualization to be closest to language are generally much further ahead in terms of formal modeling of conceptual structures and its relationship with linguistic structure (as in the case of Dik or Jackendoff) than are views which maintain that conceptualization is more abstract or at least less directly language-based. However, Pederson—Nuyts are careful to indicate that this does not mean that the more abstract or less language-related views are mistaken since the language view obviously cannot account for the