

Krister Lindén  
Tutkimusjohtaja  
FIN-CLARINin kansallinen koordinaattori  
Ohjaaja tohtoriohjelmassa



Kieliteknologia  
Digitaalisten ihmistieteiden osasto  
Centre of Excellence in Ancient Near Eastern Empires (ANEE)

Postiosoite:  
PL 24 (Unioninkatu 40)  
00014  
Suomi

Sposti: krister.linden@helsinki.fi  
Matkapuhelin: +358504154717  
Puhelin: +358294129303

## Pätevyyydet

### Vastuullinen tutkija

Ajanjakso : 01.06.2010 - 31.12.2017 organisaatiossa Nykykielten laitos 2010-2017

### Vastuullinen tutkija

Ajanjakso : 15.05.2018 - \* organisaatiossa Digitaalisten ihmistieteiden osasto

## Työsuhteet

### tutkimusjohtaja

Digitaalisten ihmistieteiden osasto  
Helsingin yliopisto  
Suomi  
1 tammikuuta 2018 → present

### dosenttuuri

Tiedekunnan yhteiset  
Helsingin yliopisto  
Suomi  
1 tammikuuta 2010 → present

### Centre of Excellence in Ancient Near Eastern Empires (ANEE)

Helsingin yliopisto  
Suomi  
1 tammikuuta 2018 → 31 joulukuuta 2025

## Tutkimustuotos

2021

### Findings of the VarDial Evaluation Campaign 2021

Chakravarthi, B. R., Gaman, M., Ionescu, R. T., Jauhiainen, H., Jauhiainen, T., Linden, K., Ljubešić, N., Partanen, N., Priyadharshini, R., Purschke, C., Rajagopal, E., Scherrer, Y. & Zampieri, M., huhtikuuta 2021, *Proceedings of the 8th VarDial Workshop on NLP for Similar Languages, Varieties and Dialects*. Zampieri, M., Nakov, P., Ljubešić, N., Tiedemann, J., Scherrer, Y. & Jauhiainen, T. (toim.). Stroudsburg: The Association for Computational Linguistics, s. 1-11  
11 Sivumäärä

### Fear in Akkadian Texts: New Digital Perspectives on Lexical Semantics

Svärd, S., Alstola, T., Jauhiainen, H., Sahala, A. & Linden, K., 2021, *The Expression of Emotions in Ancient Egypt and Mesopotamia*. Hsu, S-W. & Llop-Raduà, J. (toim.). Leiden: Brill, s. 470-502 33 Sivumäärä (Culture and History of the Ancient Near East; painos 116).

### **Naive Bayes-based Experiments in Romanian Dialect Identification**

Jauhiainen, T., Jauhiainen, H. & Linden, K., 2021, *Proceedings of the 8th VarDial Workshop on NLP for Similar Languages, Varieties and Dialects*. Zampieri, M., Nakov, P., Ljubešić, N., Tiedemann, J., Scherrer, Y. & Jauhiainen, T. (toim.). Stroudsburg: ACL Anthology, s. 76-83 8 Sivumäärä

### **Naive Bayes-based Experiments in Romanian Dialect Identification**

Jauhiainen, T., Jauhiainen, H. & Linden, K., 2021.

### **Suomalais-ugrilaiset kielet ja internet -projekti 2013-2019**

Jauhiainen, T., Jauhiainen, H. & Linden, K., 2021, *Multilingual Facilitation*. Hämäläinen, M., Partanen, N. & Alnajjar, K. (toim.). Helsinki: University of Helsinki

2020

### **A Report on the VarDial Evaluation Campaign 2020**

Gaman, M., Hovy, D., Ionescu, R. T., Jauhiainen, H., Jauhiainen, T., Linden, K., Ljubešić, N., Partanen, N., Purschke, C., Scherrer, Y. & Zampieri, M., 13 joulukuuta 2020, *Proceedings of the 7th Workshop on NLP for Similar Languages, Varieties and Dialects*. Zampieri, M., Nakov, P., Ljubešić, N., Tiedemann, J. & Scherrer, Y. (toim.). International Committee on Computational Linguistics, s. 1-14

### **Optical character recognition with neural networks and post-correction with finite state methods**

Drobac, S. & Linden, K., joulukuuta 2020, julkaisussa: *International Journal on Document Analysis and Recognition*. 23, 4, s. 279-295 17 Sivumäärä

### **Improving Word Association Measures in Repetitive Corpora with Context Similarity Weighting**

Sahala, A. & Lindén, K., marraskuuta 2020, *Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2020, Volume 1: KDIR, Budapest, Hungary, November 2-4, 2020*. Fred, A. L. N. & Filipe, J. (toim.). Setúbal: SCITEPRESS Science And Technology Publications, s. 48-58 11 Sivumäärä

### **Akkadian Treebank for early Neo-Assyrian Royal Inscriptions**

Luukko, M., Sahala, A., Hardwick, S. & Lindén, K., 27 lokakuuta 2020, *Proceedings of the 19th International Workshop on Treebanks and Linguistic Theories*. Evang, K., Kallmeyer, L., Ehren, R., Petitjean, S., Seyffarth, E. & Seddah, D. (toim.). Stroudsburg: The Association for Computational Linguistics, s. 124-134 11 Sivumäärä

### **CLARIN contractual framework for sharing language data: the perspective of personal data protection**

Kelli, A., Linden, K., Vider, K., Kamocki, P., Tavast, A., Birštonas, R., Tavits, G., Keskküla, M. & Labropoulou, P., 5 lokakuuta 2020, *Proceedings of CLARIN Annual Conference 2020. 05 – 07 October 2020, Online Edition*. Navarretta, C. & Eskevich, M. (toim.). Utrecht: CLARIN ERIC, s. 171-177 7 Sivumäärä

### **A CLARIN Contractual Framework for Sharing Personal Data for Scientific Research**

Linden, K., Kelli, A. & Nousias, A., 3 heinäkuuta 2020, *Selected Papers from the CLARIN Annual Conference 2019*. Simov, K. & Eskevich, M. (toim.). Linköping: Linköping University Electronic Press, s. 75–84 10 Sivumäärä 10. (Linköping Electronic Conference Proceedings; nro 172).

### **The Impact of Copyright and Personal Data Laws on the Creation and Use of Models for Language Technologies**

Kelli, A., Tavast, A., Lindén, K., Vider, K., Birštonas, R., Labropoulou, P., Kull, I., Tavits, G., Värvi, A., Stranák, P. & Hajic, J., 3 heinäkuuta 2020, *Selected Papers from the CLARIN Annual Conference 2019*. Simov, K. & Eskevich, M. (toim.). Linköping: Linköping University Electronic Press, s. 53-65 13 Sivumäärä (Linköping Electronic Conference Proceedings; nro 172).

### **Automated Phonological Transcription of Akkadian Cuneiform Text**

Sahala, A., Silfverberg, M., Arppe, A. & Linden, K., 17 toukokuuta 2020, *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*. Calzolari ... [ et al.], N. (toim.). Paris: European Language Resources Association (ELRA), s. 3528-3934 7 Sivumäärä

### **BabyFST: Towards a Finite-State Based Computational Model of Ancient Babylonian**

Sahala, A., Silfverberg, M., Arppe, A. & Linden, K., 17 toukokuuta 2020, *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*. Calzolari ... [et al.], N. (toim.). Paris: European Language Resources Association (ELRA), s. 3886-3894 9 Sivumäärä

### **The European Language Technology Landscape in 2020: Language-Centric and Human-Centric AI for Cross-Cultural Communication in Multilingual Europe**

Rehm, G., Marheinecke, K., Hegele, S., Piperidis, S., Bontcheva, K., Hajic, J., Choukri, K., Vasiljevs, A., Backfried, G., Prinz, C., Gomez-Perez, J. M., Meertens, L., Lukowicz, P., van Genabith, J., Lösch, A., Slusallek, P., Irgens, M., Gatellier, P., Köhler, J., Le Bars, L. & 27 muuta, Anastasiou, D., Auksoiūtė, A., Bel, N., Branco, A., Budin, G., Daelemans, W., De Smedt, K., Garabik, R., Gavriilidou, M., Gromann, D., Koeva, S., Krek, S., Krstev, C., Lindén, K., Magnini, B., Odijk, J., Ogrodniczuk, M., Rögnvaldsson, E., Rosner, M., Pedersen, B., Skadina, I., Tadić, M., Tufi, D., Váradi, T., Vider, K., Way, A. & Yvon, F., 1 toukokuuta 2020, *Proceedings of The 12th Language Resources and Evaluation Conference*. Calzolari [et al.], N. (toim.). Paris: European Language Resources Association (ELRA), s. 3322-3332 11 Sivumäärä

### **A Finnish news corpus for named entity recognition**

Ruokolainen, T., Kauppinen, P., Silfverberg, M. & Lindén, K., maaliskuuta 2020, julkaisussa: *Language Resources and Evaluation*. 54, 1, s. 247-272 26 Sivumäärä

### **Building Web Corpora for Minority Languages**

Jauhiainen, H., Jauhiainen, T. & Linden, K., 2020, *Proceedings of the 12th Web as Corpus Workshop*. Barbaresi, A., Bildhauer, F., Schäfer, R. & Stemle, E. (toim.). Stroudsburg: The Association for Computational Linguistics, s. 23-32 10 Sivumäärä

### **Compatibility of business models of Estonian language technology entrepreneurs with regulatory framework on the use of digital content**

Kull, I., Vider, K., Mets, T., Kelli, A. & Linden, K., 2020, julkaisussa: *International Comparative Jurisprudence*. 6, 1, s. 70-83 14 Sivumäärä

### **Experiments in Language Variety Geolocation and Dialect Identification**

Jauhiainen, T., Jauhiainen, H. & Linden, K., 2020, *Proceedings of the 7th Workshop on NLP for Similar Languages, Varieties and Dialects*. Zampieri, M., Nakov, P., Ljubešić, N., Tiedemann, J. & Scherrer, Y. (toim.). International Committee on Computational Linguistics (ICCL), s. 220-231 12 Sivumäärä

### **Impact of Legal Status of Data on Development of Data-Intensive Products: Example of Language Technologies**

Kelli, A., Tavast, A., Linden, K., Bristonas, R., Labropoulou, P., Vider, K., Kull, I., Tavits, G., Värvi, A. & Mantrov, V., 2020, *Legal Science: Functions, Significance and Future in Legal Systems II : Collection of Research Papers in Conjunction with the 7th International Scientific Conference of the Faculty of Law of the University of Latvia (16-18 October 2019, Riga)*. Damberg, A. (toim.). Riga: University of Latvia press, s. 383-400 18 Sivumäärä

### **Keeleandmete õigusliku režiimi mõju nende abil loodud keelemudelitele**

Kelli, A., Vider, K., Tavast, A., Lindén, K., Bristonas, R., Labropoulou, P., Värvi, A., Kull, I., Tavits, G. & Ginter, C., 2020, julkaisussa: *Eesti rakenduslingvistika ühingu aastaraamat*. 16, s. 59-76 18 Sivumäärä

### **Uralic Language Identification (ULI) 2020 shared task dataset and the Wanca 2017 corpora**

Jauhiainen, T., Jauhiainen, H., Partanen, N. & Linden, K., 2020, *Proceedings of the 7th Workshop on NLP for Similar Languages, Varieties and Dialects*. Zampieri, M., Nakov, P., Ljubešić, N., Tiedemann, J. & Scherrer, Y. (toim.). International Committee on Computational Linguistics (ICCL), s. 173-185 13 Sivumäärä

### **Vestlusrobotid ja autoriõigus**

Kelli, A., Tavast, A. & Linden, K., 2020, julkaisussa: *Juridica*. 5, s. 345-354 10 Sivumäärä

### **The Extent of Legal Control over Language Data: the Case of Language Technologies**

Kelli, A., Tavast, A., Linden, K., Vider, K., Birštonas, R., Labropoulou, P., Kull, I., Tavits, G. & Värvi, A., 30 syyskuuta 2019, *Proceedings of CLARIN Annual Conference 2019*. Simov, K. & Eskevich, M. (toim.). Utrecht: CLARIN ERIC, s. 69-74 6 Sivumäärä

### **To Ask or not to Ask: Informed Consent to Participate and Using Data in the Public Interest**

Linden, K., Kelli, A. & Nousias, A., 30 syyskuuta 2019, *Proceedings of CLARIN Annual Conference 2019*. Simov, K. & Eskevich, M. (toim.). Leipzig: CLARIN ERIC, s. 56-60 5 Sivumäärä

### **Language Model Adaptation for Language and Dialect Identification of Text**

Jauhiainen, T. S., Linden, B. K. J. & Jauhiainen, H. A., syyskuuta 2019, julkaisussa: *Natural Language Engineering*. 25, 5, s. 561-583 23 Sivumäärä, 135132491900038.

### **Automatic Language Identification in Texts: A Survey**

Jauhiainen, T., Lui, M., Zampieri, M., Baldwin, T. & Lindén, K., 25 elokuuta 2019, julkaisussa: *Journal of Artificial Intelligence Research*. 65, s. 675-782 108 Sivumäärä

### **Processing personal data without the consent of the data subject for the development and use of language resources**

Kelli, A., Lindén, K., Vider, K., Kamocki, P., Birštonas, R., Calamai, S., Labropoulou, P., Gavriilidou, M. & Stranák, P., 28 toukokuuta 2019, *Selected papers from the CLARIN Annual Conference 2018, Pisa, 8-10 October 2018*. Skadina, I. & Eskevich, M. (toim.). Linköping: Linköping University Electronic Press, s. 72-82 11 Sivumäärä (Linköping Electronic Conference Proceedings; nro 159).

### **Discriminating between Mandarin Chinese and Swiss-German varieties using adaptive language models**

Jauhiainen, T. S., Jauhiainen, H. A. & Linden, B. K. J., 30 huhtikuuta 2019, *Proceedings of the Sixth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2019)*. Stroudsburg: The Association for Computational Linguistics, s. 178-187 10 Sivumäärä

### **Language and Dialect Identification of Cuneiform Texts**

Jauhiainen, T. S., Jauhiainen, H. A., Alstola, T. & Linden, B. K. J., 30 huhtikuuta 2019, *Proceedings of the Sixth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2019)*. Stroudsburg: The Association for Computational Linguistics, s. 89-98 10 Sivumäärä

### **Aššur and His Friends: A Statistical Analysis of Neo-Assyrian Texts**

Alstola, T., Zaia, S., Sahala, A., Jauhiainen, H., Svärd, S. & Linden, K., 2019, julkaisussa: *Journal of Cuneiform Studies*. 71, s. 159-180 22 Sivumäärä

### **FinnTransFrame: translating frames in the FinnFrameNet project**

Lindén, K., Haltia, H., Laine, A., Luukkonen, J., Piitulainen, J. & Väisänen, N., 2019, julkaisussa: *Language Resources and Evaluation*. 53, 1, s. 141-171 31 Sivumäärä

### **Improving OCR of historical newspapers and journals published in Finland**

Drobac, S., Kauppinen, P. & Linden, K., 2019, *Proceedings of the 3rd International Conference on Digital Access to Textual Cultural Heritage*. New York: ACM, s. 97-102 6 Sivumäärä

### **Optical Font Family Recognition using a Neural Network**

Drobac, S. & Lindén, K., 2019, *Proceedings of the Research Data And Humanities (Rdhum) 2019 Conference: Data, Methods And Tools*. Jantunen, J. H., Brunni, S., Kunnas, N., Palviainen, S. & Västi, K. (toim.). Oulu: Studia Humaniora Ouluensia, s. 115-123 9 Sivumäärä (Studia humaniora ouluensia; nro 17).

### **Wanca in Korp: Text corpora for underresourced Uralic languages**

Jauhiainen, H., Jauhiainen, T. & Linden, K., 2019, *Proceedings of the Research data and humanities (RDHUM) 2019 conference : data, methods and tools*. Jantunen, J. H., Brunni, S., Kunnas, N., Palviainen, S. & Västi, K. (toim.). Oulu: University of Oulu, s. 21-40 20 Sivumäärä (Studia Humaniora Ouluensia; nro 17).

2018

### **Vad får AI för bild av mänskligheten?**

Linden, B. K. J., 26 marraskuuta 2018, julkaisussa: Yliopistolainen : Helsingin yliopiston henkilöstölehti. 4/2018, s. 9 1 Sivumäärä

### **Challenges of transformation of research data into open data: The perspective of social sciences and humanities**

Kelli, A., Mets, T., Vider, K., Värvi, A., Jonsson, L., Lindén, K. & Birtonas, R., syyskuuta 2018, julkaisussa: International Journal of Technology Management & Sustainable Development. 17, 3, s. 227-251 25 Sivumäärä

### **Semantic Domains in Akkadian Text**

Svärd, S. S., Jauhiainen, H. A., Linden, B. K. J. & Sahala, A. J. A., 7 elokuuta 2018, *CyberResearch on the Ancient Near East and Neighboring Regions: Case Studies on Archaeological Data, Objects, Texts, and Digital Archiving*. Juloux, V. B., Gansell, A. R. & di Ludovico, A. (toim.). Leiden: Brill, s. 224-256 33 Sivumäärä (Digital Biblical Studies; nro 2).

### **HeLI-based Experiments in Discriminating Between Dutch and Flemish Subtitles**

Jauhiainen, T. S., Jauhiainen, H. A. & Linden, B. K. J., elokuuta 2018, *Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2018)*. Zampieri, M., Nakov, P., Ljubešić, N., Tiedemann, J., Malmasi, S. & Ali, A. (toim.). Santa Fe: The Association for Computational Linguistics, s. 137-144 8 Sivumäärä

### **HeLI-based Experiments in Swiss German Dialect Identification**

Jauhiainen, T. S., Jauhiainen, H. A. & Linden, B. K. J., elokuuta 2018, *Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2018)*. Zampieri, M., Nakov, P., Ljubešić, N., Tiedemann, J., Malmasi, S. & Ali, A. (toim.). Santa Fe: The Association for Computational Linguistics, s. 254-262 9 Sivumäärä

### **Iterative Language Model Adaptation for Indo-Aryan Language Identification**

Jauhiainen, T. S., Jauhiainen, H. A. & Linden, B. K. J., elokuuta 2018, *Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2018)*. Zampieri, M., Nakov, P., Ljubešić, N., Tiedemann, J., Malmasi, S. & Ali, A. (toim.). Santa Fe: The Association for Computational Linguistics, s. 66-75 10 Sivumäärä

### **The Dagstuhl Perspectives Workshop on Performance Modeling and Prediction**

Ferro, N., Fuhr, N., Grefenstette, G., Konstan, J. A., Castells, P., Daly, E. M., Declerck, T., Ekstrand, M. D., Geyer, W., Gonzalo, J., Kuflik, T., Lindén, K., Magnini, B., Nie, J.-Y., Perego, R., Shapira, B., Soboroff, I., Tintarev, N., Verspoor, K., Willemsen, M. C. & 1 muuta, Zobel, J., kesäkuuta 2018, julkaisussa: SIGIR Forum. 52, 1, s. 91-101 11 Sivumäärä

### **Implementation of an Open Science Policy in the context of management of CLARIN language resources: a need for changes?**

Kelli, A., Lindén, K., Vider, K., Labropoulou, P., Ketzan, E., Kamocki, P. & Stranák, P., 16 toukokuuta 2018, *Selected papers from the CLARIN Annual Conference 2017, Budapest, 18–20 September 2017*. Piasecki, M. (toim.). Linköping: Linköping University Electronic Press, Vuosikerta 147. s. 102-111 10 Sivumäärä 009. (Linköping Electronic Conference Proceedings; nro 147).

### **Tieteen kansallinen termipankki (TTP) ja tiedonlouhinnan apuneuvot**

Hardwick, S., Enqvist, E. J., Onikki-Rantajääskö, T. A. & Linden, B. K. J., 7 toukokuuta 2018. 1 Sivumäärä.

### **Rethinking Summarization and Storytelling for Modern Social Multimedia**

Rudinac, S., Chua, T.-S., Diaz-Ferreyra, N., Friedland, G., Gornostaja, T., Huet, B., Kaptein, R., Lindén, K., Moens, M.-F., Peltonen, J., Redi, M., Schedl, M., Shamma, D. A., Smeaton, A. & Xie, L., 28 tammikuuta 2018, *MultiMedia Modeling: 24th International Conference, MMM 2018, Bangkok, Thailand, February 5-7, 2018, Proceedings, Part I*. Schoeffmann, K., Chalidabhongse, T. H., Ngo, C. W., Aramvith, S., O'Connor, N. E., Ho, Y.-S., Gabbouj, M. & Elgammal, A. (toim.). Cham: Springer International Publishing AG, Vuosikerta 1. s. 632-644 13 Sivumäärä (LNCS; nro 10704).

### **From Evaluating to Forecasting Performance: How to Turn Information Retrieval, Natural Language Processing and Recommender Systems into Predictive Sciences**

Ferro, N., Fuhr, N., Grefenstette, G., Konstan, J. A., Castells, P., Daly, E. M., Declerck, T., Ekstrand, M. D., Geyer, W., Gonzalo, J., Kuflik, T., Lindén, K., Magnini, B., Nie, J.-Y., Perego, R., Shapira, B., Soboroff, I., Tintarev, N., Verspoor, K., Willemsen, M. C. & 1 muuta, Zobel, J., 2018, julkaisussa: Dagstuhl manifestos. 7, 1, s. 96-139 44 Sivumäärä

### **Keeleressursside loomise ja kasutamise seonduvaid isikuandmete kaitse küsimusi**

Kelli, A., Vider, K., Kull, I., Siil, T., Lindén, K., Tavast, A., Värvi, A., Ginter, C. & Meister, E., 2018, julkaisussa: Eesti rakenduslingvistika ühingu aastaraamat. 14, s. 77-94 18 Sivumäärä

### **Rethinking Summarization and Storytelling for Modern Social Multimedia**

Rudinac, S., Chua, T-S., Diaz-Ferreira, N., Friedland, G., Gornostaja, T., Huet, B., Kaptein, R., Linden, K., Moens, M-F., Peltonen, J., Redi, M., Schedl, M., Shamma, D. A., Smeaton, A. & Xie, L., 2018, s. 141-153. 13 Sivumäärä.

## **2017**

### **FinnFN 1.0: The Finnish frame semantic database**

Lindén, K., Haltia, H., Luukkonen, J., Laine, A., Roivainen, H. & Väisänen, N., 14 elokuuta 2017, julkaisussa: Nordic journal of linguistics. 40, 3, s. 1-25 25 Sivumäärä

### **Evaluating HeLI with non-linear mappings**

Jauhiainen, T. S., Linden, B. K. J. & Jauhiainen, H. A., 2017, *Fourth Workshop on NLP for Similar Languages, Varieties and Dialects - Proceedings of the Workshop*. Stroudsburg: The Association for Computational Linguistics, s. 102-108 7 Sivumäärä

### **Evaluation of language identification methods using 285 languages**

Jauhiainen, T. S., Linden, B. K. J. & Jauhiainen, H. A., 2017, *21st Nordic Conference of Computational Linguistics: Proceedings of the Conference*. Tiedemann, J. (toim.). Linköping: Linköping University Electronic Press, s. 183-191 9 Sivumäärä (Linköping Electronic Conference Proceedings; nro 31).

### **OCR and post-correction of historical Finnish texts**

Drobac, S., Kauppinen, P. S. & Linden, B. K. J., 2017, *Proceedings of the 21st Nordic Conference on Computational Linguistics, NoDaLiDa, 22-24 May 2017, Gothenburg, Sweden*. Tiedemann, J. (toim.). Linköping: Linköping University Electronic Press, s. 70-76 7 Sivumäärä (Linköping Electronic Conference Proceedings; nro 131).

## **2016**

### **FinnPos: an open-source morphological tagging and lemmatization toolkit for Finnish**

Silfverberg, M., Ruokolainen, T., Linden, K. & Kurimo, M., joulukuuta 2016, julkaisussa: Language Resources and Evaluation. 50, 4, s. 863-878 16 Sivumäärä

### **In-Document Adaptation for a Human Guided Automatic Transcription Service**

Mansikkaniemi, A., Kurimo, M. & Linden, B. K. J., 13 elokuuta 2016, *Speech and Computer : 18th International Conference, SPECOM 2016, Budapest, Hungary, August 23-27, 2016, Proceedings*. Ronzhin, A., Potapova, R. & Németh, G. (toim.). Cham: Springer International Publishing AG, s. 395-402 8 Sivumäärä (Lecture Notes in Computer Science; painos 9811).

### **Data-Driven Spelling Correction using Weighted Finite-State Methods**

Silfverberg, M., Kauppinen, P. & Linden, K., 12 elokuuta 2016, *The 54th Annual Meeting of the Association for Computational Linguistics: Proceedings of the SIGFSM Workshop on Statistical NLP and Weighted Automata*. Stroudsburg, PA: ACL, s. 51-59 9 Sivumäärä

### **The strategic impact of META-NET on the regional, national and international level**

Rehm, G., Uszkoreit, H., Ananiadou, S., Bel, N., Bieleveciene, A., Borin, L., Branco, A., Budin, G., Calzolari, N., Daelemans, W., Garabik, R., Grobelnik, M., Garcia-Mateo, C., van Genabith, J., Hajic, J., Hernaez, I., Judge, J., Koeva, S., Krek, S., Krstev, C. & 24 muuta, Linden, K., Magnini, B., Mariani, J., McNaught, J., Melero, M., Monachini, M., Moreno, A., Odijk, J., Ogrodniczuk, M., Pezik, P., Piperidis, S., Przepiorkowski, A., Rognvaldsson, E., Rosner, M., Pedersen, B. S., Skadina, I., De Smedt, K., Tadic, M., Thompson, P., Tufis, D., Varadi, T., Vasiljevs, A., Vider, K. & Zabarskaite, J., kesäkuuta 2016, julkaisussa: Language Resources and Evaluation. 50, 2, s. 351-374 24 Sivumäärä

**Eduskunnan täysistuntovideoiden ja keskustelupöytäkirjojen automaattinen kohdistaminen AaltoASR-työkalulla**  
Mansikkaniemi, A., Kurimo, M., Lennes, M. E. & Linden, B. K. J., 23 toukokuuta 2016.

#### **The Regulatory and Contractual Framework as an Integral Part of the CLARIN Infrastructure**

Kelli, A., Vider, K. & Lindén, K., 11 huhtikuuta 2016, *Selected Papers from the CLARIN Annual Conference 2015, October 14–16, 2015, Wrocław, Poland: NEALT Proceedings Series*. DeSmedt, K. (toim.). Linköping: Linköping University Electronic Press, Vuosikerta 123. s. 13-24 12 Sivumäärä 002. (Linköping Electronic Conference Proceedings; painos 123).

#### **Digitaalsete keeleressursside loomist ja kasutamist määrav õiguslik raamistik Eestis ja selle ühildumine CLARIN-i infrastruktuuriga**

Kelli, A., Vider, K., Pisuke, H. & Linden, B. K. J., 2016, julkaisussa: Eesti rakenduslingvistika ühingu aastaraamat. 12, s. 81-98 18 Sivumäärä

#### **HeLI, a Word-Based Backoff Method for Language Identification**

Jauhiainen, T. S., Linden, B. K. J. & Jauhiainen, H. A., 2016, *Proceedings of the Third Workshop on NLP for Similar Languages, Varieties and Dialects: VarDial3, Osaka, Japan, December 12 2016*. s. 153-162 10 Sivumäärä

2015

#### **Using HFST—Helsinki Finite-State Technology for Recognizing Semantic Frames**

Lindén, K., Hardwick, S., Silfverberg, M. & Axelson, E., 9 joulukuuta 2015, *Systems and Frameworks for Computational Morphology*. Springer International Publishing AG, s. 124-136 13 Sivumäärä (Communications in Computer and Information Science; painos 537).

#### **Regulatory and Contractual Framework as an Integral Part of CLARIN Infrastructure: the Estonian and Finnish Perspectives**

Kelli, A., Vider, K. & Linden, K., lokakuuta 2015, *CLARIN Annual Conference 2015: Book of Abstracts*. De Smedt, K. (toim.). Wrocław, Poland: CLARIN ERIC, s. 32-36 5 Sivumäärä

#### **Extracting Semantic Frames using hfst-pmatch**

Hardwick, S., Silfverberg, M. & Linden, K., toukokuuta 2015, *Proceedings of the 20th Nordic Conference of Computational Linguistics: NODALIDA 2015*. Magyesi, B. (toim.). Linköping: Linköping University Electronic Press, s. 305-308 4 Sivumäärä 109:042. (Linköping Electronic Conference Proceedings).

#### **Automated Lossless Hyper-Minimization for Morphological Analyzers**

Drobac, S., Silfverberg, M. & Linden, K., 22 huhtikuuta 2015, *Proceedings of the 12th International Conference on Finite-State Methods and Natural Language Processing 2015: Collected papers*. Hanneforth, T. & Wurm, C. (toim.). Stroudsburg: ACL, 5 Sivumäärä

#### **The Finno-Ugric Languages and the Internet project**

Jauhiainen, H., Jauhiainen, T. & Linden, K., 15 tammikuuta 2015, *First International Workshop on Computational Linguistics for Uralic Languages: Proceedings of the Workshop*. Pirinen, T., Tyers, F. & Trosterud, T. (toim.). Tromsø: Septentrio Academic Publishing, Vuosikerta 2. s. 87–98 12 Sivumäärä (Septentrio Conference Series; painos 2015, nro 2).

#### **Discriminating similar languages with token-based backoff**

Jauhiainen, T., Jauhiainen, H. & Linden, K., 2015, *Proceedings of the Joint Workshop on Language Technology for Closely Related Languages, Varieties and Dialects*. The Association for Computational Linguistics, s. 44-51 8 Sivumäärä

#### **Language Set Identification in Noisy Synthetic Multilingual Documents**

Jauhiainen, T. S., Linden, K. & Jauhiainen, H. A., 2015, *Computational Linguistics and Intelligent Text Processing*. Gelbukh, A. (toim.). Springer International Publishing AG, Vuosikerta Part I. s. 633-643 11 Sivumäärä (Lecture Notes in Computer Science; painos 9041).

### **Tutkimuksen haasteet digitaalisessa ajassa**

Linden, K., 2015, julkaisussa: Kansalliskirjasto. 57, s. 41-42 2 Sivumäärä

2014

### **Analyzing and Improving the Quality of a Historical News Collection using Language Technology and Statistical Machine Learning Methods**

Kettunen, K., Honkela, T., Linden, K., Kauppinen, P., Pääkkönen, T. & Kervinen, J., 16 elokuuta 2014, *IFLA World Library and Information Congress Proceedings: 80th IFLA General Conference and Assembly*. Lyon, France: IFLA, 23 Sivumäärä

### **FIN-CLARIN – en humanistisk forskningsinfrastruktur med betoning på språk**

Linden, B. K. J., heinäkuuta 2014, julkaisussa: Språk i Norden = Sprog i Norden : årsskrift for Nordisk språkråd og språknemdene i Norden. s. 126-132 7 Sivumäärä

### **Part-of-Speech Tagging using Conditional Random Fields: Exploiting Sub-Label Dependencies for Improved Accuracy**

Silfverberg, M., Ruokolainen, T., Linden, K. & Kurimo, M., 22 kesäkuuta 2014, *Unknown host publication*.

### **CLARA: A New Generation of Researchers in Common Language Resources and Their Applications**

Koenraad De Smedt, Erhard Hinrichs, Detmar Meurers, Inguna Skadina, Bolette Pedersen, Costanza Navarretta, Núria Bel, Krister Linden, Marketa Lopatkova, Jan Hajic, Gisle andersen and Przemyslaw Lenkiewicz, 26 toukokuuta 2014, *Proceedings of LREC 2014*. Reykjavik, Iceland: European Language Resources Association (ELRA), 9 Sivumäärä #410

### **HFST-SweNER – A New NER Resource for Swedish**

Kokkinakis, D., Niemi, J., Hardwick, S., Linden, K. & Borin, L., 26 toukokuuta 2014, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Calzolari, N., Choukri, K., Declerck, T., Loftsson, H., Maegaard, B., Mariani, J., Moreno, A., Odijk, J. & Piperidis, S. (toim.). Reykjavik, Iceland: European Language Resources Association (ELRA), 7 Sivumäärä #391

### **Heuristic Hyper-minimization of Finite State Lexicons**

Drobac, S., Linden, K., Pirinen, T. & Silfverberg, M., 26 toukokuuta 2014, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Reykjavik, Iceland: European Language Resources Association (ELRA), Vuosikerta 9. 6 Sivumäärä #784

### **The Strategic Impact of META-NET on the Regional, National and International Level**

Georg Rehm, Hans Uszkoreit, Sophia Ananiadou, Núria Bel, Audrone Bielevičienė, Lars Borin, António Branco, Gerhard Budin, Nicoletta Calzolari, Walter Daelemans, Radovan Garabik, Marko Grobelnik, Carmen Garcia-Mateo, Josef Van Genabith, Jan Hajic, Inma Hernaez, John Judge, Svetla Koeva, Simon Krek, Cvetana Krstev, Krister Lindén, Bernardo Magnini, Joseph Mariani, John McNaught, Maite Melero, Monica Monachini, Asuncion Moreno, Jan Odijk, Maciej Ogrodniczuk, Piotr Pezik, Stelios Piperidis, Adam Przepiórkowski, Eiríkur Rögnvaldsson, Michael Rosner, Bolette Sandford Pedersen, Inguna Skadina, Koenraad De Smedt, Marko Tadić, Paul Thompson, Dan Tufi, Tamás Váradi, Andrejs Vasiljevs, Kadri Vider, Jolanta Zabarskaite, 26 toukokuuta 2014, *Proceedings of LREC 2014*. Reykjavik, Iceland: European Language Resources Association (ELRA), s. 1517-1524 8 Sivumäärä #405

### **Accelerated Estimation of Conditional Random Fields using a Pseudo-Likelihood-inspired Perceptron Variant**

Ruokolainen, T., Silfverberg, M., Kurimo, M. & Linden, K., 26 huhtikuuta 2014, *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics: EACL 2014*. ACL, 5 Sivumäärä

### **State-of-the-Art in Weighted Finite-State Spell-Checking**

Pirinen, T. & Linden, K., 6 huhtikuuta 2014, *Computational Linguistics and Intelligent Text Processing: 15th International Conference, CICLing 2014, Kathmandu, Nepal, April 6-12, 2014, Proceedings, Part II*. Gelbukh, A. (toim.). Berlin Heidelberg: Springer-Verlag, Vuosikerta 2. s. 519-532 14 Sivumäärä (Lecture Notes in Computer Science; painos 8404).

2013



### **Software Newsroom – an approach to automation of news search and editing**

Huovelin, J., Gross, O., Solin, O., Linden, K., Maisala, S. P. T., Oittinen, T., Toivonen, H., Niemi, J. & Silfverberg, M., 7 marraskuuta 2013, julkaisussa: *Journal of Print Media Technology research*. 2, 3, s. 141-156 15 Sivumäärä

### **HFST—a System for Creating NLP Tools**

Linden, K., Axelson, E., Drobac, S., Hardwick, S., Kuokkala, J., Niemi, J., Pirinen, T. & Silfverberg, M., syyskuuta 2013, *Systems and Frameworks for Computational Morphology: Communications in Computer and Information Science*. Mahlow, C. & Piotrowski, M. (toim.). Springer-Verlag, s. 53-71 20 Sivumäärä (Communications in Computer and Information Science).

### **Baltic and Nordic Parts of the European Linguistic Infrastructure**

Skadina, I., Vasiljevs, A., Borin, L., Linden, K., Losnegaard, G., Olsen, S., Pedersen, B., Rozis, R. & De Smedt, K., 20 toukokuuta 2013, *Proceedings of NODALIDA 2013*. Linköping University Electronic Press, 16 Sivumäärä

### **Nordic and Baltic wordnets aligned and compared through “WordTies”**

Pedersen, B., Linden, K., Vider, K., Forsberg, M., Kahusk, N., Niemi, J., Nygaard, L., Seaton, M., Orav, H., Borin, L., Voionmaa, K., Nisbeth, N. & Rognvaldsson, E., 20 toukokuuta 2013, *Proceedings of NODALIDA 2013*. Linköping University Press, 16 Sivumäärä (Linköping Electronic Conference Proceedings; nro 85).

### **Is It Possible to Create a Very Large WordNet in 100 days? – an Evaluation**

Linden, K. & Niemi, J., 2013, julkaisussa: *Language Resources and Evaluation*. 48, 2, s. 191-201 10 Sivumäärä

## **2012**

### **Building shared language research environments inside the European Union: How to optimize the system based on experiences from real life**

Oksanen, V. & Linden, K., 12 syyskuuta 2012, *Unknown host publication*. 12 Sivumäärä

### **Extending and Updating the Finnish Wordnet**

Linden, K., Niemi, J. & Hyvärinen, M., 14 elokuuta 2012, *Shall We Play the Festschrift Game? Essays on the Occasion of Lauri Carlson's 60th Birthday*. Springer-Verlag, s. 67-98 32 Sivumäärä

### **Shall We Play the Festschrift Game? Essays on the Occasion of Lauri Carlson's 60th Birthday**

Sousa, D., Linden, K. & Nganga, W., 14 elokuuta 2012, Springer-Verlag. 241 Sivumäärä

### **Creation of Shared Language Resource Repository in the Nordic and Baltic Countries**

Vasiljevs, A., Forsberg, M., Goronstay, T., Hansen, D., Jóhannsdóttir, K., Linden, K., Lyse, G., Offersgaard, L., Olsen, S., Oksanen, V., Pedersen, B., Rognvaldsson, E., Skadina, I. & de Smedt, K., 23 toukokuuta 2012, *The eighth international conference on Language Resources and Evaluation (LREC): LREC 2012*. Istanbul, Vuosikerta 8. 8 Sivumäärä

### **Representing the Translation Relation in a Bilingual Wordnet**

Niemi, J. & Linden, K., 23 toukokuuta 2012, *The eighth international conference on Language Resources and Evaluation (LREC): LREC 2012*. Istanbul: European Language Resources Association (ELRA), Vuosikerta 8. 8 Sivumäärä

### **Specifying Treebanks, Outsourcing Parsebanks: FinnTreeBank 3**

Voutilainen, A., Muhonen, K., Purtonen, T. K. & Linden, K., 23 toukokuuta 2012, *Proceedings of LREC 2012: 8th ELRA Conference on Language Resources and Evaluation*. Calzolari, N., Choukri, K., Declerck, T., Dogan, M. U., Maegaard, B., Mariani, J., Odijk, J. & Piperidis, S. (toim.). Paris: European Language Resources Association (ELRA), s. 1927-1931 5 Sivumäärä

### **Improving Finite-State Spell-Checker Suggestions with Part of Speech N-Grams**

Pirinen, T., Silfverberg, M. & Linden, K., 9 maaliskuuta 2012, *Computational Linguistics and Intelligent Text Processing: 13th International Conference, CICLing 2012*. Gelbukh, A. (toim.). Delhi, India, 11 Sivumäärä

### **Predictive Text Entry for Agglutinative Languages Using Unsupervised Morphological Segmentation**

Silfverberg, M., Linden, K. & Hyvärinen, M., 9 maaliskuuta 2012, julkaisussa: *Lecture Notes in Computer Science*. 12 Sivumäärä

### **FinnTreeBank 2**

Voutilainen, A., Purtonen, T. K., Muhonen, K., Hardwick, S. & Linden, K., helmikuuta 2012

### **Finding a Location for a New Word in WordNet**

Pääkkö, P. & Linden, K., 9 tammikuuta 2012, *Proceedings of the Global Wordnet Conference*. Matsue, Japan, 8 Sivumäärä

### **Linking and Validating Nordic and Baltic Wordnets - A Multilingual Action in META-NORD**

Sanford Pedersen, B., Borin, L., Forsberg, M., Linden, K., Orav, H. & Rognvaldsson, E., 9 tammikuuta 2012, *Proceedings of the Global Wordnet Conference*. Matsue, Japan, 5 Sivumäärä

### **Using a Bilingual Resource to Add Synonyms to a Wordnet: FinnWordNet and Wikipedia as an Example**

Niemi, J., Linden, K. & Hyvärinen, M., 9 tammikuuta 2012, *Proceedings of the Global Wordnet Conference*. Matsue, Japan, s. 227–231 5 Sivumäärä

### **HFST - an Environment for Creating Language Technology Applications**

Linden, K., Silfverberg, M., Pirinen, T., Hardwick, S., Drobac, S. & Axelson, E., 2012, *COMPUTATIONAL LINGUISTICS – APPLICATIONS*. Piasecki, M. & Przepiórkowski, A. (toim.). Springer-Verlag, 20 Sivumäärä (Studies in Computational Intelligence).

### **Suomen kieli digitaalisella aikakaudella**

Koskeniemi, K., Linden, K., Carlson, L., Vainio, M., Arppe, A., Lennes, M., Westerlund, H., Hyvärinen, M., Bartis, I., Nuolijärvi, P. & Piehl, A., 2012, Berlin: Springer-Verlag. 85 Sivumäärä

## **2011**

### **META-NORD: Towards Sharing of Language Resources in Nordic and Baltic Countries**

Skadina, I., Vasiljevs, A., Borin, L., De Smedt, K., Linden, K. & Rognvaldsson, E., 12 marraskuuta 2011, *The 5th International Joint Conference on Natural Language Processing*. Chiang Mai, Thailand, 9 Sivumäärä

### **HFST—Framework for Compiling and Applying Morphologies**

Linden, K., Silfverberg, M., Axelson, E., Hardwick, S. & Pirinen, T., 26 elokuuta 2011, *Systems and Frameworks for Computational Morphology*. Mahlow, C. & Pietrowski, M. (toim.). Springer, Vuosikerta Vol. 100. s. 67-85 18 Sivumäärä (Communications in Computer and Information Science; painos 100).

### **Do wordnets also improve human performance on NLP tasks?**

Muhonen, K. & Linden, K., 12 toukokuuta 2011, *NEALT Proceedings Series Vol. 11*. Riga, Latvia: Northern European Association for Language Technology, Vuosikerta 11. 7 Sivumäärä (NEALT Proceedings Series).

### **FIN-CLARIN: A Framework for Depositing and Disseminating Language Resources for R&D**

Voutilainen, A. & Linden, K., 11 toukokuuta 2011, *NEALT Proceedings Series*. Riga, Latvia, 4 Sivumäärä (NEALT Proceedings Series).

### **Open Content Licenses: How to choose the right one**

Oksanen, V. & Linden, K., 11 toukokuuta 2011, *NEALT Proceedings Series*. Riga, Latvia, Vuosikerta 13. 7 Sivumäärä (NEALT Proceedings Series).

### **Combining Statistical Models for POS Tagging using Finite-State Calculus**

Silfverberg, M. & Linden, K., 9 toukokuuta 2011, *Proceedings of the 18th Conference of Computational Linguistics NODALIDA 2011*. Pedersen, B. & Skadiņa, I. (toim.). Northern European Association for Language Technology, s.

183–190 8 Sivumäärä (NEALT Proceedings Series; painos 11).

#### **Designing a Dependency Representation and Grammar Definition Corpus for Finnish**

Voutilainen, A., Linden, K. & Purtonen, T. K., 8 toukokuuta 2011, *Las tecnologías de la información y las comunicaciones: Presente y futuro en el análisis de corpora: Actas del III Congreso Internacional de Lingüística de Corpus*. Pastor, C. & Mora, C. (toim.). Valencia, s. 151-158 9 Sivumäärä

#### **HFST-Framework for Compiling and Applying Morphologies**

Linden, K., Axelson, E., Hardwick, S., Pirinen, T. A. & Silfverberg, M., 2011, *SYSTEMS AND FRAMEWORKS FOR COMPUTATIONAL MORPHOLOGY*. Mahlow, C. & Piotrowski, M. (toim.). Springer-Verlag, s. 67-85 19 Sivumäärä (Communications in Computer and Information Science; painos 100).

### **2010**

#### **FinnTreeBank 1**

Voutilainen, A., Purtonen, T. K., Leisko-Järvinen, S., Kumlander, M., Linden, K., Nissinen, M. & Hardwick, S., 29 joulukuuta 2010

#### **FinnWordNet - a Finnish WordNet database**

Linden, K., Carlson, L. & Niemi, J. A., joulukuuta 2010

#### **Creating and Weighting Hunspell Dictionaries as Finite-State Automata**

Pirinen, T. & Linden, K., marraskuuta 2010, julkaisussa: *Investigationes Linguisticae* (Online Edition). 19

#### **FinnWordNet - WordNet på finska via översättning**

Linden, K. & Carlson, L., marraskuuta 2010, julkaisussa: *LexicoNordica*. 17, 16 Sivumäärä

#### **Building and Using Existing Hunspell Dictionaries and TEX Hyphenators as Finite-State Automata**

Pirinen, T. & Linden, K., lokakuuta 2010, *Proceedings of International Multiconference on Computer Science and Information Technology: Computational Linguistics—Applications (CLA'10)*. Ganzha, M. & Paprzycki, M. (toim.). Wisla, Poland, Vuosikerta 5. s. 477–484 8 Sivumäärä (Proceedings of the International Multiconference on Computer Science and Information Technology).

#### **Part-of-Speech Tagging using Parallel Weighted Finite-State Transducers**

Silfverberg, M. & Linden, K., elokuuta 2010, *Proceedings of IceTAL 2010: 7th International Conference on Natural Language Processing*. Reykjavik, 12 Sivumäärä

#### **Finite-State Spell-Checking with Weighted Language and Error Models: Building and Evaluating Spell-Checkers with Wikipedia as Corpus**

Pirinen, T. & Linden, K., toukokuuta 2010, *Proceedings of LREC 2010: Workshop on Creation and use of basic lexical resources for less-resourced languages*.

#### **Laundry Symbols and License Management: Practical Considerations for the Distribution of LRs based on experiences from CLARIN**

Oksanen, V., Linden, K. & Westerlund, H., toukokuuta 2010, *Proceedings of LREC 2010 : Workshop on Language Resources: From Storyboard to Sustainability and LR Lifecycle Management*.

### **2009**

#### **HFST Library**

Silfverberg, M., Linden, K., Pirinen, T. & Axelson, E., kesäkuuta 2009

#### **HFST-optimized-lookup**

Silfverberg, M., Hardwick, S. & Linden, K., kesäkuuta 2009

### **Conflict resolution using weighted rules in HFST-TWOLC**

Silfverberg, M. & Linden, K., toukokuuta 2009, julkaisussa: NEALT Proceedings Series. 4, s. 174-181 8 Sivumäärä

### **Guessers for Finite-State Transducer Lexicons**

Lindén, K., maaliskuuta 2009, *Computational Linguistics and Intelligent Text Processing: 10th International Conference, CICLing 2009*. Gelbukh, A. (toim.). Springer-Verlag, s. 158-169 12 Sivumäärä (Lecture Notes in Computer Science).

### **Corpus-based lexeme ranking for morphological guessers**

Linden, K. & Tuovila, J., 2009, *State of the Art in Computational Morphology edited by Cerstin Mahlow, Michael Piotrowski*. Berlin, Heidelberg: Springer-Verlag, s. 118-135 18 Sivumäärä (Communications in computer and information science).

### **Corpus-based paradigm Selection for morphological entries**

Linden, K. & Tuovila, J., 2009, julkaisussa: NEALT Proceedings Series. 4, s. 96-102 7 Sivumäärä

### **Entry Generation for New Words by Analogy for Morphological Lexicons**

Linden, K., 2009, julkaisussa: Northern European Journal of Language Technology. 1, s. 1-25 25 Sivumäärä

### **HFST runtime format: A compacted transducer format allowing for fast lookup**

Silfverberg, M. & Linden, K., 2009, *Finite-State Methods and Natural Language Processing - FSMNLP 2009 Eight International Workshop*.

### **HFST tool for morphology: An efficient open-source package for construction of morphological analyzers**

Linden, K., Silfverberg, M. & Pirinen, T., 2009, *State of the Art in Computational Morphology edited by Cerstin Mahlow, Michael Piotrowski*. Berlin, Heidelberg: Springer-Verlag, s. 28-47 20 Sivumäärä (Communications in computer and information science).

### **Weighted finite-state morphological analysis of finnish compounding with HFST-LEXC**

Linden, K. & Pirinen, T., 2009, julkaisussa: NEALT Proceedings Series. 4, s. 89-95 7 Sivumäärä

### **Weighting finite-state morphological analyzers using HFST tools**

Linden, K. & Pirinen, T., 2009, *Finite-State Methods and Natural Language Processing - FSMNLP 2009 Eight International Workshop*.

## **2008**

### **A probabilistic model for guessing base forms of new words by analogy**

Linden, K., 2008, julkaisussa: Lecture Notes in Computer Science. 4919, s. 106-116 11 Sivumäärä

### **Assigning an Inflectional Paradigm using the Longest Matching Affix**

Linden, K., 2008, *Ei mitään ongelmia: Juhlakirja Juhani Reimanille 50-vuotispäiväksi 23.1.2008*. Wiberg, M. & Koura, A. (toim.). Turku

## **2007**

### **Expert panel report: the Nordic countries : a leading region in language technology**

Koskeniemi, K. (toim.), Linden, K. (toim.) & Nordgård, T. (toim.), 2007, Helsinki: University of Helsinki. Department of General Linguistics. 76 Sivumäärä (Publications; painos 40)

### **Opiskelijapalautteen keräämisestä**

Kauhanen, I. & Linden, K., 2007, Helsinki: Helsingin yliopisto, yleisen kielitieteen laitos. 18 Sivumäärä (Opintomonisteet / Yleisen kielitieteen laitos. Helsingin yliopisto; nro No. 4)

**SpråkVis - Språkteknologisk vismansrapport: En utgåvad sammanfattning av The Nordic Countries - A leading Region in language Technology**

Koskenniemi, K. (toim.), Linden, K. (toim.) & Nordgård, T. (toim.), 2007, Helsinki: Helsingfors universitet. Institutionen för allmän språkvetenskap. 12 Sivumäärä

**2006**

**Common Infrastructure for Finite-State Based Methods and Linguistics Descriptions**

Yli-Jyrä, A. M., Koskenniemi, K. & Linden, K., 2006, *International Workshop Towards a Research Infrastructure for Language Resources*. 2 Sivumäärä

**Multilingual modelling of cross-lingual spelling variants**

Linden, K., 2006, julkaisussa: *Information Retrieval Journal*. 9, 3, s. 295-310 16 Sivumäärä

**2005**

**Inquiries into words, constraints and contexts: Festschrift in the honour of Kimmo Koskenniemi on his 60th birthday**

Arppe, A., Carlson, L., Linden, K., Piitulainen, J. O., Suominen, M., Vainio, M., Westerlund, H. & Yli-Jyrä, A. M., 2005, [S.I.]: CSLI publications. 288 Sivumäärä (CSLI Studies in Computational Linguistics ONLINE)

**Morfessor and Hutmegs: Unsupervised Morpheme Segmentation for Highly-Inflecting and Compounding Languages**

Creutz, M., Lagus, K., Linden, K. & Virpioja, S. P., 2005, *The Second Baltic Conference on Human Language Technologies: Proceedings*. Langemets, M. & Penjam, P. (toim.). Tallinn: Institute of Cybernetics; Institute of the Estonian Language

**Word sense discovery and disambiguation**

Linden, K., 2005, Helsinki: University of Helsinki.

**Word senses**

Linden, K., 2005, *Inquiries into words, constraints and contexts: Festschrift in the honour of Kimmo Koskenniemi on his 60th birthday*. [S.I.]: [s.n.], s. 249-258 10 Sivumäärä

**2004**

**Evaluation of Linguistic Features for Word Sense Disambiguation with Self-Organized Document Maps**

Linden, K., marraskuuta 2004, julkaisussa: *Language Resources and Evaluation*. 38, 4, s. 417-435 19 Sivumäärä

**Discovering Synonyms and Other Related Words**

Linden, K. & Piitulainen, J. O., elokuuta 2004, *Proceedings of COLING 2004: CompuTerm 2004: 3rd International Workshop on Computational Terminology*. Ananadiou, S. & Zweigenbaum, P. (toim.). s. 63-70

**Morpheme Segmentation Gold Standards for Finnish and English**

Creutz, M. J. P. & Linden, B. K. J., 2004, *Publications in Computer and Information Science: Report A77*. Helsinki University of Technology, 33 Sivumäärä

**2003**

**Word Sense Disambiguation with THESSOM**

Linden, K., syyskuuta 2003, *Proceedings of WSOM'03 - Intelligent Systems and Innovational Computing*.

2002

### Word Sense Disambiguation in Document Space

Linden, K. & Lagus, K., lokakuuta 2002, *Proceedings of 2002 IEEE International Conference on Systems, Man and Cybernetics*.

1997

### Language Applications with Finite-State Technology

Linden, K., 1997, julkaisussa: *International Journal of Corpus Linguistics*. s. 281-288

1987

### Unification as a grammatical tool

Carlson, L. & Linden, K., 1987, julkaisussa: *Nordic journal of linguistics*. 10, s. 111-136

## Projektit

- 1. FIN-CLARIN**  
Linden, K., Piitulainen, J., Niemi, J., Lennes, M., Bartis, I., Westerlund, H., Drobac, S., Axelson, E. & Kauppinen, P.  
05/06/2006 → ...
- 2. FinnWordNet - A Finnish WordNet**  
Linden, K., Niemi, J., Hyvärinen, M., Muhonen, K. & Pääkkö, P.  
01/01/2010 → ...
- 3. HFST - Helsinki Finite-State Technology**  
Linden, K., Koskenniemi, K., Yli-Jyrä, A., Huldén, M., Silfverberg, M., Pirinen, T., Axelson, E., Hardwick, S., Niemi, J. & Hurskainen, A.  
01/01/2005 → ...
- 4. Lahjoita puhetta -kampanjan suunnittelu ja toteutus**  
Jauhiainen, T., Lennes, M. & Linden, K.  
01/11/2019 → ...
- 5. ANEE: ANEE tiimi 1: Digitaaliset ihmistieteet**  
Svärd, S., Linden, K., Bach, J., Jauhiainen, T., Jauhiainen, H., Sahala, A., Konstantopoulos, G., Fink, S., Zaia, S., Luukko, M. I., Bennett, E. R., Alstola, T., Vanderstraeten, E., Autere, K., Giessler, J., Debourse, C. H. I. & Tambs, L.  
01/01/2018 → 31/12/2025
- 6. ANEE: Centre of Excellence in Ancient Near Eastern Empires**  
Silverman, J. M., Svärd, S., Lahelma, A., Nikki, N., Töyräänvuori, J., Jokiranta, J., Laine, L., Kletter, R., Bonnie, R., Thomas, S., Holmqvist-Sipilä, E., Kouki, P., Wasmuth, M., Wallis, C., Lorenzon, M., Jauhiainen, H., Jauhiainen, T., Alstola, T., Bach, J., Fink, S., Konstantopoulos, G., Zaia, S., Sahala, A., Miettunen, P. H., Avneri Meir, R., Aissaoui, A. I., Arpe, O. P., Luukko, M. I., Kovanen, H., Linden, K., Bennett, E. R., Saari, S., Vanderstraeten, E., Autere, K., Debourse, C. H. I., Kailamäki, U., Kautonen, S. K., Spunangle, A. & Tambs, L.  
01/01/2018 → 31/12/2025
- 7. LIST: Language Identification of Speech and Text**  
Jauhiainen, T. & Linden, K.  
01/01/2021 → 31/12/2022
- 8. Semantic Domains in Akkadian Texts**  
Linden, K., Alstola, T., Jauhiainen, H., Jauhiainen, T., Sahala, A., Svärd, S. & Zaia, S.  
01/09/2016 → 31/08/2020
- 9. FIN-CLARIN UPGRADE - CLARINin vahvistaminen Suomessa FIN-CLARIN UPGRADE - Förstärkning av CLARIN i Finland**  
Linden, K., Lennes, M., Westerlund, H., Piitulainen, J., Niemi, J., Axelson, E., Jauhiainen, T., Drobac, S., Hardwick, S., Kauppinen, P., Voutilainen, A., Dieckmann, U., Palolahti, M., Matthiesen, M. & Aalto, T.  
01/01/2017 → 31/12/2019
- 10. Suomalais-ugrilaiset kielet ja internet**  
Linden, K., Jauhiainen, H. & Jauhiainen, T.  
01/01/2013 → 31/12/2018

11. **FIN-CLARIN-INVEST2 - Yhteiskäyttöisten kielivarojen ja -teknologian kohentaminen CLARINia varten Suomessa**  
**FIN-CLARIN-INVEST2 - Uppgradering av de gemensamma språkresurserna och -teknologin för CLARIN i Finland**  
 Linden, K.  
 01/01/2015 → 28/02/2017
12. **FINNCLARIN CONTENT**  
 Linden, K. & Koskenniemi, K.  
 01/01/2011 → 31/12/2015
13. **FIN-CLARIN-INVEST - Yhteiskäyttöisten kielivarojen ja -teknologian kohentaminen CLARINia varten Suomessa**  
**FIN-CLARIN-INVEST - Uppgradering av de gemensamma språkresurserna och -teknologin för CLARIN i Finland**  
 Linden, K.  
 01/01/2014 → 31/12/2014
14. **CLARA - Common Language Resources and their Applications**  
 Koskenniemi, K. & Linden, K.  
 Unknown funder  
 01/12/2009 → 31/03/2014
15. **Uudelleenpuhuminen - tulkkaus tekstitystä varten**  
 Tiittula, L., Linden, K., Bacon, H., Carlson, L., Eskelinen, J., Hartama-Heinonen, R., Koponen, M., Kukkonen, P.,  
 Mikkonen, K. & Pekkilä, E. O.  
 Humanistinen tdk  
 01/02/2013 → 31/12/2013
16. **FIN-CLARIN-BUILD - CLARINin vahvistaminen Suomessa****FIN-CLARIN-BUILD - etablering av CLARIN i Finland**  
 Linden, K.  
 14/12/2012 → 31/12/2013
17. **META-NORD: Baltic and Nordic Parts of the European Open Linguistic Infrastructure**  
 Linden, K., Koskenniemi, K. & Bartis, I.  
 EU FP7  
 01/02/2011 → 31/01/2013
18. **FIN-CLARIN: Finnish Language Resource Infrastructure**  
 Koskenniemi, K. & Linden, K.  
 01/01/2007 → 31/12/2012
19. **Common Language Resources and Technology Infrastructure (CLARIN)**  
 Koskenniemi, K., Linden, K., Oksanen, V. & Yangarber, R.  
 01/01/2008 → 30/06/2011

## Aktiviteetit

### Post-doctoral mentor / Måns Huldén

Krister Linden (Post doc mentor)

12 tammikuuta 2009 → 11 joulukuuta 2011

Aktiviteetti: Tutkimustyytit > Post doctoral mentor

### Examination of doctoral dissertation / Grégoire Détrez

Krister Linden (Vastaväittäjä) & Bo Krister Johan Linden (Esitarkastaja)

2016

Aktiviteetti: Tutkimustyytit > Vastaväittäjä

### Exmination of doctoral dissertation / Sami Virpioja

Krister Linden (Vastaväittäjä) & Bo Krister Johan Linden (Esitarkastaja)

2012

Aktiviteetti: Tutkimustyytit > Vastaväittäjä

### Examination of doctoral dissertation

Krister Linden (Esitarkastaja)

15 toukokuuta 2007 → 31 toukokuuta 2007

Aktiviteetti: Tutkimustyytit > Väitöskirjan esitarkastaja

### Examination of doctoral dissertation

Krister Linden (Esitarkastaja)

1 syyskuuta 2007 → 15 syyskuuta 2007

Aktiviteetti: Tutkimustyytit > Väitöskirjan esitarkastaja

**Supervisor of PhD Thesis in Language Technology**

Bo Krister Johan Linden (Ohjaaja) & Anssi Yli-Jyrä (Ohjaaja)

1 tammikuuta 2010 → 22 lokakuuta 2016

Aktiviteetti: Tutkimustyytit › Väitöskirjan ohjaaja tai sivuohjaaja

**Supervisor of doctoral thesis in Language Technology / Emily Öhman**

Krister Linden (Sivuohjaaja)

1 elokuuta 2016 → ...

Aktiviteetti: Tutkimustyytit › Väitöskirjan ohjaaja tai sivuohjaaja

**Supervisor of doctoral thesis in Language Technology / Senka Drobac**

Krister Linden (Ohjaaja)

1 tammikuuta 2012 → ...

Aktiviteetti: Tutkimustyytit › Väitöskirjan ohjaaja tai sivuohjaaja

**Supervisor of doctoral thesis in Language Technology / Tommi Jauhiainen**

Krister Linden (Ohjaaja)

1 kesäkuuta 2011 → ...

Aktiviteetti: Tutkimustyytit › Väitöskirjan ohjaaja tai sivuohjaaja

**Supervisor of doctoral thesis in Language Technology / Tommi Pirinen**

Krister Linden (Ohjaaja)

1 tammikuuta 2010 → 2 joulukuuta 2014

Aktiviteetti: Tutkimustyytit › Väitöskirjan ohjaaja tai sivuohjaaja

**Cabinet of Commissioner Leonard Orban at the European Commission (Ulkoisen yksikkö)**

Krister Linden (Asiantuntijajäsen)

1 huhtikuuta 2008 → 31 maaliskuuta 2009

Aktiviteetti: Jäsenyytystyytit › Jäsenyys tai muu rooli suomalaisessa tai kansainvälisessä julkisen sektorin organisaatiossa

**Nordiska Ministerrådet (Ulkoisen yksikkö)**

Krister Linden (Asiantuntijajäsen)

1 tammikuuta 2006 → 31 joulukuuta 2006

Aktiviteetti: Jäsenyytystyytit › Jäsenyys tai muu rooli suomalaisessa tai kansainvälisessä julkisen sektorin organisaatiossa

**CLARIN - Common Language Resources and Technology Infrastructure (Ulkoisen yksikkö)**

Krister Linden (Hallituksen jäsen)

1 tammikuuta 2010 → 31 joulukuuta 2011

Aktiviteetti: Jäsenyytystyytit › Jäsenyys tai muu rooli kansallisessa/kansainvälisessä komiteassa, toimikunnassa tai hallituksessa

**ESF - European Science Foundation (Ulkoisen yksikkö)**

Krister Linden (Asiantuntijajäsen)

24 toukokuuta 2011 → 23 toukokuuta 2015

Aktiviteetti: Jäsenyytystyytit › Jäsenyys tai muu rooli kansallisessa/kansainvälisessä komiteassa, toimikunnassa tai hallituksessa

**HAKA - Identity federation of the Finnish universities, polytechnics and research institutions (Ulkoisen yksikkö)**

Krister Linden (Hallituksen jäsen)

2012 → 2013

Aktiviteetti: Jäsenyytystyytit › Jäsenyys tai muu rooli kansallisessa/kansainvälisessä komiteassa, toimikunnassa tai hallituksessa

**KIT Language Technology Graduate School of Finland (Ulkoisen yksikkö)**

Krister Linden (Asiantuntijajäsen)



1 tammikuuta 2001 → 31 joulukuuta 2005

Aktiviteetti: Jäsenyyssyypit › Jäsenyys tai muu rooli kansallisessa/kansainvälisessä komiteassa, toimikunnassa tai hallituksessa

#### **Langnet (Ulkoinen yksikkö)**

Krister Linden (Hallituksen jäsen)

1 tammikuuta 2012 → 31 joulukuuta 2015

Aktiviteetti: Jäsenyyssyypit › Jäsenyys tai muu rooli kansallisessa/kansainvälisessä komiteassa, toimikunnassa tai hallituksessa

#### **Lantern - Language Technology Research Network (Ulkoinen yksikkö)**

Krister Linden (Sihteeri)

1 tammikuuta 2005 → 31 joulukuuta 2007

Aktiviteetti: Jäsenyyssyypit › Jäsenyys tai muu rooli kansallisessa/kansainvälisessä komiteassa, toimikunnassa tai hallituksessa

#### **META-TRUST (Ulkoinen yksikkö)**

Krister Linden (Hallituksen jäsen)

12 syyskuuta 2012 → 31 tammikuuta 2017

Aktiviteetti: Jäsenyyssyypit › Jäsenyys tai muu rooli kansallisessa/kansainvälisessä komiteassa, toimikunnassa tai hallituksessa

#### **NYKYKIELET 2010-2017 (Yksikkö)**

Krister Linden (Hallituksen jäsen)

1 tammikuuta 2014 → 31 joulukuuta 2017

Aktiviteetti: Jäsenyyssyypit › Jäsenyys tai muu rooli kansallisessa/kansainvälisessä komiteassa, toimikunnassa tai hallituksessa

#### **Northern European Association for Language Technology (Ulkoinen yksikkö)**

Krister Linden (Puheenjohtaja)

1 tammikuuta 2010 → 31 joulukuuta 2013

Aktiviteetti: Jäsenyyssyypit › Jäsenyys tai muu rooli kansallisessa/kansainvälisessä komiteassa, toimikunnassa tai hallituksessa

#### **Northern European Association for Language Technology (Ulkoinen yksikkö)**

Krister Linden (Asiantuntijajäsen)

1 tammikuuta 2014 → 31 joulukuuta 2017

Aktiviteetti: Jäsenyyssyypit › Jäsenyys tai muu rooli kansallisessa/kansainvälisessä komiteassa, toimikunnassa tai hallituksessa

#### **Global Wordnet Association (Ulkoinen yksikkö)**

Krister Linden (Jäsen)

2012 → ...

Aktiviteetti: Jäsenyyssyypit › Jäsenyys tai muu rooli tutkimusverkostossa

#### **NEALT - North European Association for Language Technology (Ulkoinen yksikkö)**

Krister Linden (Jäsen)

2009 → ...

Aktiviteetti: Jäsenyyssyypit › Jäsenyys tai muu rooli tutkimusverkostossa

#### **Nykykielten laitos (Yksikkö)**

Krister Linden (Assessment)

2014

Aktiviteetti: Jäsenyyssyypit › Lausunnot virantäytöissä