
Språkbanken

Årsrapport 2013

ÖVERSIKT

I DENNA ÅRSRAPPORT redovisas merparten av de språkteknologiverksamheter som bedrivs vid institutionen för svenska språket. I vårt identitetsarbete för vi samman dessa verksamheter under rubriken ”Språkbanken” (1) för att markera att de aktiviteter som beskrivs här hör så nära ihop att de i praktiken utgör en forskningsmiljö, samt (2) därför att Språkbanken är en väl inarbetad benämning (som går tillbaka till 1970-talet) på dessa aktiviteter.

När vi talar om Språkbankens verksamhet menar vi således språkteknologisk forskningsverksamhet vid institutionen för svenska språket som finansieras av fakulteten under benämningarna *Språkbanken* och *språkvetenskaplig databehandling* (forsknings- och forskarutbildningsämne som deltar i grundutbildningen inom masterprogrammet i språkteknologi), samt av Göteborgs universitet genom en riktad satsning på styrkeområdet språkteknologi under rubriken *Centre for Language Technology* (CLT), samt ett antal externfinansierade forskningsprojekt inom språkteknologi.

KONFERENSER/PRESENTATIONER

Under denna rubrik har vi samlat en viktig del av vår utåtriktade verksamhet. Denna del består i att vi presenterar vårt arbete i form av föredrag eller posterpresentationer vid konferenser och workshoppar. Dessa har en öppen inbjudan att inkomma med bidrag (inklusive sådana som vi själva arrangerar eller är med om att arrangera). Inom området språkteknologi är sådana konferenser huvudpubliceringskanalen. Bidrag skickas först in i fulltext (typiskt 4–8 sidor) och bedöms i normalfallet dubbelt anonymt (anonymiserade bidrag bedöms av anonyma granskare) av två till tre fackgranskare. Sådana presentationer återfinns i nästa avsnitt.

En annan viktig del av den utåtriktade verksamheten består i presentationer efter inbjudan eller egen anmälan vid seminarier, projektmöten o.dyl. Huvudskillnaden mot föregående kategori är avsaknaden av explicit fackgranskning. Naturligtvis avspeglar mängden och den geografiska spridningen av den här sortens presentationer ändå forskargruppens aktivitet och rykte. Dessa redovisas i avsnittet *Andra presentationer* nedan.

2 *Språkbanken*

Konferens/workshop, plats/månad (antal presentationer)

- The 4th International Louhi Workshop on Health Document Text Mining and Information Analysis (Louhi 2013), Sydney, Australien/februari (1)
- The 35th European Conference on IR Research (ECIR 2013), Moskva, Ryssland/mars (1)
- The 10th International Conference on Computational Semantics, Potsdam, Tyskland/mars (1)
- The 14th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing), Samos, Grekland/mars (1)
- The Language Testing and CEFR conference, Antwerpen, Belgien/maj (1)
- The 19th Nordic Conference of Computational Linguistics (NODALIDA 2013), Oslo, Norge/maj (3)
- Workshop on Computational Historical Linguistics (NODALIDA 2013), Oslo, Norge/maj (2)
- Workshop on Lexical Semantic Resources for NLP (NODALIDA 2013), Oslo, Norge/maj (1)
- The 2nd Workshop on NLP for Computer-Assisted Language Learning (NODALIDA 2013), Oslo, Norge/maj (1)
- The 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT), Atlanta, USA/juni (1)
- The 7th international Corpus Linguistics conference (CL2013), Lancaster, Storbritannien/juli (2)
- Digital Humanities, Lincoln, Nebraska, USA/juli (1)
- The 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH), Sofia, Bulgarien/augusti (1)
- The 14th World Congress on Medical and Health Informatics (MEDINFO 2013), Köpenhamn, Danmark/augusti (2)
- The 2013 EUROCALL Conference, Évora, Portugal/september (2)
- The 5th Conference on Quantitative Investigations in Theoretical Linguistics (QITL-5), Leuven, Belgien/september (1)
- The International Conference Recent Advances in Natural Language Processing (RANLP), Hissar, Bulgarien/september (1)
- The 22nd ACM International Conference on Information and Knowledge Management (CIKM 2013), San Francisco, USA/oktober (2)
- The 6th Workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR 2013), held at CIKM, San Francisco, USA/oktober (1)
- Electronic lexicography in the 21st century: thinking outside the paper. eLex 2013 conference, Tallinn, Estland/oktober (3)

- The 3rd Workshop on Annotation of Corpora for Research in the Humanities (ACRH-3), Sofia, Bulgarien/december (1)

Andra presentationer, plats/månad

- Centre for Language Technology (CLT) seminarium, Göteborg/under året (8 seminarier)
- Exploring the past: Language technology for Old Swedish corpora, Uppsala/februari
- Exploring the past: Language technology for Old Swedish corpora, Gent, Belgien/mars
- Lärka vid Centrum för Språkintröduktion, Göteborg/april
- Presentation av Korp, Svenska litteratursällskapet i Finland, Helsingfors, Finland/april
- Presentation vid Molto final project meeting, Barcelona, Spanien/maj
- From CEFR-relevant course book corpus to CEFR-aware readability measures vid Läromedelsworkshop, Göteborg/juni
- Presentation vid Third GF Summer School, Chiemsee, Tyskland/augusti
- Institutionen för svenska språkets personalkonferens, Uddevalla/augusti (12 bidrag)
- Mot svensk språkteknologi för forskning och utveckling – en rapport från fronten, Nordiskt språkmöte, Akureyri, Island/augusti
- Presentation of Språkbanken vid Social Sensing Workshop, Göteborg/september
- GRUS för lärare, Göteborg/september
- Presentation vid FrameNet-gruppens veckomöte på ICSI, Berkeley, USA/oktober
- Centre for Language Technology (CLT) årlig workshop, Uddevalla/oktober (12 bidrag)
- Text mining vid EUROPEANA meeting, Göteborg/oktober
- Presentation vid Nationell termkonferens 2013, Göteborg/oktober (1 bidrag)
- The resource infrastructure at Språkbanken vid CLARIN annual meeting, Prag, Tjeckien/oktober
- Bra (språk)(forsknings)datahantering – vad är det?, SND:s workshop Forskningsmaterial – en framtidsfråga, Göteborg/november
- En fri molntjänst för OCR vid Digitalisera Vidare, Uppsala/december

I media

Utöver regelrätta publikationer och presentationer har Språkbanken och språkbankens resurser omnämmts i media i flera fall.

- Språket i P1
- Språktidningen

4 *Språkbanken*

Organiserad konferens/workshop/seminarium, plats/månad (totalt antal deltagare)

- Flerordsworkshop, Göteborg/mars (ca 25)
- International FrameNet Workshop, Berkeley, USA/april (ca 30)
- 2nd workshop on NLP for computer-assisted language learning, i samband med NODALIDA, Oslo, Norge/maj (ca 20)
- Workshop on computational historical linguistics, i samband med NODALIDA, Oslo, Norge/maj (ca 25)
- Workshop on lexical semantic resources for NLP, i samband med NODALIDA, Oslo, Norge/maj (ca 20)
- Workshop on Nordic language research infrastructure, i samband med NODALIDA, Oslo, Norge/maj (ca 20)
- Språkbankens årliga höstworkshop, Göteborg/oktober (ca 40)
- Workshop om digitala språkhistoriska resurser, Göteborg/november (ca 25)
- Idéseminarium: Skönlitteratur som vetenskapligt källmaterial - Hur analyserar vi kulturarvet utan att låsas av kanon?, Göteborg/november (ca 25)

PROJEKT

Språkbankens arbete utförs inom en mängd olika projekt. Några av dessa projekt är interna, medan flertalet innebär samarbeten på nationell eller internationell nivå. Här redovisar vi de forskningsprojekt som vi deltog i under året.

- AO – en akademisk ordlista för svenska (år 3); finansiär ISA
- Automatic identification of language change; egen finansiering
- A System Architecture for Intelligent Computer-Assisted Language Learning (år 3 av 3); finansiär NordPlus Sprog
- CLT:s textteknologilaboratorium (år 5 av 7); finansiär CLT (Göteborgs universitet)
- Diabase – Towards a diachronic BLARK; egen finansiering
- Digital areallingvistik (år 4); finansiär VR
- En fri molntjänst för OCR (år 1 av 1); finansiär Kungliga biblioteket, Göteborgs universitet
- Förbättring av testkollektionen MedEval; finansiär CLT (Göteborgs universitet), egen finansiering
- GRUS – Gymnasie- och grundskoleelevers skrivande i ”en dator till en elev”-satsning (år 3); finansiär ISA
- Kulturomik: Mot kunskapsbaserad storskalig kunskapsutvinning ur svensk text (år 2 av 5); finansiär VR

- Litteraturbanken (permanent); finansiär Svenska Akademien, Kungl. Vetenskapsakademien
- META-NORD (år 2 av 2); finansiär EU
- MOLTO (år 4); finansiär EU
- PROMISE (år 3 av 3); finansiär EU
- SweCxn – ett svenskt konstruktikon (år 1 av 3); finansiär RJ
- SweFN++ (år 3); finansiär CLT (Göteborgs universitet), VR

Flera projektansökningar har inlämnats för pågående och kommande projekt.

- Construction Linguistics in Natural Language Processing. Ansökt till ERC Starting Grant (ansökt 2012, avslagen)
- Distributionella metoder för ramsemantik och konstruktionsgrammatik. Ansökt till Vetenskapsrådet (beviljad)
- From reading to writing: evidence-based studies of CEFR texts. Ansökt till Vetenskapsrådet (avslagen)
- Koala – Korps lingvistiska annotationer. Ansökt till Riksbankens Jubileumsfond (beviljad)
- LTLOD@SB. Ansökt till VINNOVA (beviljad)
- MA π iR: Methods for the automatic analysis of text in digital historical resources. Ansökt till Markus och Amalia Wallenberg (ansökt 2012, beviljad)
- MA π iR: Methods for the automatic analysis of text in digital historical resources. Ansökt till Riksbankens Jubileumsfond (avslagen)
- MA π iR: Methods for the automatic analysis of text in digital historical resources. Ansökt till Vetenskapsrådet (avslagen)
- SWE-CLARIN: Svensk språkteknologi för humaniora och samhällsvetenskap. Ansökt till Vetenskapsrådet (beviljad)

INFRASTRUKTUR

Språkbanken arbetar aktivt för att utveckla en språkteknologisk infrastruktur. Detta arbete omfattar i dagsläget följande komponenter:

- lexikal infrastruktur,
- korpusinfrastruktur, samt
- metadata.

Uppbyggnaden av infrastruktur för lexikala resurser och korpora handlar för närvarande om att harmonisera och standardisera så många fria språkliga resurser som möjligt, samt att göra dem tillgängliga för forskningsvärlden. Vi skapar dessutom verktyg för att utforska dessa resurser, t ex Karp (tidigare SBLEX) för lexikala resurser och Korp för korpora. Målet är vidare att alla dessa resurser ska vara väl beskrivna i ett metadata-repositorium som följer de

6 *Språkbanken*

standarder som finns inom området.

Öppenhet är ett av våra ledord, en filosofisk ståndpunkt vi försöker att tillämpa i största möjliga mån. Vi anser att forskning ska utföras öppet, för att tillåta granskning och samarbete. I denna öppenhet ingår att använda öppna standarder och licenser samt att använda och skapa verktyg med öppen källkod. Även om alla äldre resurser inte kan göras fritt tillgängliga, p.g.a. restriktiva licenser, strävar vi efter att samla fler och större fria resurser, för att främja språklig forskning och utveckling av språkteknologiska tillämpningar, i Sverige och världen.

Resurserna används flitigt, och t ex Korp har fler än 1000 användare varje månad. Majoriteten av dessa finns av naturliga skäl i Sverige, men många sökningar görs även från andra delar av Europa och resten av världen.

SAMVERKAN

En central del i samverkan med det omgivande samhället består i att vi är representerade i relevanta externa organisationer. Här redovisas denna medverkan.

- CLT:s styrgrupp – Centre for Language Technology
- Dramawebbens referensgrupp
- Litteraturbankens styrelse
- MLT:s styrgrupp – Masterprogrammet i språkteknologi
- SIS TK115 – Swedish Standards Institute, Terminologi och språkliga resurser
- SND:s styrelse – Svensk Nationell Datatjänst
- Språkrådets rådgivargrupp för språkteknologi

Utöver medverkan i dessa organisationer samarbetar vi, genom Språkbanken som helhet eller genom projekt, med följande organisationer, institutioner, och företag.

Nationella samarbeten

- Digisam
- Dramawebben
- Gothia Forum för klinisk forskning
- GPCC – Centrum för personcentrerad vård, Göteborgs universitet
- Institutet för svenska som andraspråk, Göteborgs universitet
- Kungliga biblioteket
- Kungliga tekniska högskolan
- Lexikaliska institutet, Göteborgs universitet
- Linköpings universitet
- Lunds universitet

- Riksarkivet mediakonverteringscentrum
- Sahlgrenska universitetssjukhuset
- SND – Svensk Nationell Datatjänst
- Socialstyrelsen
- Språkrådet, Institutet för folk- och språkminnen
- Stockholms universitet
- Uppsala universitet

Internationella samarbeten

- Berkeley FrameNet
- Centrum för internationalisering och parallellspråklighet, Köpenhamns universitet
- CLARIN – Common language resources and technology infrastructure
- Kelly – Keywords for language learning for young and adults alike
- Max Planck-institutet för evolutionär antropologi
- META-NORD – The multilingual Europe technology alliance, specifically the Baltic and Nordic parts
- MOLTO – Multilingual online translation
- Tekstlaboratoriet, UiO – Universitetet i Oslo

BESÖK

Ett flertal gästforskare och seminariegäster har besökt oss under året.

- Graeme Hirst, Torontos universitet
- Måns Huldén, Helsingfors universitet
- Valia Kordoni, Humboldt-universitetet Berlin/Saarlands universitet
- Marco Kuhlmann, Uppsala universitet
- Maria Liakata, Warwicks universitet
- John McCrae, Bielefelds universitet
- Eleni Miltsakaki, Pennsylvanias universitet
- Miriam R L Petruck, International Computer Science Institute (ICSI)
- Maciej Piasecki, Wroclaws tekniska universitet
- Barbara Plank, Köpenhamns universitet
- Caroline Sporleder, Triers universitet
- Nina Tahmasebi, Leibniz universitet Hannover
- He Tan, Linköpings universitet
- Martin Volk, Zürichs universitet

BEDÖMNINGSUPPDRAG

Utöver deltagande med egna presentationer vid konferenser och workshopar deltar vi även som fackgranskare av andra bidrag. Ofta tilldelas man flera bidrag att bedöma. Här redovisas sådana bedömningsuppdrag.

- Tidskriftsartiklar
 - Extended CLEF and Louhi papers submitted to the special issue of the Artificial Intelligence in Medicine
 - CALICO journal
 - Constructions and Frames, special issue based on the IFNW'13 presentations
 - Information Systems
 - Journal of Logic and Computation
 - Linguistic Issues in Language Technology
 - Linguistics
 - Northern European Journal of Language Technology
 - Veredas Special Issue on Frame Semantics
- Konferensbidrag
 - ACL
 - CLEFeHealth 2013
 - Digital Humanities
 - EACL
 - ECIR
 - EMNLP
 - International FrameNet Workshop 2013
 - The 4th International Workshop on Health Document Text Mining and Information Analysis with the Focus of Cross-Language Evaluation
 - HLT4LL symposium
 - Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCh)
 - LREC
 - *SEM
 - NAACL-HLT workshop: Innovative Use of NLP for Building Educational Applications
 - NODALIDA
 - NODALIDA 2013 workshop: Computational historical linguistics
 - NODALIDA 2013 workshop: Lexical semantic resources for NLP
 - NODALIDA 2013 workshop: NLP for computer-assisted language learning

- NODALIDA 2013 workshop: Nordic language research infrastructure

10 *Språkbanken*

- Övrigt
 - Projektansökan för det tjeckiska vetenskapsrådet (GAČR)

NY PERSONAL

Under året har Språkbanken anställt två nya doktorander i Språkvetenskaplig databehandling.

- Malin Ahlberg
- Ildikó Pilán

UTBILDNING

Personal från Språkbanken har deltagit i undervisning om olika aspekter av språkteknologi, ibland gällande enstaka föreläsningar, ibland med kursansvar, vilket innefattar planering och examination.

Grundläggande och avancerad nivå

- Skandinavisk grammatik, vid Gents universitet (kandidatnivå) vt 2013
- Handledning av masteruppsatser (masternivå) vt 2013
- Språkteknologiska moment på kurser i svenska och svenska som andraspråk (grundnivå) ht2013
- Natural language processing (masternivå) ht 2013
- Introduction to formal linguistics (masternivå) ht 2013
- Introduktion till programmering (masternivå) ht 2013 (kursansvar)
- Planering av Introduktion till språkteknologi för språkvetare

Forskarutbildning

- Kurser:
 - Korpuslingvistik (forskarnivå) ht2013 (kursansvar)
- Doktorander:
 - Malin Ahlberg, Computational historical linguistics, började 2013, handledare Lars Borin, Markus Forsberg och Henrik Rosenkvist
 - Dana Dannélls, Multilingual text generation from structured formal representations, började 2007, handledare Lars Borin och Arne Ranta
 - Ann-Marie Eklund, Analysis of queries to medical web sites and databases as a basis for knowledge sharing, började 2010, handledare Dimitrios Kokkinakis och Jussi Karlgren

- Katarina Heimann Mühlenbock, Enhancing readability by text simplification, började 2008, handledare Sofie Johansson Kokkinakis och Lars Borin
 - Ildikó Pilán, Swedish second language readability from a semantic perspective, började 2013, handledare Lars Borin och Elena Volodina
 - Taraka Rama Kasicheyanula, Computational historical linguistics, började 2010, handledare Lars Borin och Søren Wichmann
 - Judy Ribbeck, Syntactic analysis of textbooks used in Sweden when teaching mathematics and natural science, började 2010, handledare Emma Sköldberg och Lars Borin
- Disputationer:
 - Dana Dannélls, Multilingual text generation from structured formal representations, disputerade februari 2013
 - Katarina Heimann Mühlenbock, I see what you mean. Assessing readability for specific target groups, disputerade april 2013

PUBLIKATIONER

1. Malin Ahlberg, Lars Borin, Markus Forsberg, Martin Hammarstedt, Leif-Jöran Olsson, Olof Olsson, Johan Roxendal och Jonatan Uppström. Korp and Karp – a bestiary of language resources: the research infrastructure of Språkbanken. I: *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2013), May 22-24, 2013, Oslo, Norway*, nummer 16 i NEALT Proceedings, ss 429–433, Linköping, Sweden, 2013. Linköping University Electronic Press.
2. Peter Andersson och Malin Ahlberg. Towards automatic tracking of lexical change: linking historical lexical resources. I: *Proceedings of the Workshop on Computational Historical Linguistics (NODALIDA 2013), May 22-24, 2013, Oslo, Norway*, nummer 18 i NEALT Proceedings, ss 1–10, Linköping, Sweden, 2013. Linköping University Electronic Press.
3. Lars Borin. The why and how of measuring linguistic differences. I: Lars Borin och Anju Saxena, redaktörer, *Approaches to Measuring Linguistic Differences*, ss 3–26. De Gruyter Mouton, Berlin, Germany, 2013.
4. Lars Borin, Bernard Comrie och Anju Saxena. The Intercontinental Dictionary Series – a rich and principled database for language comparison. I: Lars Borin och Anju Saxena, redaktörer, *Approaches to Measuring Linguistic Differences*, ss 285–302. De Gruyter Mouton, Berlin, Germany, 2013.

5. Lars Borin, Devdatt Dubhashi, Markus Forsberg, Richard Johansson, Dimitrios Kokkinakis och Pierre Nugues. Mining semantics for culturo-mics: towards a knowledge-based approach. I: *Proceedings of the 2013 international workshop on Mining unstructured big data using natural language processing, held October 27th-November 1st at CIKM 2013 in San Francisco, CA, USA*, ss 3–10, New York, NY, USA, 2013. Association for Computing Machinery (ACM).
6. Lars Borin, Ruth Vatvedt Fjeld, Markus Forsberg, Sanni Nimb, Pierre Nugues och Bolette Sandford Pedersen, redaktörer. *Proceedings of the Workshop on Lexical Semantic Resources for NLP at NODALIDA 2013, May 22-24, 2013, Oslo, Norway*, nummer 19 i NEALT Proceedings, Linköping, Sweden, 2013. Linköping University Electronic Press.
7. Lars Borin, Markus Forsberg och Benjamin Lyngfelt. Close encounters of the fifth kind: Some linguistic and computational aspects of the Swedish FrameNet++ project. *Veredas*, 17(1):28–43, 2013.
8. Lars Borin, Markus Forsberg och Lennart Lönngren. Saldo: a touch of yin to WordNet’s yang. *Language resources and evaluation*, 47(4):1191–1211, 2013.
9. Lars Borin, Markus Forsberg, Leif-Jöran Olsson, Olof Olsson och Jonatan Uppström. The lexical editing system of Karp. I: I. Kosem, J. Kallas, P. Gantar, S. Krek, M. Langemets och M. Tuulik, redaktörer, *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference, 17-19 October 2013, Tallinn, Estonia*, ss 503–516, 2013.
10. Lars Borin och Anju Saxena, redaktörer. *Approaches to Measuring Linguistic Differences*. De Gruyter Mouton, Berlin, Germany, 2013.
11. Lars Borin och Elena Volodina, redaktörer. *Proceedings of the SLTC 2012 workshop on NLP for CALL*, Linköping, Sweden, 2012. Linköping University Electronic Press.
12. Gerlof Bouma och Yvonne Adesam. Experiments on sentence segmentation in Old Swedish editions. I: *Proceedings of the Workshop on Computational Historical Linguistics (NODALIDA 2013), May 22-24, 2013, Oslo, Norway*, nummer 18 i NEALT Proceedings, ss 11–26, Linköping, Sweden, 2013. Linköping University Electronic Press.
13. Linnéa Bäckström, Lars Borin, Markus Forsberg, Benjamin Lyngfelt, Julia Prentice och Emma Sköldberg. Automatic identification of construction candidates for a Swedish constructicon. I: *Proceedings of the Workshop on Lexical Semantic Resources for NLP (NODALIDA 2013), May 22-24, 2013, Oslo, Norway*, nummer 19 i NEALT Proceedings, ss 2–11,

- Linköping, Sweden, 2013. Linköping University Electronic Press.
14. Mariana Damova, Dana Dannélls, Maria Mateva, Ramona Enache och Aarne Ranta. Natural language interaction with Semantic Web Knowledge Bases and LOD. I: *Towards multilingual Semantic Web*. Springer, Berlin, Germany, 2013.
 15. Dana Dannélls, Lars Borin och Leif-Jöran Olsson. MapServer for Swedish language technology. I: *Digital Humanities*, 2013.
 16. Dana Dannélls, Aarne Ranta, Ramona Enache, Mariana Damova och Maria Mateva. Multilingual access to cultural heritage content on the semantic web. I: *Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH)*, 2013.
 17. Dana Dannélls, Aarne Ranta, Ramona Enache, Mariana Damova och Maria Mateva. Translation and retrieval system for museum object descriptions. MOLTO project deliverable, 2013.
 18. Dana Dannélls, Aarne Ranta, Ramona Enache, Inari Listenmaa, Laura Tolosi och Maria Mateva. Grammar-ontology interoperability – Final work and overview. MOLTO project deliverable, 2013.
 19. Koenrad De Smedt, Lars Borin, Krister Lindén, Bente Maegaard, Eiríkur Rögnvaldsson och Kadri Vider, redaktörer. *Proceedings of the Workshop on Nordic Language Research Infrastructure at NODALIDA 2013, May 22-24, 2013, Oslo, Norway*, nummer 20 i NEALT Proceedings, Linköping, Sweden, 2013. Linköping University Electronic Press.
 20. Ann-Marie Eklund. Mobility and health information searches - a Swedish perspective. I: *MEDINFO 2013 - Proceedings of the 14th World Congress on Medical and Health Informatics, 20-13 August 2013, Copenhagen, Denmark*, band 192 av *Studies in Health Technology and Informatics*, s 1079. IOS Press, 2013.
 21. Ann-Marie Eklund. On challenges with mobile e-health – lessons from a game-theoretic perspective. I: *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management, CIKM 2013, October 27th-November 1st, San Francisco, CA, USA*, New York, NY, USA, 2013. Association for Computing Machinery (ACM).
 22. Þórhallur Eypórsson, Lars Borin, Dag Haug och Eiríkur Rögnvaldsson, redaktörer. *Proceedings of the Workshop on Computational Historical Linguistics at NODALIDA 2013, May 22-24, 2013, Oslo, Norway*, nummer 18 i NEALT Proceedings, Linköping, Sweden, 2013. Linköping University Electronic Press.
 23. Johan Falkenjack, Katarina Heimann Mühlenbock och Arne Jönsson. Fe-

- atures indicating readability in Swedish text. I: *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2013), May 22-24, 2013, Oslo, Norway*, nummer 085, ss 27–40, Linköping, Sweden, 2013. Linköping University Electronic Press.
24. Nicola Ferro, Richard Berentsen, Allan Hanbury, Mihai Lupu, Vivien Petras, Marten De Rijke, Gianmaria Silvello, Maristella Agosti, Toine Bogers, Martin Braschler, Paul Buitelaar, Khalid Choukri, Giorgio Maria Di Nunzio, Pamela Forner, Karin Friberg Heppin, Preben Hansen, Anni Järvelin, Birger Larsen, Ivano Masiero, Henning Müller, Simone Peruzzo, Florina Piori, Guiseppa Santucci och Elaine Thoms. PROMISE retreat report prospects and opportunities for information access evaluation. *ACM SIGIR Forum*, 46(2):60–84, 2013.
 25. Karin Friberg Heppin. Search using semantic FrameNet frames as variables. I: *Proceedings of Sixth Workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR 2013), held at CIKM 2013, October 27th-November 1st, San Francisco, CA, USA*, ss 25–28, New York, NY, USA, 2013. Association for Computing Machinery (ACM).
 26. Karin Friberg Heppin och Anni Järvelin. Assessors assessing assessments. I: *Proceedings of the 4th International Louhi Workshop on Health Document Text Mining and Information Analysis (Louhi 2013)*, 2013.
 27. Birgitta Frändberg och Sofie Johansson Kokkinakis. Högstadielävers användning av naturvetenskapligt språkbruk i kemiämnet i TIMSS. *Utbildning och Demokrati*, 22(3):53–68, 2013.
 28. Sucheta Ghosh, Sara Tonelli och Richard Johansson. Mining fine-grained opinion expressions with shallow parsing. I: *Proceedings of the International Conference Recent Advances in Natural Language Processing*, ss 302–310, 2013.
 29. Thierry Hamon, Natalia Grabar och Dimitrios Kokkinakis. Medication extraction and guessing in Swedish, French and English. I: *MEDINFO 2013 - Proceedings of the 14th World Congress on Medical and Health Informatics, 20-13 August 2013, Copenhagen, Denmark*, band 192 av *Studies in Health Technology and Informatics*. IOS Press, 2013.
 30. Katarina Heimann Mühlenbock. *I see what you mean*. University of Gothenburg, Gothenburg, Sweden, 2013.
 31. Richard Johansson. Training parsers on incompatible treebanks. I: *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, ss 127–137, 2013.
 32. Qi Ju, Alessandro Moschitti och Richard Johansson. Learning to rank

- from structures in hierarchical text classification. I: P. Serdyukov et al., redaktör, *Advances in Information Retrieval; 35th European Conference on IR Research, ECIR 2013, Moscow, Russia, March 24-27*, ss 183–194, Berlin, Germany, 2013. Springer.
33. Adam Kilgariff, Frieda Charalabopoulou, Maria Gavrilidou, Janne Bondi Johannessen, Saoussen Khalil, Sofie Johansson Kokkinakis, Robert Lew, Serge Sharoff, R. Vadlapudi och Elena Volodina. Corpus-based vocabulary lists for language learners for nine languages. *Language resources and evaluation*, 2013.
 34. Dimitrios Kokkinakis. Annotation of interpersonal relations in Swedish prose fiction. I: *Proceedings of the 3rd Workshop on Annotation of Corpora for Research in the Humanities (ACRH-3)*. Sofia, Bulgaria, ss 37–47, 2013.
 35. Dimitrios Kokkinakis. Figurative language in Swedish clinical texts. I: *Computational Semantics in Clinical Text workshop. Part of the 10th International Conference on Computational Semantics, Potsdam, Germany*, s 6, 2013.
 36. Dimitrios Kokkinakis. Medical event extraction using frame semantics - challenges and opportunities. I: *Proceedings of the 14th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*, Samos, Greece, 2013.
 37. Dimitrios Kokkinakis och Ann-Marie Eklund. Query logs as a corpus. I: Andrew Hardie och Robbie Love, redaktörer, *Proceedings of the 7th international Corpus Linguistics conference (CL2013)*, Lancaster, UK, ss 329–330, 2013.
 38. Dimitrios Kokkinakis och Mats Malm. A macroanalytic view of Swedish literature using topic modeling. I: Andrew Hardie och Robbie Love, redaktörer, *Proceedings of the 7th international Corpus Linguistics conference (CL2013)*, Lancaster, UK, ss 144–146, 2013.
 39. D. Oelke, Dimitrios Kokkinakis och D. A. Keim. Fingerprint matrices: Uncovering the dynamics of social networks in prose literature. *Computer Graphics Forum*, 32(3):371–380, 2013.
 40. Bolette Pedersen, Lars Borin, Markus Forsberg, Neeme Kahusk, Krister Lindén, Jyrki Niemi, Niklas Nisbeth, Lars Nygaard, Heili Orav, Eiríkur Rögnvaldsson, Mitchel Seaton, Kadri Vider och Kaarlo Voionmaa. Nordic and Baltic wordnets aligned and compared through “WordTies”. I: *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2013)*, May 22-24, 2013, Oslo, Norway, nummer 16 i NEALT Proceedings, ss 147–162, Linköping, Sweden, 2013. Linköping

University Electronic Press.

41. Dijana Pijetlovic och Elena Volodina. Developing a Swedish spelling game on an ICALL platform. I: Linda Bradley och Sylvie Thoësny, redaktörer, *20 Years of EUROCALL: Learning from the Past, Looking to the Future. Proceedings of the 2013 EUROCALL Conference, Évora, Portugal, 11-14 September 2013*, Dublin, Ireland; Voillans, France, 2013. Research-publishing.net.
42. Ildikó Pilán, Elena Volodina och Richard Johansson. Automatic selection of suitable sentences for language learning exercises. I: Linda Bradley och Sylvie Thoësny, redaktörer, *20 Years of EUROCALL: Learning from the Past, Looking to the Future. Proceedings of the 2013 EUROCALL Conference, Évora, Portugal, 11-14 September 2013*, Dublin, Ireland; Voillans, France, 2013. Research-publishing.net.
43. Taraka Rama. Phonotactic diversity predicts the time depth of the world's language families. *PLoS ONE*, 8(5):e63238, 2013.
44. Taraka Rama, Prasanth Kolachina och Sudheer Kolachina. Two methods for automatic cognate identification. I: Thomas Wielfaert, Kris Heylen och Dirk Speelman, redaktörer, *Proceedings of the 5th Conference on Quantitative Investigations in Theoretical Linguistics (QITL-5), University of Leuven, Belgium, 12-14 September 2013*, ss 76–80, Leuven, Belgium, 2013.
45. Taraka Rama och Sudheer Kolachina. Distance-based phylogenetic inference algorithms in the subgrouping of Dravidian languages. I: Lars Borin och Anju Saxena, redaktörer, *Approaches to Measuring Linguistic Differences*. De Gruyter Mouton, Berlin, Germany, 2013.
46. Johan Roxendal. State chart XML (SCXML): State machine notation for control abstraction – W3C last call working draft 1 august 2013. Teknisk rapport, MIT, Cambridge, MA, USA, 2013.
47. Johan Roxendal. State chart XML (SCXML): State machine notation for control abstraction – W3C working draft 6 december 2012. Teknisk rapport, MIT, Cambridge, MA, USA, 2013.
48. Anju Saxena och Lars Borin. Carving Tibeto-Kanauri by its joints: Using basic vocabulary lists for genetic grouping of languages. I: *Approaches to Measuring Linguistic Differences*, ss 175–198. De Gruyter Mouton, Berlin, Germany, 2013.
49. Emma Sköldberg, Linnéa Bäckström, Lars Borin, Markus Forsberg, Benjamin Lyngfelt, Leif-Jöran Olsson, Julia Prentice, Rudolf Rystedt, Sofia Tingsell och Jonatan Uppström. Between grammars and dictionaries: a Swedish constructicon. I: I. Kosem, J. Kallas, P. Gantar, S. Krek,

- M. Langemets och M. Tuulik, redaktörer, *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference, 17-19 October, Tallinn, Estonia*, ss 310–327, Ljubljana/Tallinn, 2013. Trojina, Institute for Applied Slovene Studies/Eesti Keele Instituut.
50. Frida Smith, Eva Carlsson, Dimitrios Kokkinakis, Markus Forsberg, Karl Kodeda, Richard Sawatzky, Febe Friberg och Joakim Ohlén. Readability, suitability and comprehensibility in patient education materials for Swedish patients with colorectal cancer undergoing elective surgery: A mixed method design. *Patient education and counseling*, 2013.
 51. Kaarlo Voionmaa och Karin Friberg Heppin. Use of support verbs in FrameNet annotations. I: I. Kosem, J. Kallas, P. Gantar, S. Krek, M. Langemets och M. Tuulik, redaktörer, *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference, 17-19 October, Tallinn, Estonia*, Ljubljana/Tallinn, 2013. Trojina, Institute for Applied Slovene Studies/Eesti Keele Instituut.
 52. Elena Volodina, Lars Borin och Hrafn Loftsson, redaktörer. *Proceedings of the second workshop on NLP for computer-assisted language learning at NODALIDA 2013 May 22-24, 2013, Oslo, Norway*, nummer 17 i NEALT Proceedings, Linköping, Sweden, 2013. Linköping University Press.
 53. Elena Volodina och Sofie Johansson Kokkinakis. Compiling a corpus of CEFR-related texts. I: *Language Testing in Europe: Time for a New Framework? Proceedings of LT-CEFR2013, 27-29 May 2013, Antwerp, Belgium*, ss 248–259, Antwerpen, Belgium, 2013. Universiteit Antwerpen.
 54. Elena Volodina, Dijana Pijetlovic, Ildikó Pilán och Sofie Johansson Kokkinakis. Towards a gold standard for Swedish CEFR-based ICALL. I: *Proceedings of the Second Workshop on NLP for Computer-Assisted Language Learning, held at NODALIDA 2013, May 22-24, 2013, Oslo, Norway*, nummer 17 i NEALT Proceedings, Linköping, Sweden, 2013. Linköping University Electronic Press.