Briefly Noted

Natural Language Processing Using Very Large Corpora

Susan Armstrong, Kenneth Church, Pierre Isabelle, Sandra Manzi, Evelyne Tzoukermann, and David Yarowsky (editors)
(University of Geneva, AT&T Labs–Research, Xerox Research Centre Europe, University of Geneva, Bell Laboratories, and Johns Hopkins University)

Dordrecht: Kluwer Academic Publishers (Text, speech and language technology series, edited by Nancy Ide and Jean Véronis, volume 11), 1999, xvii+304 pp; hardbound, ISBN 0-7923-6055-9, $128.00, £79, Dfl 240.00

This volume is a collection of important papers selected from the proceedings of the first three Workshops on Very Large Corpora (Columbus, 1993; Kyoto, 1994; Cambridge, MA, 1995) and the 1995 workshop entitled From Text to Tags (Dublin). The editors write in their introduction: "The success of these workshops was in some measure a reflection of the growing popularity of corpus-based methods in the NLP community. But first and foremost, it was due to the fact that the workshops attracted so many high-quality papers. The importance of this material for the field is such that it deserves to be made more readily available than harder-to-find or out-of-print workshop proceedings." The contents of the volume are as follows:

"Implementation and evaluation of a German HMM for POS disambiguation" by Helmut Feldweg

"Improvements in part-of-speech tagging with an application to German" by Helmut Schmid

"Unsupervised learning of disambiguation rules for part-of-speech tagging" by Eric Brill and Mihai Pop

"Tagging French without lexical probabilities —Combining linguistic knowledge and statistical learning" by Evelyne Tzoukermann, Dragomir Radev, and William Gale

"Example-based sense tagging of running Chinese text" by Xiang Tong, Chang-ning Huang, and Cheng-ming Guo

"Disambiguating noun groupings with respect to WordNet senses" by Philip Resnik

"A comparison of corpus-based techniques for restoring accents in Spanish and French

text" by David Yarowsky

"Beyond word *N*-grams" by Fernando Pereira, Yoram Singer, and Naftali Tishby

"Statistical augmentation of a Chinese machine-readable dictionary" by Pascale Fung and Dekai Wu

"Text chunking using transformation-based learning" by Lance Ramshaw and Mitchell P. Marcus

"Prepositional phrase attachment through a backed-off model" by Michael Collins and James Brooks

"On the unsupervised induction of phrase-structure grammars" by Carl de Marcken

"Robust bilingual word alignment for machine aided translation" by Ido Dagan, Kenneth Church, and William Gale

"Iterative alignment of syntactic structures for a bilingual corpus" by Ralph Grishman

"Trainable coarse bilingual grammars for parallel text bracketing" by Dekai Wu

"Comparative discourse analysis of parallel texts" by Pim van der Eijk

"Comparing the retrieval performance of English and Japanese text databases" by Hideo Fujii and W. B. Croft

"Inverse document frequency (IDF): A measure of deviations from Poisson" by Kenneth Church and William Gale

Techniques in Speech Acoustics

Jonathan Harrington and Steve Cassidy
(Macquarie University)

Dordrecht: Kluwer Academic Publishers (Text, speech and language technology series, edited by Nancy Ide and Jean Véronis, volume 8), 1999, xiv+319 pp and CD-ROM; hardbound, ISBN 0-7923-5731-0, $150.00, £88, Dfl 250.00

"This book is the development of a series of lectures to undergraduate and post-graduate students at Macquarie University on basic principles in acoustic phonetics and speech signal processing. The first part of the book is intended to provide students with the ability to interpret acoustic records of speech signals in their various forms. These chapters include a review of elementary wave motion and frequency analysis as applied to speech, a summary of the relationship between speech production and its acoustic consequences, and a review of the principal cues to speech sounds and prosodic units.

294