

1. Record Nr.	UNINA9910464192003321
Autore	Sampson Geoffrey <1944->
Titolo	Corpus Linguistics [[electronic resource] ] : Readings in a Widening Discipline
Pubbl/distr/stampa	London, : Bloomsbury Publishing, 2005
ISBN	1-4411-3937-0
Descrizione fisica	1 online resource (541 p.)
Collana	Open linguistics series Corpus linguistics
Altri autori (Persone)	McCarthyDiana
Disciplina	410
Soggetti	Computational linguistics Language and languages Linguistics Languages & Literatures Philology & Linguistics Electronic books.
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Description based upon print version of record.
Nota di contenuto	Contents; Sources and acknowledgements; Abbreviations used in this book; 1 Introduction; 2 From The Structure of English (1952); 3 A standard corpus of edited present-day American English (1965); 4 On the distribution of noun-phrase types in English clause-structure (1971); 5 Predicting text segmentation into tone units (1986); 6 Typicality and meaning potentials (1986); 7 Historical drift in three English genres (1987); 8 Corpus creation (1987); 9 Cleft and pseudo-cleft constructions in English spoken and written discourse (1987); 10 What is wrong with adding one? (1989) 11 A statistical approach to machine translation (1990)12 A point of verb syntax in south-western British English: an analysis of a dialect continuum (1991); 13 Using corpus data in the Swedish Academy grammar (1991); 14 On the history of that/zero as object clause links in English (1991); 15 Encoding the British National Corpus (1992); 16 Computer corpora - what do they tell us about culture? (1992); 17 Representativeness in corpus design (1992); 18 A corpus-driven approach to grammar: Principles, Methods, and Examples (1993); 19 Structural ambiguity and lexical relations (1993)

20 Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies (1993); 21 Building a large annotated corpus of English: the Penn Treebank (1993); 22 Automatically extracting collocations from corpora for language learning (1994); 23 Developing and evaluating a probabilistic LR parser of part-of-speech and punctuation labels (1995); 24 Why a Fiji corpus? (1996); 25 Treebank grammars (1996); 26 English corpus linguistics and the foreign-language teaching syllabus (1996); 27 Data-oriented language processing: an overview (1996)  
28 Conflict talk: A comparison of the verbal disputes between adolescent females in two corpora (1996); 29 Assessing agreement on classification tasks: the kappa statistic (1996); 30 Linguistic and interactional features of Internet Relay Chat (1996); 31 Distinguishing systems and distinguishing senses: New evaluation methods for word-sense disambiguation (1997); 32 Qualification and certainty in L1 and L2 students' writing (1997); 33 Analysing and predicting patterns of DAMSL utterance tags (1998); 34 Assessing claims about language use with corpus data - swearing and abuse (1998)  
35 The syntax of disfluency in spontaneous spoken language (1998); 36 The use of large text corpora for evaluating text-to-speech systems (1998); 37 The Prague Dependency Treebank: how much of the underlying syntactic structure can be tagged automatically? (1999); 38 Reflections of a dendrographer (1999); 39 A generic approach to software support for linguistic annotation using XML (2000); 40 Europe's ignored languages (2001); 41 Semi-automatic tagging of intonation in French spoken corpora (2001); 42 Web as corpus (2001); 43 Intonational variation in the British Isles (2002); Bibliography  
URL List

---

## Sommario/riassunto

Corpus Linguistics seeks to provide a comprehensive sampling of real-life usage in a given language, and to use these empirical data to test language hypotheses. Modern corpus linguistics began fifty years ago, but the subject has seen explosive growth since the early 1990s. These days corpora are being used to advance virtually every aspect of language study, from computer processing techniques such as machine translation, to literary stylistics, social aspects of language use, and improved language-teaching methods. Because corpus linguistics has grown fast from small beginnings, newcomers t

---