

Linguisticæ Investigationes  
Supplementa 28

# *Lexical-Semantic Relations*

Theoretical and practical perspectives

Edited by  
Petra Storjohann

JOHN BENJAMINS PUBLISHING COMPANY

# LEXICAL-SEMANTIC RELATIONS

THEORETICAL AND PRACTICAL  
PERSPECTIVES

Edited by  
PETRA STORJOHANN  
Institut für Deutsche Sprache, Mannheim



JOHN BENJAMINS PUBLISHING COMPANY  
AMSTERDAM/PHILADELPHIA



TM

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

### Library of Congress Cataloging-in-Publication Data

Lexical-semantic relations : theoretical and practical perspectives / edited by Petra Storjohann.

p. cm. -- (Linguisticae investigationes. Supplementa ISSN; 0165-7569; v. 28)

Includes bibliographical references and index.

1. Semantics. 2. Lexicology. 3. Relational grammar. I. Storjohann, Petra, 1972-

P325.L482 2010

401'.43--dc22

2010009954

ISBN 978 90 272 3138 3 (Hb: alk. paper)

ISBN 978 90 272 8816 5 (Eb)

© 2010 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. • P.O.Box 36224 • 1020 ME Amsterdam • The Netherlands

John Benjamins North America • P.O.Box 27519 • Philadelphia PA 19118-0519 • USA

## LEXICAL-SEMANTIC RELATIONS

# LINGVISTICÆ INVESTIGATIONES: SUPPLEMENTA

*Studies in French & General Linguistics /  
Études en Linguistique Française et Générale*

This series has been established as a companion  
series to the periodical "LINGVISTICÆ INVESTIGATIONES",  
which started publication in 1977.

## *Series Editors:*

Éric Laporte (Université Paris-Est Marne-la-Vallée & CNRS)  
Annibale Elia (Università di Salerno)  
Gaston Gross (Université Paris-Nord & CNRS)  
Elisabete Ranchhod (Universidade de Lisboa)

Volume 28

Petra Storjohann (ed.)  
*Lexical-Semantic Relations. Theoretical and practical perspectives*

## Preface

The availability of corpus-guided methods and the emergence of new semantic models, particularly cognitive and psycholinguistic frameworks, have prompted linguists to develop a range of immensely fruitful new approaches to sense relations. Not only does the field of sense relations have immediate relevance for the study of paradigmatic structures in lexicology, it is also a much discussed field for a variety of other application-oriented areas such as lexicography, Natural Language Processing and database engineering of lexical-semantic webs. It was in this context that the *Institut für Deutsche Sprache* in Mannheim (Germany) held an international colloquium from 5th–6th June 2008 on the subject of “Lexical-Semantic Relations from Theoretical and Practical Perspectives”. This event brought together researchers with an interest in semantic theory and experts with a more practical, application-based background looking at different languages.

The papers in this volume derive from the colloquium and address specific semantic, lexicographic, computational and technological approaches to a range of meaning relations, particularly those which have traditionally been classified as “paradigmatic” sense relations, as well as exploring the construction, representation, retrieval and documentation of relations of contrast and meaning equivalence in a variety of languages including German, English and Swedish. This book provides specialists from different disciplines and areas with the opportunity to gain an insight into current cross-linguistic research in semantics, corpus and computer linguistics, lexicology, applied teaching and learning, and lexical typology as well as technological applications such as computational lexical-semantic wordnets.

The overall aim of this book is to make up for some of the shortcomings of more traditional and often non-empirical studies, by providing an overview of current theoretical perspectives on lexical-semantic relations and presenting recent application-oriented research. Above all, its aim is to stimulate dialogue and revive discussion on sense relations in general, a subject which requires reappraisal in the light of recent semantic theories and which merits application via contemporary linguistic/lexicographic methods and procedures.

I am appreciative of the help of all authors who have contributed to this book and who have clearly demonstrated the tremendous scope of the field and the importance of current trends in the study of paradigmatic structures. I also wish to thank some colleagues and friends for their criticisms, their help and support. My gratitude also goes to that handful of very special people for their understanding, their sense and sensitivity when making this book. Beyond these, a sincere thanks also goes out to the *Institut für Deutsche Sprache Mannheim*, for hosting the colloquium, thereby enabling semanticists, lexicographers, experts in Natural Language Programming and computer linguists to share their common interest in lexical-semantic relations.

Petra Storjohann (Institut für Deutsche Mannheim)

# Table of contents

Preface	VII
Introduction <i>Petra Storjohann</i>	1
Lexico-semantic relations in theory and practice <i>Petra Storjohann</i>	5
Swedish opposites: A multi-method approach to ‘goodness of antonymy’ <i>Caroline Willners and Carita Paradis</i>	15
Using web data to explore lexico-semantic relations <i>Steven Jones</i>	49
Synonyms in corpus texts: Conceptualisation and construction <i>Petra Storjohann</i>	69
Antonymy relations: Typical and atypical cases from the domain of speech act verbs <i>Kristel Proost</i>	95
An empiricist’s view of the ontology of lexical-semantic relations <i>Cyril Belica, Holger Keibel, Marc Kupietz and Rainer Perkuhn</i>	115
The consistency of sense-related items in dictionaries: Current status, proposals for modelling and applications in lexicographic practice <i>Carolin Müller-Spitzer</i>	145
Lexical-semantic and conceptual relations in GermaNet <i>Claudia Kunze and Lothar Lemnitzer</i>	163
Index	185



# Introduction

Petra Storjohann

This collective volume focuses on what have traditionally been termed the “paradigmatics” or “sense relations” of a lexical unit. These include relations of contrast and opposition, meaning equivalence, hyponymy, hyperonymy etc., all of which have captured the interest of researchers from a range of disciplines. In the existing literature, studies on sense relations often just cover one specific phenomenon of one specific language, stressing specific semantic or methodological aspects. The present book covers different languages and different paradigmatic phenomena. It outlines the full complexity of the subject, combining linguistic and methodological elucidations with discussions on current practical and application-oriented research.

The papers in this volume which are concerned with lexicological questions examine a range of linguistic models and semantic modelling, and the use of data such as corpora or the Internet as a lexical resource for information retrieval. Various authors demonstrate that research on language and cognition calls for evidence from different sources. They explain the nature of different lexical resources and the working methods associated with them, and they suggest some theoretical implications of a larger semantic model. The lexicological papers have as a common theme contextualised and dynamically constructed structures and look at the phenomenon as it is governed by conventional, contextual and cognitive constraints which favour specific choices. The semantic approaches used here concentrate on questions of mental representation, linguistic conventionalisation, cognitive processes and ideas of constructions, and they respond to the opportunities presented by methodologies from psycholinguistics or corpus studies.

Moreover, the book explores recent developments in building large lexical resources as well as some lexicographic and text-technological aspects. These include for example elucidations on the structure and possible applications of reference databases such as GermaNet in natural language processing and computational linguistics. This database has been constructed in the style of its English counterpart WordNet and it is an integral part of EuroWordNet. This volume

is also concern with sense-related items in dictionaries. Recent insights into the paradigmatics of a word from a lexical-semantic point make a compelling case for dictionary makers to include more appropriate and innovative descriptions of sense-related items in reference works. Additionally, new technical standards and text-technological facilities are still largely being ignored by lexicographers although they offer opportunities to enhance dictionaries and to make them more consistent. These too are some of the concerns of this book.

Together, these papers not only give an impression of the scope of the field by looking at lexico-semantic relations from a range of positions and for different purposes, but they also demonstrate how cross-linguistic examinations benefit each other, how research areas fertilise and complement each other and how results in one field have an impact on research in other disciplines, each enriching the insights and developments of the others.

In “Lexico-semantic relations in theory and practice” **Petra Storjohann** provides an overview on the subject of sense relations in different linguistic fields such as lexicology, corpus studies, lexicography and computer linguistics. The paper particularly focuses on the shift of approaches and methodologies to lexico-semantic relations in lexical semantics and concentrates on perspectives taken in German and in English linguistics. The paper is thought as a general discussion on the subject and it reveals some of the open questions and the challenges that need to be approached in the future.

Both psycholinguistic and corpus-linguistic approaches are taken in the paper “Swedish opposites. A multi-method approach to goodness of antonymy” by **Caroline Willners** and **Carita Paradis** where the nature of English and Swedish antonymy and the degree of conventionalisation of antonymic word pairs in language and in memory are examined. Methodologically, their analyses are conducted on the basis of data from textual co-occurrence, and from judgement and elicitation experiments. Both types of evidence are used as a means of substantiating semantic theories. The paper not only examines differences in applying various methods, but also addresses the meanings of conventionalised canonical antonym pairings, including issues such as dimensional clarity, symmetry and contextual range. In terms of theoretical implications, it is argued that opposite meaning is construed and the authors show this with help of both highly conventionalised and less conventionalised binary opposites.

Whether the Internet can be used as a valid corpus source for linguistic analyses is a question addressed by **Steven Jones** in “Using web data to explore lexico-semantic relations”. Taking English antonymy as an example, he explores how the web can be used as a lexical resource to reveal and quantify relational structures, and raises the question of whether prior semantic statements on canonicity can be based on such a methodology. He starts out from the assumption that specific

lexico-grammatical frames are characteristic contextual surroundings for lexical items in relations of antonymy. Using these as model constructions, they can be applied to identify and quantify opposite word-pairs in mass data.

How corpus data offer different insights into the manifestation of German meaning equivalence is the subject of “Synonyms in corpus texts: conceptualisation and construction” by **Petra Storjohann**. In this article, traditional descriptions of synonymy which focus on categorical systems of common semantic traits and the degree of semantic overlap are criticised. Here, it is argued that it is the specific cognitive prerequisites which establish contextual meaning equivalence that need to be analysed in order to explain the construction of synonymy, rather than attempting to subclassify this sense relation. The paper investigates principles of synonymy construction, as they appear in corpora, and concentrates on questions of how meaning equivalence might be conceptualised and construed by speakers in actual language use. In fact, this also raises the question of how these conceptualisation principles can be pinpointed, and how these are applied by speakers in discourse. With the help of corpus data, it is shown how different conventionalised conceptualisations operate on various structures construing relations of meaning equivalence. An attempt is made to demonstrate how these are grounded in experience and perception, and to explore the links between language and thought.

The subject of typical and atypical German antonymy relations is addressed in **Kristel Proost’s** “Antonymy relations: typical and atypical cases from the domain of speech act verbs”. She explores the subject of typical and atypical German examples of antonymy from the domain of speech act verbs, and discusses different classification models. The use of a large German corpus provides ample evidence for the fact that some antonyms sometimes lack typical attributes in terms of gradability. As a result, the article stresses the diverging evidence presenting different proposals for the classification of atypical opposites.

Research results on the methodology of quantitative and qualitative analysis of large corpora and the modelling of processes of linguistic interpretation at the lexical level are presented in “An empiricist’s view of the ontology of lexical-semantic relations” by **Cyril Belica, Holger Keibel, Marc Kupietz and Rainer Perkuhn**. They address questions of how lexical-semantic relations emerge from language use and how they manifest themselves in language data. Their usage-based view covers reflections on collocation profiles, semantic proximity and self-organising maps, and explains how these can be modelled and how they are used to create topographic profiles. Although their elucidations concern German, they do not primarily focus on a specific language and, as such, their research results are readily transferable to other languages. Their findings have a threefold impact. Not only do they contribute to the general propaedeutic discussions of linguistic

theory and semantic models, they also propose an empirically-driven methodology specifically for the explorative examination of lexical-semantic relations, and they support the development of methods for empirical work with corpora and the detection of language structures in comprehensive linguistic data in more general terms.

An interesting insight into the problems of inconsistent lexicographic information is provided in **Carolín Müller-Spitzer's** paper "The consistency of sense-relations in dictionaries. Current status, proposals for modelling and applications in lexicographic practice". The paper reveals how inconsistent reference structures (e.g. 'non-reversed reference') are in German dictionaries of synonymy and antonymy. Problems such as bidirectional linking that is realised as unidirectional references are challenges for lexicographers as well as dictionary users. Although computational procedures are available to solve the problem, these have not been implemented so far. As Müller-Spitzer argues, a coherent lexicographic XML-based modelling architecture is a prerequisite for an effective data structure. With the help of *eleXiko*, a specific corpus-based, electronic reference work, she illustrates how text-technological methods can provide support for the overall consistency of sense-related pairings during the process of compiling a dictionary. Her discussion also outlines the technical requirements for achieving consistency in data-managing and data-linking.

In their article "Lexical-semantic and conceptual relations in GermaNet", **Claudia Kunze** and **Lothar Lemnitzer** discuss the relevance of some lexical as well as conceptual relations for a lexical resource of German. An overview of how these relations have been integrated in the construction and maintenance of GermaNet is given, and recent developments and their repercussions in terms of theoretical and application-oriented research are discussed. Other practical perspectives on the subject of meaning relations include the possible innovative and beneficial applications of GermaNet.

It is beyond any doubt that this book remains a comparatively brief account of a complex field with a number of issues that are not even touched upon. Nevertheless, the intention is to stimulate further discussion and promote closer collaboration between the different fields. This collective volume attempts to show what research on lexical-semantic relations has to offer and to demonstrate, as Alan Cruse (2004: 141) asserts, "that sense relations are a worthwhile object of study". As such, it is an invitation to scholars from every field whose interests involve words, meaning, the mind and/or language technology and who have a shared interest in lexical-semantic relations and approach the subject from various positions in philosophy, psychology, neuroscience, linguistics, lexicography, computer science, early childhood language acquisition and second language education, to name but a few.

# Lexico-semantic relations in theory and practice

Petra Storjohann

This paper provides a general overview of the treatment of lexico-semantic relations in different fields of research including theoretical and application-oriented disciplines. At the same time, it sketches the development of the descriptions and explanations of sense relations in various approaches as well as some methodologies which have been used to retrieve and analyse paradigmatic patterns.

## 1. Lexicology: From structural to cognitive approaches

### 1.1 Structuralist approaches

From a lexicological point of view, the subject of sense relations has long been closely linked with several traditions of structural semantics and lexical field analysis, particularly within German linguistics. For decades, the theory of lexical field analysis was a very popular area of research, reaching its peak in the 1970s and 80s.<sup>1</sup> Hence, it is automatically associated with the classical notion of the study of a language system, with atomised and isolated approaches, and the semantics of lexemes in terms of distinctive features. The emphasis is simultaneously on fixed and inherent semantic properties, componential meaning analysis and the idea that meaning can be neatly decomposed and described. The view was held that language is as an “externalized object” (Paradis 2009) with clearly recognisable structures. Sense relations were of particular interest since the basic assumption was that lexical meaning is constituted by the relations a lexeme holds with other lexemes in the same lexical-semantic paradigm. Structuralists not only made use of language as a system but also referred to lexical relations in terms of paradigmatic and syntagmatic structures implying strict distinctions between them.

---

1. For an overview on the development of lexical field theory see Storjohann (2003:25–40).

Formalist linguists sought to define the meaning of lexical items by decompositional approaches, which worked well for modelling structural aspects such as phonology or syntax. But classical decompositional theories suffered from a number of problems, above all the belief that vocabulary has a definitional structure with distinct boundaries that can be precisely delimited. The traditional conception of sense relations was that of semantic connections between words and semantic interrelations among members of paradigmatic sets were viewed as stable and context-independent structures. Today, as a result, the phenomenon of sense relations is stigmatised and too closely linked to traditional or old-fashioned models. In German linguistics in particular, where once research on lexical-semantic structures flourished, the chapter on sense-related lexical terms was essentially closed by the works of Lutzeier (1981, 1985). His studies not only offered systematic examinations of lexical fields and their sense relations but they also made use of a stringent terminology and introduced the notion of contextual restrictions by bringing in key elements such as verbal context, syntactic category and semantic aspects. Particularly his later work pointed out the discrepancy between structuralist descriptions and textual structures in language use and prepared the ground for more empirical research on lexico-semantic relations in actual discourse. Nonetheless, the perception that the subject is obsolete persists to the present day and German semanticists have not further contributed to the general discussion on sense-related items in more recent contexts.

In contrast, the situation with regard to semantics was never as bleak in the case of English linguistics, where scholars around the world were not so keen to avoid the subject of sense relations after the decline of the structuralist period. Lexical semantics in general has thrived in the UK, and its tradition is best exemplified by names such as John Lyons and Alan Cruse, both of whom have developed exhaustive definitions and descriptions of semantic relations.<sup>2</sup> For them, the study of sense relations was central to the study of meaning. At the same time, the Firthian tradition developed which concentrated on syntagmatic relations.<sup>3</sup> Collocations become the key notion and later the centre of attention to corpus linguists. Generally, more contextualised approaches to sense relations were encountered at that time with Cruse's (1986) approach as a central piece of work in the tradition of the British Contextualism and consequently, it is studies on the English language which have succeeded in further advancing theories about lexico-semantic relations.

---

2. Cf. Lyons (1968, 1977) and Cruse (1986).

3. "You shall know the meaning of a word by the company it keeps" (Firth 1957: 179).

## 1.2 Cognitive approaches

As the notion of the lexicon started to be of interest to approaches to syntax which left behind the division between grammar and lexis, the nature of lexical semantics and the notion of the mental lexicon changed. New methodologies were introduced which looked at language from a usage-based perspective. However, corpus linguistics has largely focused its efforts on collocations and co-occurrences, and although linguistic theories have progressed, particularly in the area of cognitive linguistics, most semantic research has centred around issues such as polysemy and metaphor. And although the cognitive strand generally has had a major impact on lexical studies (cf. Geeraerts/Cuyckens 2007), the study of sense relations has not been a central component in the new semantic paradigm, and, as Cruse (2004: 141) concludes, “cognitive linguists, for the most part, have had very little to say on the topic”. Throughout his later work, Cruse himself has been concerned with bringing the cognitive aspect into his theory of meaning (Cruse 1992; Cruse/Togia 1995; Croft/Cruse 2004), unfortunately without incorporating new methodological approaches to substantiate his ideas.

New guiding principles, assumptions and foundational hypotheses have become points of departure for semantic research in general, and they have gradually been transferred to the understanding of how sense relations are established in text and discourse. These concern how meaning is constructed. According to the cognitive school, meaning construction is equated with knowledge representation, categorisation and conceptualisation. Meaning is a process, it is dynamic, and it draws upon encyclopaedic knowledge and the subject of sense relations has started to be re-examined from a cognitive point of view. We now have a different understanding of how semantic relations are mentally represented and linguistically expressed, notions that are owed to the proliferation of research in the field of cognitive linguistics. Today, a number of linguists, mostly outside German linguistics, with a particular interest in lexical semantic relations, have reopened the chapter on sense relations offering new perspectives, employing new methodologies and using empirical evidence for their work. In particular, the *Group for Comparative Lexicology*<sup>4</sup> has sought to advance theories around English and Swedish lexical relations. They have succeeded in showing how sense relations materialise in text and discourse. The question of whether sense relations are lexical relations, or rather conceptual-semantic relations, or relations among contextual construals, has been addressed. As a result, classical notions of the paradigmatics

---

4. Steven Jones (e.g. 2002), Lynne Murphy (e.g. 2003, 2006), Carita Paradis (e.g. 2005) and Caroline Willners (e.g. 2001) are particularly concerned with the study of English and Swedish opposites.

of a lexical item have largely been abandoned. Recent semantic theories now account for lexical-semantic relations and are capable of accommodating all kinds of relations “ranging from highly conventionalized lexico-semantic couplings to strongly contextually motivated pairings” (Paradis forthcoming).

### 1.3 Corpus material and language in use

As Paradis (forthcoming) notes, it is methodologies which have radically changed studies on meaning and semantic relations. The basis of investigations is now determined by corpus procedures, by different observational and experimental techniques and by computational facilities and these contribute profitably to insights into the nature of the paradigmatics. A particularly promising trend within the new linguistic context is the fact that recent theories have also brought lexical semantics, and thus the subject of lexical-semantic relations, much closer to language in use and thought. Through the use of corpora, for example, we gain a different notion of language as it emerges from language use. The central function of language as a means of natural communication and its role in social interaction are no longer ignored. Conclusions are drawn not on the basis of intuitive judgement, but from real data and on the basis of mass data which account for recurrence, variability and the distribution of patterns. Generally, semanticists from various schools of thought have for a long time proved to be immune to corpus methods, and it is only recently that some researchers have made a compelling case for incorporating methods of corpus linguistics into semantics. This is all the more astonishing since both cognitive linguists and corpus linguists share an interest in contextualised, dynamically constructed meaning and in the grounding of language use in cognitive and social-interactional processes.

Language in natural communicative situations involving speakers and addressees has come to occupy the seat of honour in cognitive linguistic research and the combination of the theoretical and empirical developments has sparked new interest in research on lexico-semantic relations and their functions in language and thought. (Paradis forthcoming)

In terms of empirical corpus studies, it is however predominantly the subject of English opposites that has attracted interest from a corpus-based perspective (e.g. Jones 2002; Murphy 2006), demonstrating how corpus evidence can be used to derive semantic models. Until now, corpus-oriented studies of sense relations have been rather few and far between. However, systematic corpus-guided investigations have shown that corpus methodologies have contributed greatly to the study of lexical-semantic paradigms, and yielded new insights into issues such as how these relational patterns behave and function in discourse.



## 2. Lexicography

The field where findings on semantic relations demand to be accounted for and where they are of potential utility is lexicography. Sense relations are documented in dictionaries of synonymy and antonymy or in onomasiological reference books such as a thesaurus. There is a striking clash between the findings of theoretical semantic research on the one hand, and the commercial and practical missions of dictionaries on the other hand. Dictionary entries provide lists of sense-related items which are treated as stable relations between words, often not even assigned to a specific sense. And however inappropriate and inconsistent the representations of the facts about a word and its relations might be, it seems impossible to make a reference book radically different. The pressure of a dictionary is to present definite answers and clear-cut definitions. Hence, often sets of discrete synonyms or antonyms are given for words without overlapping meanings. Although it is commonly agreed that the construction of lexico-semantic relations is flexible, lexicographers continue to offer only vague descriptions and struggle to present meaning, and hence sense relations, as context-dependent, variable and dynamic.

In addition, although corpora have been available for some time now, the exploration of mass data and the use of corpus tools for lexicographic analysis are restricted to corpus-based investigations, leaving a pool of linguistic evidence to be used for acts of verification only. Corpus-driven methodologies, however, where the corpus is approached without any prior assumptions and where collocation profiles reveal insights into the use of sense-related items, are largely ignored.<sup>5</sup> As a result, as Alan Cruse comments:

No one is puzzled by the contents of a dictionary of synonymy, or by what lexicographers in standard dictionaries offer by way of synonyms, even though the great majority of these qualify neither as absolute nor as propositional synonyms.

(Cruse 2004: 156)

An analysis of dictionary consultations by Harvey and Yuill<sup>6</sup> in 1994 showed that in 10% of cases, users were looking for meaning equivalent terms. In over 36% of these situations, users were left without answers, or the information given was not satisfactory. Information on contextual conditions and situational usage was lacking. No other type of search showed the same degree of dissatisfaction. Users do

---

5. For further differences between corpus-based and corpus-driven methodologies see Tognini-Bonelli (2001).

6. Unpublished research paper quoted in Partington (1998: 29).