

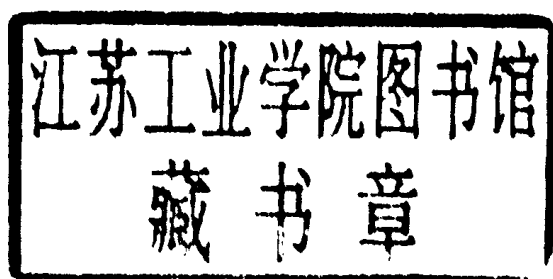
# THE GENERATIVE LEXICON

JAMES PUSTEJOVSKY



# The Generative Lexicon

James Pustejovsky



The MIT Press  
Cambridge, Massachusetts  
London, England

© 1995 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

This book was set in Computer Modern by the author and printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Pustejovsky, J. (James)

The Generative Lexicon / James Pustejovsky.

p. cm.

Based on the author's unpublished manuscript, "Towards a Generative Lexicon."

Includes bibliographical references (p. ) and indexes.

ISBN 0-262-16158-3

1. Semantics. 2. Generative grammar 3. Computational linguistics. I. Title

P325.P85 1995

401'.43-dc20

95-32875

CIP

## Preface

This book is based on a larger unpublished manuscript, *Towards a Generative Lexicon*, written between 1988–1993. Many of the basic ideas for this manuscript are first explored in a general way in an article written in 1987 with Bob Ingria, entitled “Active Objects in Syntax and Semantics.” Since the original publication of the article “Generative Lexicon” in the journal *Computational Linguistics* in 1991, what had started merely as a critique of theories of lexical sense enumeration has developed into a fairly specific proposal for how to perform lexical semantic analysis. Substantial new material has emerged from cooperative work with several colleagues. This includes the original work done with Bran Boguraev and discussion with Ted Briscoe and Ann Copestake in the context of their Acquilex research funded by Esprit in Europe. The work on unaccusativity stems from joint research with Federica Busa of Brandeis. Extensions and elaborations of the coercion analysis for aspectual predicates in French has been done in close collaboration with Pierrette Bouillon of ISSCO and the University of Paris.

Several chapters have been omitted for clarity of presentation and in some cases for clarity of the content. Hence, some topics that were expected to be included have been deleted entirely. For example, two important themes in generative lexical studies (the role played by Lexical Inheritance Theory and the theory of co-specification) are not examined in any depth in this monograph; both of these areas have become too large to make only passing reference to, and I felt the discussion possible in this monograph would do no justice to these issues. Regarding lexical inheritance, recent work in computational lexicography and semantics, much of it done in the context of the Acquilex project (and reported in Briscoe *et al.*, 1993) has pointed to new and exciting directions for how lexicons should be organized globally. The promise of realizing a *projective inheritance model*, as suggested in my 1991 article, awaits further investigation, although it is currently one of the topics of research at Brandeis in conjunction with Bran Boguraev at Apple.

The second major omission in this monograph is a comprehensive discussion of co-specification and processes of selection. This has proved to be a central concern in the applied computational research at Brandeis as well as the recent work on lexical acquisition and induction from corpora. In the present study, however, I have chosen to concentrate on the core mechanisms involved in semantic selection as they relate to

syntactic expression. Because co-specification treats the subtle use and variation in selection below the level of conventional semantic and syntactic types, it is impossible to do it justice without extensive discussion. This can be found in Pustejovsky (forthcoming) and to a certain extent in Boguraev and Pustejovsky (1996).

Finally, I should point out that many questions relating to natural language semantics are not investigated in any great detail here. In particular, issues surrounding quantification and genericity are only touched on briefly, if at all. Furthermore, details of several of the mechanisms of composition are to be found not here but in other works, including Pustejovsky (1995b) and Pustejovsky and Johnston (forthcoming). My aim in the current work has been to outline what I feel is the necessary infrastructure for a truly generative, highly distributed, and lexically-based semantic theory for language.

James Pustejovsky  
Brandeis University

## Acknowledgments

The ideas contained in this work have been greatly influenced by conversations and interactions with numerous friends and colleagues. Those who have contributed to the specifics of this work include: Noam Chomsky, Bob Ingria, Bran Boguraev, Ted Briscoe, Federica Busa, Patrick Saint-Dizier, Nicholas Asher, Ray Jackendoff, Geoffrey Nunberg, Michael Johnston, Pierrette Bouillon, Scott Waterman, Sabine Bergler, Peter Anick, Paul Buitelaar, Evelyne Viegas, and David McDonald.

Early encouragement from Dave Waltz, Jane Grimshaw, Jim Higginbotham, Remko Scha, and Edwin Williams helped shape the direction of the work, and their influence is seen throughout the pages of this book. Critical interactions with many colleagues have been useful in sharpening my proposal and the specific analyses. In particular, George Miller, Jerry Hobbs, Manfred Bierwisch, Jürgen Kunze, Ewald Lang, Beth Levin, Barbara Grosz, Sergei Nirenburg, Aravind Joshi, Chuck Fillmore, Rick Alterman, Yorick Wilks, David Waltz, and Peter Cariani have provided various degrees of commentary and criticism, making, I believe, the resulting work much clearer than it might have been.

Large portions of the material reported on in this book have been presented in front of audiences from linguistics and computer science conferences and departments, including Brandeis, UMIST, MIT, University of Pennsylvania, Princeton, McGill, the University of Texas, Stuttgart, Humboldt University in Berlin, the CUNY Sentence Processing Conference in New York, IBM, SUNY Buffalo, Toronto, Toulouse, Cambridge University, University of Copenhagen, and Charles University in Prague. Many of the comments and questions from these interactions have found their way into the work presented here, and I would like to thank the participants of these talks for their critical and helpful remarks.

I would like to also thank the following for carefully reading the final manuscript: Bob Ingria, Federica Busa, Ann Copestake, Henk Verkuyl, Ted Briscoe, Nicholas Asher, Alex Lascarides, Anne-Marie Mineur, Michael Johnston, and Marc Verhagen.

I would like to give particular thanks to MIT Press, for their encouragement and editorial support of this project. To Harry Stanton, for his patience and faith, and to Amy Pierce, for her enormous help with guiding the book through the Press on time, and further expediting the publication. And finally, to Deborah Cantor-Adams, for her wonderful editorial help and support with the manuscript.

For moral support, there are really too many people to thank. I would like to thank my mother, Frances, for her constant energy and encouragement; Charlie and Helga Marqusee, for their wonderful creature comforts and Charlie's personal mail and delivery service; John Broglio, for lending me his apartment in Northampton to rewrite the manuscript; and to Brent Cochran, Peter Cariani, Bob Ingria, and Fede Busa, for their undeserved faith in me.

And finally and most significantly, to my family, Cathie, Zachary, and Sophia for always being there, and enduring the tedious hours and my testy moodiness.

This work was supported in part by grants from The Advanced Research Project Agency (MDA904-91-C-9328); the National Science Foundation (IRI-93-14955); Digital Equipment Corporation; and the National Institutes for Health.

I would like to dedicate this book to the memory of my father, George Harry Pustejovsky.

At this point, as it often happens in philosophy, we suddenly realize that the path of inquiry we hoped to open is already marked by the footprints of Aristotle.

Zeno Vendler

*Linguistics in Philosophy*



# Contents

Preface	ix
Acknowledgments	xi
<b>1 Introduction</b>	<b>1</b>
<b>2 The Nature of Lexical Knowledge</b>	<b>5</b>
2.1 Semantic Classes and Categorial Alternation	8
2.2 Interlexical Relations	23
<b>3 The Logical Problem of Polysemy</b>	<b>27</b>
3.1 Varieties of Sense Extension	27
3.2 Contrastive Ambiguity	29
3.3 Complementary Polysemy	31
3.4 An Elementary Lexical Semantic Theory	33
<b>4 Limitations of Sense Enumerative Lexicons</b>	<b>39</b>
4.1 The Goals of Lexical Semantic Theory	40
4.2 The Creative Use of Words	42
4.3 Permeability of Word Senses	47
4.4 Difference in Syntactic Forms	50
4.5 Semantic Expressiveness	55
4.6 Generative Lexical Models	57
4.7 Strong vs. Weak Compositionality	59
<b>5 The Semantic Type System</b>	<b>61</b>
5.1 Levels of Representation	61
5.2 Argument Structure	62
5.3 Extended Event Structure	67
5.4 Qualia Structure	76
5.5 The Interaction of Semantic Levels	81

<b>6</b>	<b>Qualia Structure</b>	<b>85</b>
6.1	Modes of Explanation	85
6.2	The Qualia Structure of Nominals	90
<b>7</b>	<b>Generative Mechanisms in Semantics</b>	<b>105</b>
7.1	Coercion and Type Shifting	106
7.2	Co-composition	122
7.3	Selective Binding	127
7.4	Semantic Selection	131
<b>8</b>	<b>The Semantics of Nominals</b>	<b>141</b>
8.1	Basic Issues	141
8.2	Nominals with Unified Types	142
8.3	Nominals with Complex Types	149
8.4	Propositions and Event Descriptions	157
8.5	Varieties of Nominalization	165
8.6	Lexicalization and LCPs	177
<b>9</b>	<b>The Lexical Semantics of Causation</b>	<b>183</b>
9.1	How Language Encodes Causation	183
9.2	Causation and Unaccusativity	188
9.3	Aspectual Causatives and Coercion	198
9.4	Experiencer Predicates	208
9.5	Modal Causatives	215
9.6	Conclusion	218
<b>10</b>	<b>Consequences of a Generative Lexicon</b>	<b>221</b>
10.1	Co-composition and Verbal Semantics	221
10.2	Stage-Level Predication	225
10.3	Further Applications of Coercion	230
10.4	Linguistic versus Commonsense Knowledge	232

10.5 Lexical Inference and Rhetoric	236
10.6 Summary	238
Notes	241
Bibliography	269
Author Index	289
Subject Index	293

# 1 Introduction

This book deals with natural language semantics, and in particular the semantics of words, both alone and in combination, i.e. the problem of compositionality. Lexical semantics is the study of how and what the words of a language denote. Computational and theoretical linguists have largely treated the lexicon as a static set of word senses, tagged with features for syntactic, morphological, and semantic information. Under this view, different word senses have been generally associated with distinct lexical items. Nevertheless, formal theories of natural language semantics have done little to address two important issues:

- the creative use of words in novel contexts;
- an evaluation of lexical semantic models on the basis of compositionality.

In this study I examine the interaction of word meaning and compositionality as they relate to these concerns. I will argue that, by adequately accounting for the problem of creative word senses, we directly address the issue of compositionality. Our theory of lexical meaning will affect the general design of a semantic theory in several ways. If we view the goal of a semantic theory as being able to recursively assign meanings to expressions, accounting for phenomena such as synonymy, antonymy, polysemy, and metonymy, then compositionality depends ultimately on what the basic lexical categories of the language denote. The traditional view has been that words behave as either active functors or passive arguments. But we will see that if we change the way in which categories can denote, then the form of compositionality itself changes. Hence, if studied comprehensively, lexical semantics can be a means to reevaluate the very nature of semantic composition in language, in order to satisfy the goals of semantic theory.

First, I review some basic issues in lexical representation and present the current view on how to represent lexical ambiguity, both in theoretical and computational models. This view, incorporating “sense enumerative techniques,” distinguishes word senses on the basis of finite feature distinctions. As I argue in chapter 3, however, such an approach, makes no distinction between what Weinreich (1964) calls *contrastive* and *complementary ambiguity*.<sup>1</sup> The former is basic homonymy, where a lexical item accidentally carries several distinct and unrelated meanings, whereas

the latter refers to logically related word senses of the same lexical item. I then turn to some further problems with the enumeration method for lexical description illustrated in chapter 3. It will be shown that the representations assumed by current theories are inadequate to account for the richness of natural language semantics.

As I show in chapters 2 and 3, most of the careful representation work has been done on verb classes (e.g., Levin, 1993). In fact, the semantic weight in both lexical and compositional terms usually falls on the verb. This has obvious consequences for how lexical ambiguity has been treated. In chapter 4, I discuss several devices which simplify our semantic description, but which fall outside the conception of enumerative lexical semantics. Looking at these devices closely, we notice that they point to a very different view of lexical semantics and how word meanings are combined.

Given the discussion in these chapters, the following conception of lexical semantic systems emerges. Under such a theory, a core set of word senses, typically with greater internal structure than is assumed in previous theories, is used to generate a larger set of word senses when individual lexical items are combined with others in phrases and clauses. I will refer to such an organization as a *generative lexicon*, and the operations which generate these “extended senses” as *generative devices*, including operations such as *type coercion* and *co-composition*. I discuss how this view supports an explanatory view of semantic modeling. I then examine the goals of linguistic theory in general and lexical semantics in particular. I argue that our framework of knowledge for lexical items must be guided by a concern for semanticity in addition to grammaticality. The model of semantic interpretation we construct should reflect the particular properties and difficulties of natural language, and not simply be an application of a ready-to-wear logical formalism to a new body of data. I will view natural languages as positioned on a hierarchy of semantic descriptions, characterized in terms of their underlying polymorphic generative power. I argue that natural languages fall within the *weakly polymorphic* languages, more expressive than *monomorphic*, but well below the power of unrestricted polymorphic languages. This particular characterization is rich enough to capture the behavior of logical polysemy as well as effects of co-compositionality.

Next, in chapter 5, I outline the type system for our semantics. A generative theory of the lexicon includes multiple levels of representation for

the different types of lexical information needed. Among such levels are Argument Structure (for the representation of adicity information for functional elements), Event Structure (for the representation of information related to Aktionsarten and event type, in the sense of Vendler, 1967, and related work), Qualia Structure (for the representation of the defining attributes of an object, such as its constituent parts, purpose and function, mode of creation, etc.), and Inheritance Structure (for the representation of the relation between the lexical item and others in the lexicon). Chapter 6 presents in more detail the structure of qualia, and the role they play in distributing the functional behavior of words and phrases in composition.

Chapter 7 presents the application of the mechanisms outlined in chapters 5 and 6 to the polymorphic behavior of language. A variety of polymorphic types is studied and I consider what operations are needed to adequately account for the syntactic expressiveness of semantic types. In particular, I examine the role of coercion in the grammar as well as the need for other generative devices, such as selective binding and co-composition. There is no single form of polymorphism; rather, polysemy and type ambiguity are a result of several semantic phenomena in specific interaction.

Chapter 8 examines briefly what the consequences of qualia structure are for the semantics of nominals. Nouns can be formally characterizable in terms of three dimensions of analysis, involving argument structure, event type, and qualia structure. An analysis of nominal polysemy is presented, making use of the type system outlined in the previous chapters, and explaining in more detail the distinction between unified types and dot objects.

In the next two chapters, I outline some areas of grammar that can be greatly simplified if we apply to them principles of generative lexical analysis through the use of the generative devices and the type system presented in chapter 5. In particular, I treat argument selection as driven by semantic types, modulated by constraints on coercion rules, selective binding, and co-composition operations in the grammar. This approach will permit us to explain the polymorphic nature of verbs taking multiple syntactic types. In chapter 9, I discuss the role that qualia and event structure have in describing the way causal relations are lexicalized in language. Specifically, I look at the semantics of causative/inchoative verbs, aspectual predicates, experiencer predicates, and *modal causatives*

such as *risk*.

Finally, I discuss how this view of lexical organization relates to current theories of metaphor and pragmatically-induced metonymy. I argue, on methodological grounds, for a strong distinction between common-sense knowledge and lexical structure, although the issue is clearly an empirical one. The types of creative polysemy examined in this work exhibit a regularity and systematicity across languages that is absent from patterns of pragmatic sense extension or modes of metaphor.

## 2 The Nature of Lexical Knowledge

Only a few years ago, it was conventional practice in both theoretical and computational linguistics textbooks to cover all that needed to be said regarding the lexicon in one quick chapter, before getting to the more interesting and substantive topics of syntactic form and semantic interpretation. Such an impoverished coverage today would scarcely reflect the vibrancy of the field of lexical research or the central role played by lexical knowledge in linguistic theory and processing models. It is now standardly assumed by most linguistic frameworks (both computational and theoretical) that much of the structural information of a sentence is best encoded from a lexicalized perspective.<sup>1</sup>

The most pressing problems for lexical semantics, I believe, are the following:

- (a) Explaining the *polymorphic nature* of language;
- (b) Characterizing the *semanticity* of natural language utterances;
- (c) Capturing the *creative use of words* in novel contexts;
- (d) Developing a richer, *co-compositional* semantic representation.

I believe we have reached an interesting turning point in research, where linguistic studies can be informed by computational tools for lexicology as well as an appreciation of the computational complexity of large lexical databases. Likewise, computational research can profit from an awareness of the grammatical and syntactic distinctions of lexical items; natural language processing (NLP) systems must account for these differences in their lexicons and grammars. The wedding of these disciplines is so important, in fact, that I believe it will soon be difficult to carry out serious computational research in the fields of linguistics and NLP without the help of electronic dictionaries and computational lexicographic resources (cf. Zampolli and Atkins, 1994, Boguraev and Briscoe, 1988). Positioned at the center of this synthesis is the study of word meaning, lexical semantics.

Before addressing these questions, I would like to discuss two assumptions that will figure prominently in my suggestions for a lexical semantics framework. The first is that, without an appreciation of the syntactic structure of a language, the study of lexical semantics is bound to fail. There is no way in which meaning can be completely divorced from the structure that carries it. This is an important methodological



point, since grammatical distinctions are a useful metric in evaluating competing semantic theories.

The second point is that the meanings of words should somehow reflect the deeper conceptual structures in the cognitive system, and the domain it operates in. This is tantamount to stating that the semantics of natural language should be the image of nonlinguistic conceptual organizing principles, whatever their structure.

Computational lexical semantics should be guided by the following principles. First, a clear notion of semantic well-formedness will be necessary in order to characterize a theory of possible word meaning. This may entail abstracting the notion of lexical meaning away from other semantic influences. For instance, this might suggest that discourse and pragmatic factors should be handled differently or separately from the semantic contributions of lexical items in composition.<sup>2</sup> Although this is not a necessary assumption and may in fact be wrong, it will help narrow our focus on what is important for lexical semantic descriptions.

Secondly, lexical semantics must look for representations that are richer than thematic role descriptions (cf. Gruber, 1965, Fillmore, 1968). As argued in Levin and Rappaport (1986), named roles are useful at best for establishing fairly general mapping strategies to the syntactic structures in language. The distinctions possible with thematic roles are much too coarse-grained to provide a useful semantic interpretation of a sentence. What is needed, I will argue, is a principled method of lexical decomposition. This presupposes, if it is to work at all, (1) a rich, recursive theory of semantic composition, (2) the notion of semantic well-formedness mentioned above, and (3) an appeal to several levels of interpretation in the semantics (cf. Scha, 1983).

Thirdly, and related to the preceding point, lexical semantics must study all syntactic categories in order to characterize the semantics of natural language. That is, contrary to the recent trends in semantic representation, the lexicon must encode information for categories other than verbs. Recent work has done much to clarify the nature of verb classes and the syntactic constructions that each allows (cf. Levin 1985, 1993). Yet it is not clear whether we are any closer to understanding the underlying nature of verb meaning, why the classes develop as they do, and what consequences these distinctions have for the rest of the lexicon and grammar. The curious thing is that there has been little attention paid to the other lexical categories (but cf. Miller and Johnson-Laird,