1. Record Nr. UNISA996208491403316

Autore	Peter Spyns Jan Odijk
Titolo	Essential Speech and Language Technology for Dutch : Results by the STEVIN-programme / / edited by Peter Spyns, Jan Odijk
Pubbl/distr/stampa	Cham, : Springer Nature, 2013 Berlin, Heidelberg : , : Springer Berlin Heidelberg : , : Imprint : Springer, , 2013
ISBN	3-642-30910-0
Edizione	[First edition 2013.]
Descrizione fisica	1 online resource (xvii, 413 pages) : 79 illustrations, 29 illustrations in colour
Collana	Theory and Applications of Natural Language Processing
Disciplina	410.285
Soggetti	Computational linguistics Germanic languages Artificial intelligence Computational Linguistics Germanic Languages Artificial Intelligence
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	Foreword. by Linde van den Bosch -- Introduction -- Peter Spyns -- Part I. How it started -- 2.The STEVIN Programme: Result of Five Years Cross-Border HLT for Dutch Policy Preparation. P.Spyns and E.D' Halleweyn -- Part II. HLT Resource-project Related Papers -- 3.The JASMIN Speech Corpus: Recordings of Children, Non-Natives and Elderly People. C. Cucchiarini and H. Van Hamme -- 4.Resources Developed in the Autonomata Projects. H.van den Heuvel, J-P.Martens, G.Bloothooft, M.Schraagen, N.Konings, K.D'hanens, and Q.Yang -- 5. STEVIN can Praat. D.Weenink -- 6.SPRAAK: Speech Processing, Recognition and Automatic Annotation Kit. P.Wambacq, K.Demuynck, and D.Van Compernolle -- 7.COREA: Coreference Resolution for Extracting Answers for Dutch. I.Hendrickx, G.Bouma, W.Daelemans and V.Hoste -- 8.Automatic Tree Matching for Analysing Semantic Similarity in Comparable Text. E.Marsi and E.Krahmer -- 9.Large Scale Syntactic Annotation of Written Dutch: Lassy. G.van Noord, G.Bouma, F.van

Eynde, D.de Kok, J.van der Linde, I.Schuurman, E.Tjong Kim Sang, and V.Vandeghinste -- 10.Cornetto: a Combinatorial Lexical Semantic Database for Dutch. P.Vossen, I.Maks, R.Segers, H.van der Vliet, M-F. Moens, K.Hofmann, E.Tjong Kim Sang, and M.de Rijke -- 11.Dutch Parallel Corpus: a Balanced Parallel Corpus for Dutch-English and Dutch-French. H.Paulusen, L.Macken, W.Vandeweghe, and P.Desmet -- 12.Identification and Lexical Representation of Multiword Expressions. J.Odijk -- 13.The Construction of a 500-million-word Reference Corpus of Contemporary Written Dutch. N.Oostdijk, M.Reynaert, V. Hoste, and I.Schuurman -- Part III. HLT Technology Related Papers -- 14.Lexical Modeling for Proper Name Recognition in Autonomata Too -- B.Réveil, J-P.Martens, H.van den Heuvel, G.Bloothooft, and M. Schraagen -- 15.N-Best 2008: a Benchmark Evaluation for Large Vocabulary Speech Recognition in Dutch. D.A. van Leeuwen -- 16. Missing Data Solutions for Robust Speech Recognition. Y.Wang, J.F. Gemmeke, K.Demuynck, and H.Van Hamme -- 17.Parse and Corpus-based Machine Translation. V.Vandeghinste, S.Martens, G.Kotzé, J. Tiedemann, J.Van den Bogaert, K.De Smet, F.Van Eynde, and G.van Noord -- Part IV.HLT Application Related Papers -- 18.Development and Integration of Speech technology into COurseware for Language Learning: the DISCO Project. H. Strik, J. van Doremalen, J. Colpaert, and C. Cucchiarini -- 19.Question Answering of Informative Web Pages: How Summarisation Technology Helps. J.De Belder, D.de Kok, G.van Noord, F.Nauze, L.van der Beek, and M-F.Moens -- 20.Generating, Refining and Using Sentiment Lexicons. M.de Rijke, V.Jijkoun, F.Laan, W.Weerkamp, P.Ackermans, and G.Geleijnse -- Part V. And now -- 21. The Dutch-Flemish HLT Agency: Managing the Lifecycle of STEVIN's Language Resources. R.van Veenendaal, L.van Eerten, C.Cucchiarini, and P.Spyns -- 22.Conclusions and Outlook to the Future. Jan Odijk.

Sommario/riassunto

The book provides an overview of more than a decade of joint R&D efforts in the Low Countries on HLT for Dutch. It not only presents the state of the art of HLT for Dutch in the areas covered, but, even more importantly, a description of the resources (data and tools) for Dutch that have been created are now available for both academia and industry worldwide. The contributions cover many areas of human language technology (for Dutch): corpus collection (including IPR issues) and building (in particular one corpus aiming at a collection of 500M word tokens), lexicology, anaphora resolution, a semantic network, parsing technology, speech recognition, machine translation, text (summaries) generation, web mining, information extraction, and text to speech to name the most important ones. The book also shows how a medium-sized language community (spanning two territories) can create a digital language infrastructure (resources, tools, etc.) as a basis for subsequent R&D. At the same time, it bundles contributions of almost all the HLT research groups in Flanders and the Netherlands, hence offers a view of their recent research activities. Targeted readers are mainly researchers in human language technology, in particular those focusing on Dutch. It concerns researchers active in larger networks such as the CLARIN, META-NET, FLaReNet and participating in conferences such as ACL, EACL, NAACL, COLING, RANLP, CICling, LREC, CLIN and DIR (both in the Low Countries), InterSpeech, ASRU, ICASSP, ISCA, EUSIPCO, CLEF, TREC, etc. In addition, some chapters are interesting for human language technology policy makers and even for science policy makers in general.