

1. Record Nr.	UNINA9910787912603321
Autore	Kuebler Sandra
Titolo	Corpus Linguistics and Linguistically Annotated Corpora // Sandra Kuebler and Heike Zinsmeister
Pubbl/distr/stampa	London ; ; New York : , : Bloomsbury Academic, , [2015]
ISBN	1-4411-1980-9 1-4411-6447-2 1-4725-9357-X 1-4411-1991-4
Descrizione fisica	1 online resource (321 p.)
Classificazione	LAN009000
Disciplina	410.1/88
Soggetti	Corpora (Linguistics) Computational linguistics
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Bibliographic Level Mode of Issuance: Monograph
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Machine generated contents note: Preface -- Part I Introduction 1. Corpus Linguistics ; 2. Corpora and Linguistic Annotation -- Part II Linguistic Annotation ; 3. Linguistic Annotation on the Word Level ; 4. Syntactic Annotation ; 5. Semantic Annotation ; 6. Discourse Annotation -- Part III Using Linguistic Annotation in Corpus Linguistics. 7. Advantages and Limitations of Using Linguistically Annotated Corpora ; 8. Corpus Linguistics Using Linguistically Annotated Corpora -- Part IV Querying Linguistically Annotated Corpora. 9. Concordances ; 10. Regular Expressions ; 11. Searching on the Word Level ; 12. Querying Syntactic Structures ; 13. Searching for Semantic and Discourse Phenomena -- Appendix A. Penn Treebank POS Tagset 343 -- Appendix B. ICE POS Tagset 345 -- Bibliography -- Index.
Sommario/riassunto	"Linguistically annotated corpora are becoming a central part of the corpus linguistics field. One of their main strengths is the level of searchability they offer, but with the annotation come problems of the initial complexity of queries and query tools. This book gives a full, pedagogic account of this burgeoning field.Beginning with an overview of corpus linguistics, its prerequisites and goals, the book then introduces linguistically annotated corpora. It explores the different

levels of linguistic annotation, including morphological, parts of speech, syntactic, semantic and discourse-level, as well as advantages and challenges for such annotations. It covers the main annotated corpora for English, the Penn Treebank, the International Corpus of English, and OntoNotes, as well as a wide range of corpora for other languages. In its third part, search strategies required for different types of data are explored. All chapters are accompanied by exercises and by sections on further reading, together with an integral companion website that contains lists and guidance on contemporary annotated corpora and query tools."--Bloomsbury Publishing.
