

Reassessing Language Change: The Challenge of Real Time

Annual Report, January 2016 – December 2017

1. *Conferences, presentations, seminars*

2017

Kesäniemi, Joonas. LCD Aggregated Data Analysis workbench (LADA). ATT-hankkeen esittely Yhteinen tiedonhallinta (YTI) -hankkeelle. Ministry of Finance and Population Register Centre. Helsinki, Finland, December 2017.

Vartiainen, Turo. Emergent word classes: Gradience in some structurally complex modifiers in English. Gradience and Constructional Change. Santiago-Leuven-Edinburgh workshop on change in English, Edinburgh, United Kingdom, November 2017.

Kesäniemi, Joonas, Turo Vartiainen, Tanja Säily, Agata Dominowska, Aatu Liimatta and Terttu Nevalainen. From linked open data to meta-analysis in historical linguistics. Presentation, HELDIG Digital Humanities Summit 2017, Helsinki, Finland, October 2017.

Kesäniemi, Joonas, Turo Vartiainen, Tanja Säily, Agata Dominowska, Aatu Liimatta and Terttu Nevalainen. LCD Aggregated Data Analysis workbench (LADA). Demo, HELDIG Digital Humanities Summit 2017, Helsinki, Finland, October 2017.

Kesäniemi, Joonas. Making new use of old research: The Language Change Database. Poster, HULIB, Helsinki, Finland, June 2017.

Säily, Tanja, Eetu Mäkelä and Jukka Suomela. Social embedding of neologisms in early English correspondence. The Dynamics of Lexical Innovation: Data, Methods, Models (DynLex), Munich, Germany, June 2017.

Vartiainen, Turo. "As well" as an emerging connective: Preliminary evidence from corpus data. FINSSE-8, The Biannual Conference for Scholars of English in Finland, Turku, Finland, June 2017.

Kesäniemi, Joonas, Turo Vartiainen, Tanja Säily, Agata Dominowska, Aatu Liimatta and Terttu Nevalainen. Making new use of old research: The Language Change Database. Poster, 38th Annual Conference of the International Computer Archive of Modern and Medieval English (ICAME 38), Prague, Czech Republic, May 2017.

Säily, Tanja. Variation in the productivity of adjective comparison. Invited talk, University of Zurich, Switzerland, March 2017.

Säily, Tanja, Jukka Suomela and Eetu Mäkelä. From old-English grubbers to cheeky blighters? Variation in the productivity of *-er* in the history of English. The 5th International Symposium on History of English Lexicography and Lexicology (HEL-LEX 5), Zurich, Switzerland, February 2017.

2016

Nevalainen, Terttu, Tanja Säily and Turo Vartiainen. Language Change Database: A new online resource. Poster, Building and Using Language Technology (BAULT 2016), Helsinki, Finland, December 2016.

Säily, Tanja. Sociolinguistinen vaihtelu englannin *-er*-substantiivijohtimen produktiivisuudessa 1700-luvulta nykypäivään (Sociolinguistic variation in the productivity of the English nominal suffix *-er* from the 18th century to the present). Seminar of the Helsinki Society for Historical Lexicography, Turku, Finland, September 2016.

Nevalainen, Terttu, Tanja Säily, Turo Vartiainen, Joonas Kesäniemi and the LCD team. What can you do with 200 studies of language change? International Society for the Linguistics of English (ISLE 4), Poznan, Poland, September 2016.

Säily, Tanja. Variation in morphological productivity in the history of English: The case of *-er*. International Society for the Linguistics of English (ISLE 4), Poznan, Poland, September 2016.

Vartiainen, Turo. Word class coercion and semantic change. International Society for the Linguistics of English (ISLE 4), Poznan, Poland, September 2016.

Nevalainen, Terttu and Tanja Säily. The intensifier *so*: A recent innovation or a case of recycling gendered use? International Conference on English Historical Linguistics (ICEHL), Essen, Germany, August 2016.

Vartiainen, Turo. A constructionist approach to category change: Constraining factors in the adjectivization of participles. International Conference on English Historical Linguistics (ICEHL), Essen, Germany, August 2016.

Kesäniemi, Joonas. Small data, big impact – Hands on approach to data management in co-operation with the research group and library as the linked data hub. 40th European Library Automation Group Conference (ELAG 2016), Copenhagen, Denmark, June 2016.

Säily, Tanja and Jukka Suomela. Exploring sociolinguistic variation in the productivity of *-er* and *-or*: BNCweb, MorphoQuantics, and *types2*. Workshop on Derivational Morphology and Spoken Language, Reading, UK, June 2016.

Säily, Tanja and Turo Vartiainen. Variation in pronoun use in early English letters: Gendered styles of writing? Sociolinguistics Symposium 21, Murcia, Spain, June 2016.

Nevalainen, Terttu, Tanja Säily and Turo Vartiainen. History of English as punctuated equilibria? Round table discussion on the rate of language change, Helsinki, Finland, May 2016.

Säily, Tanja. Kielen vaihtelu ja muutos: Korpusmenetelmien ja visualisoinnin uusia tuulia (Language variation and change: New trends in corpus methods and visualisation). Guest lecture, University of Tampere, Finland, April 2016.

Nevalainen, Terttu, Tanja Säily and Turo Vartiainen. Language Change Database: A new online resource. Poster, HiSoN conference, Helsinki, Finland, March 2016.

2. Publications

Forthcoming

Kesäniemi, Joonas, Turo Vartiainen, Tanja Säily and Terttu Nevalainen. Accepted. Open science for English historical corpus linguistics: Introducing the Language Change Database. *Digital Humanities in the Nordic Countries, 3rd Conference, 7-9 March 2018, Helsinki*.

Kesäniemi, Joonas, Turo Vartiainen, Tanja Säily, Agata Dominowska, Aatu Liimatta and Terttu Nevalainen. In preparation. Exploring meta-analysis for historical corpus linguistics based on linked data.

2017

Hiltunen, Turo, Joe McVeigh and Tanja Säily (eds.). 2017. *Big and Rich Data in English Corpus Linguistics: Methods and Explorations* (Studies in Variation, Contacts and Change in English 19). Helsinki: VARIENG. <http://www.helsinki.fi/varieng/series/volumes/19/>

Lijffijt, Jeffrey and Terttu Nevalainen. 2017. A simple model for recognizing core genres in the BNC. Turo Hiltunen, Joe McVeigh and Tanja Säily (eds.), *Big and Rich Data in English Corpus Linguistics: Methods and Explorations* (Studies in Variation, Contacts and Change in English 19). Helsinki: VARIENG. http://www.helsinki.fi/varieng/series/volumes/19/lijffijt_nevalainen/

Säily, Tanja, Arja Nurmi, Minna Palander-Collin and Anita Auer. 2017. The future of historical sociolinguistics? Tanja Säily, Arja Nurmi, Minna Palander-Collin and Anita Auer (eds.), *Exploring Future Paths for Historical Sociolinguistics* (Advances in Historical Sociolinguistics 7), 1–19. Amsterdam: John Benjamins.

Säily, Tanja, Arja Nurmi, Minna Palander-Collin and Anita Auer (eds.). 2017. *Exploring Future Paths for Historical Sociolinguistics* (Advances in Historical Sociolinguistics 7). Amsterdam: John Benjamins.

Säily, Tanja and Jukka Suomela. 2017. *types2*: Exploring word-frequency differences in corpora. Turo Hiltunen, Joe McVeigh and Tanja Säily (eds.), *Big and Rich Data in English Corpus Linguistics: Methods and Explorations* (Studies in Variation, Contacts and Change in English 19). Helsinki: VARIENG. http://www.helsinki.fi/varieng/series/volumes/19/saily_suomela/

Säily, Tanja, Turo Vartiainen and Harri Siirtola. 2017. Exploring part-of-speech frequencies in a sociohistorical corpus of English. Tanja Säily, Arja Nurmi, Minna Palander-Collin and Anita Auer (eds.), *Exploring Future Paths for Historical Sociolinguistics* (Advances in Historical Sociolinguistics 7), 23–52. Amsterdam: John Benjamins.

Siirtola, Harri, Tanja Säily and Terttu Nevalainen. 2017. Interactive Principal Component Analysis. Ebad Banissi (ed.), *Proceedings of the 21st International Conference on Information Visualisation (IV 2017)*, 416–421. Los Alamitos, CA: IEEE Computer Society.

2016

Lijffijt, Jeffrey, Terttu Nevalainen, Tanja Säily, Panagiotis Papapetrou, Kai Puolamäki and Heikki Mannila. 2016. Significance testing of word frequencies in corpora. *Digital Scholarship in the Humanities* 31(2): 374–397.

Nevalainen, Terttu, Turo Vartiainen, Tanja Säily, Joonas Kesäniemi, Agata Dominowska and Emily Öhman. 2016. Language Change Database: A new online resource. *ICAME Journal* 40: 77–94. doi:10.1515/icame-2016-0006

Säily, Tanja. 2016. Sociolinguistic variation in morphological productivity in eighteenth-century English. *Corpus Linguistics and Linguistic Theory* 12(1): 129–151.

Siirtola, Harri, Poika Isokoski, Tanja Säily and Terttu Nevalainen. 2016. Interactive text visualization with Text Variation Explorer. Ebad Banissi (ed.), *Proceedings of the 20th International Conference on Information Visualisation (IV 2016)*, 330–335. Los Alamitos, CA: IEEE Computer Society.

Vartiainen Turo. 2016a. *Challenges in categorization: Corpus-based studies of adjectival premodifiers in English*. Mémoires de la Société Néophilologique de Helsinki 101. Helsinki: Société Néophilologique.

Vartiainen, Turo. 2016b. A constructionist approach to category change: Constraining factors in the adjectivization of participles. *Journal of English Linguistics* 44(1): 34–60.

Vartiainen, Turo. 2016c. Evaluative nominals in Present-day English: A corpus-based study of the definiteness and syntactic distribution of subjective and objective NPs. *Folia Linguistica* 50(1): 243–268.

3. Events organized

Tanja Säily, Turo Vartiainen. Jeffrey Lijffijt's visit. Project collaborator Jeffrey Lijffijt visited the team in Helsinki to hear about the latest developments concerning the database and to give a demonstration of the tools he has been developing. Helsinki, Finland, May 2017.

Terttu Nevalainen, Tanja Säily and Turo Vartiainen. A round table discussion with Peter Trudgill, Sali Tagliamonte and colleagues on the rate of language change. Helsinki, Finland, May 2016.

4. Research visits

Tanja Säily

University of Zurich, Switzerland: February–March 2017

University of Tampere, Finland: September–December 2016

Turo Vartiainen

Stockholm University, Sweden: November 27 – December 1 2017

Stockholm University, Sweden: May 8–12 2017

5. *Teaching*

Digital Humanities in English Language Studies (ENG331, 405280)

14.01.2017–29.04.2017, 5 credits

Terttu Nevalainen, Turo Vartiainen, Agata Dominowska

The course took place in the 3rd and 4th periods of the academic year 2016–2017. The LCD database was a recurring topic during the course, with a presentation and coursework related to its content.

A course focusing on the project itself is planned for spring 2018.

6. *Database: progress and related applications*

In 2016 and 2017, the database was further planned and expanded by gathering and processing data from research articles. The most recently added articles are organized around the central research themes such as intensifiers, relativizers, negation, and progressive in order to facilitate further comparative research focused on meta-analysis.

At the same time, the online LCD platform was further constructed and developed. The data model was developed incrementally, with corrections and modifications introduced in cycles. The back-end of the database was developed and the structure of entries modified to include publication details and content details as one unit.

A new type of application related to data filtering and composition of new datasets was created: the LCD Aggregation and Data-Analysis workbench (LADA), http://www.helsinki.fi/lcd/pub/LADA_HELDIG2017.pdf.

By the end of 2017 the database will consist of over 300 complete entries and the accompanying data fully transferred to the open-access online database. All of the data is stored and as Excel tables, with the goal of being available for downloading. Part of the data has also been annotated to be readable by LADA.

7. *Website and testing*

The official website of the Language Change Database was updated: <http://www.helsinki.fi/lcd/>

Pilot testing of the database was conducted in the autumn of 2017 on a select group of users, who gave feedback on the functionalities of the back-end of the online database and on the user manual.