Thank you extremely much for downloading *Corpus Linguistics And Linguistically Annotated Corpora Heike Zinsmeister*. Maybe you have knowledge that, people have see numerous time for their favorite books behind this *Corpus Linguistics And Linguistically Annotated Corpora Heike Zinsmeister*, but stop happening in harmful downloads.

Rather than enjoying a good book later a cup of coffee in the afternoon, then again they juggled taking into account some harmful virus inside their computer. *Corpus Linguistics And Linguistically Annotated Corpora Heike Zinsmeister* is clear in our digital library an online admission to it is set as public correspondingly you can download it instantly. Our digital library saves in compound countries, allowing you to acquire the most less latency epoch to download any of our books behind this one. Merely said, the *Corpus Linguistics And Linguistically Annotated Corpora Heike Zinsmeister* is universally compatible subsequent to any devices to read.

*Corpus Linguistics* Anke Lüdeling 2008 The series Handbooks of Linguistics and Communication Science is designed to illuminate a field which not only includes general linguistics and the study of linguistics as applied to specific languages, but also covers those more recent areas which have developed from the increasing body of research into the manifold forms of communicative action and interaction.

**Verbmobil: Foundations of Speech-to-Speech Translation** Wolfgang Wahlster 2013-04-17 In 1992 it seemed very difficult to answer the question whether it would be possible to develop a portable system for the automatic recognition and translation of spontaneous speech. Previous research work on speech processing had focused on read speech only and international projects aimed at automated text translation had just been terminated without achieving their objectives. Within this context, the German Federal Ministry of Education and Research (BMBF) made a careful analysis of all national and international research projects conducted in the field of speech and language technology before deciding to launch an eight-year basic-research lead project in which research groups were to cooperate in an interdisciplinary and international effort covering the disciplines of computer science, computational linguistics, translation science, signal processing, communication science and artificial intelligence. At some point, the project comprised up to 135 work packages with up to 33 research groups working on these packages. The project was controlled by means of a network plan. Every two years the project situation was assessed and the project goals were updated. An international scientific advisory board provided advice for BMBF. A new scientific approach was chosen for this
project: coping with the complexity of spontaneous speech with all its pertinent phenomena such as ambiguities, self-corrections, hesitations and disfluencies took precedence over the intended lexicon size. Another important aspect was that prosodic information was exploited at all processing stages.

**Corpus Linguistics** Douglas Biber 1998-04-23 This book is about investigating the way people use language in speech and writing. It introduces the corpus-based approach to linguistics, based on analysis of large databases of real language examples stored on computer. Each chapter focuses on a different area of linguistics, including lexicography, grammar, discourse, register variation, language acquisition, and historical linguistics. Example analyses are presented in each chapter to provide concrete descriptions of the research methods and advantages of corpus-based techniques. Ten methodology boxes provide clear and concise explanations of the issues in doing corpus-based research and reading corpus-based studies and there is a useful appendix of resources for corpus-based investigation. This lucid and comprehensive introduction to the subject will be welcomed by a broad range of readers, from undergraduate students to professional researchers.


**Dependency Linguistics** Kim Gerdes 2014-09-15 This volume offers the reader a unique possibility to obtain a concise introduction to dependency linguistics and to learn about the current state of the art in the field. It unites the revised and extended versions of the linguistically-oriented papers to the First International Conference on Dependency Linguistics held in Barcelona. The contributions range from the discussion of definitional challenges of dependency at different levels of the linguistic model, its role beyond the classical grammatical description, and its annotation in dependency treebanks to concrete analyses of various cross-linguistic phenomena of syntax in its interplay with phonetics, morphology, and semantics, including phenomena for which classical simple phrase-structure based models have proven to be unsatisfactory. The volume will be thus of interest to both experts and newcomers to the field of dependency linguistics and its computational applications.

**Syntax of Dutch** 2012 The Syntax of Dutch will be published in at least seven volumes in the period 2012-2016 and aims at presenting a synthesis of the currently available syntactic knowledge of Dutch. It is primarily concerned with language description and not with linguistic theory, and provides support to all researchers interested in matters relating to the syntax of Dutch, including advanced students of language and linguistics.

**Individuality in Language Change** Lynn Anthonissen 2021-10-25 Linguists have typically studied language change at the aggregate level of speech communities, yet key mechanisms of change such as analogy and automation operate within the minds of individual language users. Drawing on lifespan data from 50 authors
and the intriguing case of the special passives in the history of English, this study addresses three fundamental issues relating to individuality in language change: (i) how variation and change at the individual level interact with change at the community level; (ii) how much innovation and change is possible across the adult lifespan; (iii) and to what extent related linguistic patterns are associated in individual cognition. As one of the first large-scale empirical studies to systematically link individual- and community-based perspectives in language change, this volume breaks new ground in our understanding of language as a complex adaptive system.

**The Cambridge Handbook of Cognitive Linguistics** Barbara Dancygier 2017-06-01

The best survey of cognitive linguistics available, this Handbook provides a thorough explanation of its rich methodology, key results, and interdisciplinary context. With in-depth coverage of the research questions, basic concepts, and various theoretical approaches, the Handbook addresses newly emerging subfields and shows their contribution to the discipline. The Handbook introduces fields of study that have become central to cognitive linguistics, such as conceptual mappings and construction grammar. It explains all the main areas of linguistic analysis traditionally expected in a full linguistics framework, and includes fields of study such as language acquisition, sociolinguistics, diachronic studies, and corpus linguistics. Setting linguistic facts within the context of many other disciplines, the Handbook will be welcomed by researchers and students in a broad range of disciplines, including linguistics, cognitive science, neuroscience, gesture studies, computational linguistics, and multimodal studies.

**The Present-day English Gerund System** Charlotte Maekelberghe 2020-09-21

This study offers a first comprehensive synchronic account of the Present-day English gerundial system. Most synchronic studies of gerunds have hitherto focused on the verbal gerund, scrutinizing its categorial status or comparing it with other non-finite clausal structures. A systematic comparison with its nominal counterpart, however, is lacking. Based on a detailed empirical analysis of lexico-grammar and semantics, this study develops an innovative cognitive-constructionist model of the English gerund system which depicts the usage profiles of nominal and verbal gerunds in terms of probabilistic trends rather than by means of categorical labels. It is shown that a better understanding of the functioning of the English gerund system requires a description that operates on multiple levels, accounting for both the abstract construal gerunds can impose on an event as well as the token-level constraints on variation between the two gerund types. This multifaceted approach, it is argued, not only offers a new perspective on the configuration of ing-forms in Present-day English, it can also be of relevance to the description of other complex grammatical structures.

**New directions in corpus-based translation studies** Claudio Fantinuoli 2015-05-27

Corpus-based translation studies has become a major paradigm and research methodology and has investigated a wide variety of topics in the last two decades. The contributions to this volume add to the range of corpus-based
studies by providing examples of some less explored applications of corpus analysis methods to translation research. They show that the area keeps evolving as it constantly opens up to different frameworks and approaches, from appraisal theory to process-oriented analysis, and encompasses multiple translation settings, including (indirect) literary translation, machine (assisted)-translation and the practical work of professional legal translators. The studies included in the volume also expand the range of application of corpus applications in terms of the tools used to accomplish the research tasks outlined.

**Natural Language Processing for Historical Texts** Michael Piotrowski 2012

More and more historical texts are becoming available in digital form. Digitization of paper documents is motivated by the aim of preserving cultural heritage and making it more accessible, both to laypeople and scholars. As digital images cannot be searched for text, digitization projects increasingly strive to create digital text, which can be searched and otherwise automatically processed, in addition to facsimiles. Indeed, the emerging field of digital humanities heavily relies on the availability of digital text for its studies. Together with the increasing availability of historical texts in digital form, there is a growing interest in applying natural language processing (NLP) methods and tools to historical texts. However, the specific linguistic properties of historical texts -- the lack of standardized orthography, in particular -- pose special challenges for NLP. This book aims to give an introduction to NLP for historical texts and an overview of the state of the art in this field. The book starts with an overview of methods for the acquisition of historical texts (scanning and OCR), discusses text encoding and annotation schemes, and presents examples of corpora of historical texts in a variety of languages. The book then discusses specific methods, such as creating part-of-speech taggers for historical languages or handling spelling variation. A final chapter analyzes the relationship between NLP and the digital humanities. Certain recently emerging textual genres, such as SMS, social media, and chat messages, or newsgroup and forum postings share a number of properties with historical texts, for example, nonstandard orthography and grammar, and profuse use of abbreviations. The methods and techniques required for the effective processing of historical texts are thus also of interest for research in other domains. Table of Contents: Introduction / NLP and Digital Humanities / Spelling in Historical Texts / Acquiring Historical Texts / Text Encoding and Annotation Schemes / Handling Spelling Variation / NLP Tools for Historical Languages / Historical Corpora / Conclusion / Bibliography

**Historical Corpora** Jost Gippert 2015-03-11

The volume contains 23 papers read at the international conference "Historical Corpora 2012", which was hosted by the LOEWE Research Cluster "Digital Humanities" of the State of Hesse at the University of Frankfurt on December 6-8, 2012. The papers, which include three keynote speeches, have been duly updated for the present volume. The contributions take a broad variety of perspectives on "historical corpora", including their structuring, their management, and various other facets. In addition to this, they cover a large amount of different languages, extending
Handbook of Linguistic Annotation Nancy Ide 2017-06-16 This handbook offers a thorough treatment of the science of linguistic annotation. Leaders in the field guide the reader through the process of modeling, creating an annotation language, building a corpus and evaluating it for correctness. Essential reading for both computer scientists and linguistic researchers. Linguistic annotation is an increasingly important activity in the field of computational linguistics because of its critical role in the development of language models for natural language processing applications. Part one of this book covers all phases of the linguistic annotation process, from annotation scheme design and choice of representation format through both the manual and automatic annotation process, evaluation, and iterative improvement of annotation accuracy. The second part of the book includes case studies of annotation projects across the spectrum of linguistic annotation types, including morpho-syntactic tagging, syntactic analyses, a range of semantic analyses (semantic roles, named entities, sentiment and opinion), time and event and spatial analyses, and discourse level analyses including discourse structure, coreference, etc. Each case study addresses the various phases and processes discussed in the chapters of part one.

Challenges of Anglophone Language(s), Literatures and Cultures Alena Kačmárová 2017-01-06 This book explores scholarly challenges within the fields of Anglophone language, literature, and culture. The section focusing on language details issues falling within two areas: namely, language contact and the language-culture relationship, and stylistic and syntactic perspectives on the English language. The literature part investigates twentieth-century American, English, and Australian literature, dealing with both poetry and prose and discussing topics of identity, gender, metafiction, postmodern conditions, and other relevant theoretical issues in contemporary literature. The culture part treats theoretical approaches in cultural studies that are vital in today’s cultural context, especially in Central European universities, the Irish language and culture, and contemporary cultural phenomena inspired by the growing ubiquity of technological intrusions into various fields of cultural production.

Syntax - Theory and Analysis Tibor Kiss 2015-03-10 The handbook offers an overview of syntactic theory and analysis, in terms of different theories, different languages, and different methods. The Handbook presents the state of art in syntactic analysis, also dealing with the methodology employed, and the rules of argumentation required to achieve such analyses for a wide range of phenomena.

Cross-Linguistic Corpora for the Study of Translations Silvia Hansen-Schirra
The book specifies a corpus architecture, including annotation and querying techniques, and its implementation. The corpus architecture is developed for empirical studies of translations, and beyond those for the study of texts which are inter-lingually comparable, particularly texts of similar registers. The compiled corpus, CroCo, is a resource for research and is, with some copyright restrictions, accessible to other research projects. Most of the research was undertaken as part of a DFG-Project into linguistic properties of translations. Fundamentally, this research project was a corpus-based investigation into the language pair English-German. The long-term goal is a contribution to the study of translation as a contact variety, and beyond this to language comparison and language contact more generally with the language pair English - German as our object languages. This goal implies a thorough interest in possible specific properties of translations, and beyond this in an empirical translation theory. The methodology developed is not restricted to the traditional exclusively system-based comparison of earlier days, where real-text excerpts or constructed examples are used as mere illustrations of assumptions and claims, but instead implements an empirical research strategy involving structured data (the sub-corpora and their relationships to each other, annotated and aligned on various theoretically motivated levels of representation), the formation of hypotheses and their operationalizations, statistics on the data, critical examinations of their significance, and interpretation against the background of system-based comparisons and other independent sources of explanation for the phenomena observed. Further applications of the resource developed in computational linguistics are outlined and evaluated.

Anaphora Resolution and Text Retrieval Helene Schmolz 2015-03-30 This book covers anaphora resolution for the English language from a linguistic and computational point of view. First, a definition of anaphors that applies to linguistics as well as information technology is given. On this foundation, all types of anaphors and their characteristics for English are outlined. To examine how frequent each type of anaphor is, a corpus of different hypertexts has been established and analysed with regard to anaphors. The most frequent type are non-finite clause anaphors - a type which has never been investigated so far. Therefore, the potential of non-finite clause anaphors are further explored with respect to anaphora resolution. After presenting the fundamentals of computational anaphora resolution and its application in text retrieval, rules for resolving non-finite clause anaphors are established. Therefore, this book shows that a truly interdisciplinary approach can achieve results which would not have been possible otherwise.

Approaches to Grammaticalization Elizabeth Closs Traugott 1991-01-01 The study of grammaticalization raises a number of fundamental theoretical issues pertaining to the relation of langue and parole, creativity and automatic coding, synchrony and diachrony, categoriality and continua, typological characteristics and language-specific forms, etc., and therefore challenges some of the basic tenets of twentieth century linguistics. This two-volume work presents a number of diverse theoretical viewpoints on grammaticalization and
gives insights into the genesis, development, and organization of grammatical categories in a number of language world-wide, with particular attention to morphosyntactic and semantic-pragmatic issues. The papers in Volume I are divided into two sections, the first concerned with general method, and the second with issues of directionality. Those in Volume II are divided into five sections: verbal structure, argument structure, subordination, modality, and multiple paths of grammaticalization.

Possession and Ownership Aleksandra I︠U︡rʹevna Aĭkhenvalʹd 2013 Linguists and anthropologists explore the intriguing variety of possessive phrases denoting ownership of property, whole-part relations (such as body and plant parts), and blood and affinal kinship relations across a wide range of languages. Like others in the series this pioneering book will be equally valued in linguistics and anthropology.

Corpus Linguistics Anke Lüdeling 2009

Corpus Annotation R. G. Garside 2016-07-27 Corpus Annotation gives an up-to-date picture of this fascinating new area of research, and will provide essential reading for newcomers to the field as well as those already involved in corpus annotation. Early chapters introduce the different levels and techniques of corpus annotation. Later chapters deal with software developments, applications, and the development of standards for the evaluation of corpus annotation. While the book takes detailed account of research worldwide, its focus is particularly on the work of the UCREL (University Centre for Computer Corpus Research on Language) team at Lancaster University, which has been at the forefront of developments in the field of corpus annotation since its beginnings in the 1970s.

Topics in Language Resources for Translation and Localisation Elia Yuste Rodrigo 2008 Language Resources (LRs) are sets of language data and descriptions in machine readable form, such as written and spoken language corpora, terminological databases, computational lexica and dictionaries, and linguistic software tools. Over the past few decades, mainly within research environments, LRs have been specifically used to create, optimise or evaluate natural language processing (NLP) and human language technologies (HLT) applications, including translation-related technologies. Gradually the infrastructures and exploitation tools of LRs are being perceived as core resources in the language services industries and in localisation production settings. However, some efforts ought yet to be made to raise further awareness about LRs in general, and LRs for translation and localisation in particular to a wider audience in all corners of the world. Topics in Language Resources for Translation and Localisation sets out to establish the state of the art of this ever expanding field and underscores the usefulness that LRs can potentially have in the process of creating, adapting, managing, standardising and leveraging content for more than one language and culture from various perspectives.
The Linguistics of Newswriting Daniel Perrin 2013-09-25 The Linguistics of Newswriting focuses on text production in journalistic media as both a socially relevant field of language use and as a strategic field of applied linguistics. The book discusses and paves the way for scientific projects in the emerging field of linguistics of newswriting. From empirical micro and theoretical macro perspectives, strategies and practices of research development and knowledge transformation are discussed. Thus, the book is addressed to researchers, teachers and coaches interested in the linguistics of professional writing in general and newswriting in particular. Together with the training materials provided on the internet www.news-writing.net, the book will also be useful to anyone who wants to become a more “discerning consumer” (Perry, 2005) or a more reflective producer of language in the media.

Research Methodologies in Translation Studies Gabriela Saldanha 2014-04-08 As an interdisciplinary area of research, translation studies attracts students and scholars with a wide range of backgrounds, who then need to face the challenge of accounting for a complex object of enquiry that does not adapt itself well to traditional methods in other fields of investigation. This book addresses the needs of such scholars – whether they are students doing research at postgraduate level or more experienced researchers who want to familiarize themselves with methods outside their current field of expertise. The book promotes a discerning and critical approach to scholarly investigation by providing the reader not only with the know-how but also with insights into how new questions can be fruitfully explored through the coherent integration of different methods of research. Understanding core principles of reliability, validity and ethics is essential for any researcher no matter what methodology they adopt, and a whole chapter is therefore devoted to these issues. Research Methodologies in Translation Studies is divided into four different chapters, according to whether the research focuses on the translation product, the process of translation, the participants involved or the context in which translation takes place. An introductory chapter discusses issues of reliability, credibility, validity and ethics. The impact of our research depends not only on its quality but also on successful dissemination, and the final chapter therefore deals with what is also generally the final stage of the research process: producing a research report.

What it Takes to Talk Paul Ibbotson 2020-07-20 This book puts cognition back at the heart of the language learning process and challenges the idea that language acquisition can be meaningfully understood as a purely linguistic phenomenon. For each domain placed under the spotlight - memory, attention, inhibition, categorisation, analogy and social cognition - the book examines how they shape the development of sounds, words and grammar. The unfolding cognitive and social world of the child interacts with, constrains, and predicts language use at its deepest levels. The conclusion is that language is special, not because it is an encapsulated module separate from the rest of cognition, but because of the forms it can take rather than the parts it is made of, and because it could be nature’s finest example of cognitive recycling and reuse.
The Oxford Handbook of Endangered Languages  
Kenneth L. Rehg 2018-07-18  
The endangered languages crisis is widely acknowledged among scholars who deal with languages and indigenous peoples as one of the most pressing problems facing humanity, posing moral, practical, and scientific issues of enormous proportions. Simply put, no area of the world is immune from language endangerment. The Oxford Handbook of Endangered Languages, in 39 chapters, provides a comprehensive overview of the efforts that are being undertaken to deal with this crisis. A comprehensive reference reflecting the breadth of the field, the Handbook presents in detail both the range of thinking about language endangerment and the variety of responses to it, and broadens understanding of language endangerment, language documentation, and language revitalization, encouraging further research. The Handbook is organized into five parts. Part 1, Endangered Languages, addresses the fundamental issues that are essential to understanding the nature of the endangered languages crisis. Part 2, Language Documentation, provides an overview of the issues and activities of concern to linguists and others in their efforts to record and document endangered languages. Part 3, Language Revitalization, includes approaches, practices, and strategies for revitalizing endangered and sleeping ("dormant") languages. Part 4, Endangered Languages and Biocultural Diversity, extends the discussion of language endangerment beyond its conventional boundaries to consider the interrelationship of language, culture, and environment, and the common forces that now threaten the sustainability of their diversity. Part 5, Looking to the Future, addresses a variety of topics that are certain to be of consequence in future efforts to document and revitalize endangered languages.

Creativity in English as a Lingua Franca  
Marie-Luise Pitzl 2018-07-09  
This book investigates formal characteristics and discourse functions of linguistic creativity at the level of idioms in spoken ELF as represented in the Vienna-Oxford International Corpus of English (VOICE). Building on the findings of previous ELF research, the book proposes that creativity might serve as a fundamental concept in accounting for the variation that seems to be central to describing and understanding English as a lingua franca.

Multilayer Corpus Studies  
Amir Zeldes 2018-07-11  
This volume explores the opportunities afforded by the construction and evaluation of multilayer corpora, an emerging methodology within corpus linguistics that brings about multiple independent parallel analyses of the same linguistic phenomena, and how the interplay of these concurrent analyses can help to push the field into new frontiers. The first part of the book surveys the theoretical and methodological underpinnings of multilayer corpus work, including an exploration of various technical and data collection issues. The second part builds on the groundwork of the first half to show multilayer corpora applied to different subfields of linguistic study, including information structure research, referentiality, discourse models, and functional theories of discourse analysis, synthesizing these different discussions in a detailed case study of non-standard language in its concluding chapter. Advancing the multilayer corpus linguistic research paradigm into new and different
directions, this volume is an indispensable resource for graduate students and researchers in corpus linguistics, syntax, semantics, construction studies, and cognitive grammar.

**GECCo - German-English Contrasts in Cohesion** Kerstin Kunz 2021-10-25 In contrastive linguistics English-German, there is a tradition of accounting for contrasts with respect to grammar, and to a lesser extent, for lexis and phonetics. Moving on to discourse and text, there is a sizeable body of literature on cohesive patterns in English and German respectively - but very little in terms of a comparison. The latter, though, would be of particular interest for language learners, translators and, of course, linguists. The proposed book attempts to close this gap, based on a number of years of corpus-based study into variation and cohesion in the two languages. While there is an overall focus on the language contrast, we shall for each of the two languages also investigate variation between several different registers language-internally, and between written and spoken mode in particular. For each of the five major types of cohesion (co-reference, substitution, ellipsis, conjunctive relations and lexical cohesion), overviews will be given of contrasts in the system and of contrastive frequencies in instantiated text. The book should thus be relevant for language teaching, translation and generally corpus-based work on English and German, in terms of results and methodology.

**Twenty Years of Learner Corpus Research. Looking Back, Moving Ahead** Sylviane Granger 2013-06-21 This proceedings volume covers issues of learner corpus design, collection and annotation and contains reports on various aspects of (written and spoken) learner interlanguage as well as design of learner-corpus-informed tools.

**Multiword expressions** Manfred Sailer Multiword expressions (MWEs) are a challenge for both the natural language applications and the linguistic theory because they often defy the application of the machinery developed for free combinations where the default is that the meaning of an utterance can be predicted from its structure. There is a rich body of primarily descriptive work on MWEs for many European languages but comparative work is little. The volume brings together MWE experts to explore the benefits of a multilingual perspective on MWEs. The ten contributions in this volume look at MWEs in Bulgarian, English, French, German, Maori, Modern Greek, Romanian, Serbian, and Spanish. They discuss prominent issues in MWE research such as classification of MWEs, their formal grammatical modeling, and the description of individual MWE types from the point of view of different theoretical frameworks, such as Dependency Grammar, Generative Grammar, Head-driven Phrase Structure Grammar, Lexical Functional Grammar, Lexicon Grammar.

**Variation in Time and Space** Anna Čermáková 2020-12-07 Variation in Time and Space: Observing the World through Corpora is a collection of articles that address the theme of linguistic variation in English in its broadest sense. Current research in English language presented in the book explores a fascinating number of topics, whose unifying element is the corpus linguistic
methodology. Part I of this volume, Meaning in Time and Space, introduces the two dimensions of variation – time and space – relating them to the negotiation of meaning in discourse and questions of intertextuality. Part II, Variation in Time, approaches the English language from a diachronic point of view; the time periods covered vary considerably, ranging from 16th century up to present-day; so do the genres explored. Part III, Variation in Space, focuses on global varieties of English and includes a contrastive point of view. The range of topics is again broad – from specific lexico-grammatical structures to the variation in academic English, combining the regional and genre dimensions of variation. This is a timely volume that shows the breadth and depth in current corpus-based research of English.

Corpus Linguistics and Linguistically Annotated Corpora Sandra Kuebler 2015 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.

Annotations in Scholarly Editions and Research Julia Nantke 2020-10-26 The term ‘annotation’ is associated in the Humanities and Technical Sciences with different concepts that vary in coverage, application and direction but which also have instructive parallels. This publication mirrors the increasing cooperation that has been taking place between the two disciplines within the scope of the digitalization of the Humanities. It presents the results of an international conference on the concept of annotation that took place at the University of Wuppertal in February 2019. This publication reflects on different practices and associated concepts of annotation in an interdisciplinary perspective, puts them in relation to each other and attempts to systematize their commonalities and divergences. The following dynamic visualizations allow an interactive navigation within the volume based on keywords: Wordcloud ☁, Matrix ▭, Edge Bundling ⊛

Building and Using Comparable Corpora Serge Sharoff 2013-12-13 The 1990s saw a paradigm change in the use of corpus-driven methods in NLP. In the field of multilingual NLP (such as machine translation and terminology mining) this implied the use of parallel corpora. However, parallel resources are relatively scarce: many more texts are produced daily by native speakers of any given language than translated. This situation resulted in a natural drive towards
the use of comparable corpora, i.e. non-parallel texts in the same domain or genre. Nevertheless, this research direction has not produced a single authoritative source suitable for researchers and students coming to the field. The proposed volume provides a reference source, identifying the state of the art in the field as well as future trends. The book is intended for specialists and students in natural language processing, machine translation and computer-assisted translation.

**Developing Writers in Higher Education** Anne R Gere 2018-12-19 For undergraduates following any course of study, it is essential to develop the ability to write effectively. Yet the processes by which students become more capable and ready to meet the challenges of writing for employers, the wider public, and their own purposes remain largely invisible. Developing Writers in Higher Education shows how learning to write for various purposes in multiple disciplines leads college students to new levels of competence. This volume draws on an in-depth study of the writing and experiences of 169 University of Michigan undergraduates, using statistical analysis of 322 surveys, qualitative analysis of 131 interviews, use of corpus linguistics on 94 electronic portfolios and 2,406 pieces of student writing, and case studies of individual students to trace the multiple paths taken by student writers. Topics include student writers’ interaction with feedback; perceptions of genre; the role of disciplinary writing; generality and certainty in student writing; students’ concepts of voice and style; students’ understanding of multimodal and digital writing; high school’s influence on college writers; and writing development after college. The digital edition offers samples of student writing, electronic portfolios produced by student writers, transcripts of interviews with students, and explanations of some of the analysis conducted by the contributors. This is an important book for researchers and graduate students in multiple fields. Those in writing studies get an overview of other longitudinal studies as well as key questions currently circulating. For linguists, it demonstrates how corpus linguistics can inform writing studies. Scholars in higher education will gain a new perspective on college student development. The book also adds to current understandings of sociocultural theories of literacy and offers prospective teachers insights into how students learn to write. Finally, for high school teachers, this volume will answer questions about college writing.

**The Cambridge Handbook of World Englishes** Daniel Schreier 2020-01-02 The plural form 'Englishes' conveys the diversity of English as a global language, pinpointing the growth and existence of a large number of national, regional and social forms. The global spread of English and the new varieties that have emerged around the world has grown to be a vast area of study and research, which intersects multiple disciplines. This Handbook provides a comprehensive and authoritative survey of World Englishes from 1600 to the present day. Covering topics such as variationist sociolinguistics, pragmatics, contact linguistics, linguistic anthropology, corpus- and applied linguistics and language history, it combines discussion of traditional topics with a variety of innovative approaches. The chapters, all written by internationally...
acclaimed authorities, provide up-to-date discussions of the evolution of different Englishes around the globe, a comprehensive coverage of different models and approaches, and some original perspectives on current challenges.

*Introducing Linguistic Research* Svenja Voelkel 2021-08-31 Over the past decade, conducting empirical research in linguistics has become increasingly popular. The first of its kind, this book provides an engaging and practical introduction to this exciting versatile field, providing a comprehensive overview of research aspects in general, and covering a broad range of subdiscipline-specific methodological approaches. Subfields covered include language documentation and descriptive linguistics, language typology, corpus linguistics, sociolinguistics and anthropological linguistics, cognitive linguistics and psycholinguistics, and neurolinguistics. The book reflects on the strengths and weaknesses of each single approach and on how they interact with one-another across the study of language in its many diverse facets. It also includes exercises, example student projects and recommendations for further reading, along with additional online teaching materials. Providing hands-on experience, and written in an engaging and accessible style, this unique and comprehensive guide will give students the inspiration they need to develop their own research projects in empirical linguistics.

*The Cambridge Handbook of Learner Corpus Research* Sylviane Granger 2015-10-01 The origins of learner corpus research go back to the late 1980s when large electronic collections of written or spoken data started to be collected from foreign/second language learners, with a view to advancing our understanding of the mechanisms of second language acquisition and developing tailor-made pedagogical tools. Engaging with the interdisciplinary nature of this fast-growing field, The Cambridge Handbook of Learner Corpus Research explores the diverse and extensive applications of learner corpora, with 27 chapters written by internationally renowned experts. This comprehensive work is a vital resource for students, teachers and researchers, offering fresh perspectives and a unique overview of the field. With representative studies in each chapter which provide an essential guide on how to conduct learner corpus research in a wide range of areas, this work is a cutting-edge account of learner corpus collection, annotation, methodology, theory, analysis and applications.

*Language, Data, and Knowledge* Jorge Gracia 2017-06-08 This book constitutes the proceedings of the First International Conference on Language, Data and Knowledge, LDK 2017, held in Galway, Ireland, in June 2017. The 14 full papers and 19 short papers included in this volume were carefully reviewed and selected from 68 initial submissions. They deal with language data; knowledge graphs; applications in NLP; and use cases in digital humanities, social sciences, and BioNLP.