



LREC 2022 Workshop
Language Resources and Evaluation Conference
20-25 June 2022

**Proceedings of Globalex Workshop on Linked Lexicography
(GWLL)**

PROCEEDINGS

Editors: Ilan Kernerman, Simon Krek

globaLex

Proceedings of the LREC 2022 workshop Globalex Workshop on Linked Lexicography (GWLL 2022)

Edited by: Ilan Kernerman and Simon Krek

ISBN: 979-10-95546-92-4

EAN: 9791095546924

For more information:

European Language Resources Association (ELRA)

9 rue des Cordelières

75013, Paris

France

<http://www.elra.info>

Email: lrec@elda.org



© European Language Resources Association (ELRA)

These workshop proceedings are licensed under a Creative Commons
Attribution-NonCommercial 4.0 International License

Preface to the Proceedings of Globalex Workshop on Linked Lexicography

Globalex Workshop on Linked Lexicography (GWLL 2022) is the fourth iteration of the GLOBALEX full-day workshop series held in conjunction with LREC. It pursues and expands the topic of linking data across lexicographic resources and with other lexical resources – in the aim of enhancing language data methodologies and applications – which was the focus of our third workshop at LREC 2020 that was cancelled due to the COVID-19 pandemic. Earlier on, the second workshop at LREC 2018 had the main theme of Lexicography and WordNets, and it followed on the first workshop at LREC 2016, on Lexicographic Resources and Human Language Technology.

These proceedings feature 11 papers (including one extended abstract) highlighting lexicographic issues related to linguistic linked data, wordnets, dictionary generation, sign language, collocations, senses, and word formation and morphology, as well as two papers stemming from the fifth edition of TIAD shared task on Translation Inference Across Dictionaries.

We would like to thank all the authors, our colleagues in TIAD and the Review Committee, namely Thierry Declerck, Jorge Gracia, Besim Kabashi, Iztok Kosem, Nikola Ljubešić, and John McCrae, as well as Teja Goli, ELRA, ELEXIS and GLOBALEX, for their contribution to GWLL @ LREC 2022.

Ilan Kernerman and Simon Krek

Organizers

Ilan Kernerman – K Dictionaries – Lexicala
Simon Krek – Institut Jožef Stefan

TIAD 2022 Organizers:

Jorge Gracia – University of Zaragoza, Aragon Institute of Engineering Research (I3A)
Besim Kabashi – University of Erlangen-Nuremberg, Corpus and Computational Linguistics
Ilan Kernerman – K Dictionaries – Lexicala

Review Committee:

Thierry Declerck, DFKI (GERMANY)
Jorge Gracia, University of Zaragoza (SPAIN)
Besim Kabashi, University of Erlangen-Nuremberg (GERMANY)
Ilan Kernerman, K Dictionaries – Lexicala (ISRAEL)
Iztok Kosem, Institut Jožef Stefan (SLOVENIA)
Simon Krek, Institut Jožef Stefan (SLOVENIA)
Nikola Ljubešić, Institut Jožef Stefan (SLOVENIA)
John McCrae, National University of Ireland, Galway (IRELAND)

Table of Contents

<i>Towards the Profiling of Linked Lexicographic Resources</i> Lenka Bajcetic, Seung-bin Yim and Thierry Declerck	1
<i>Towards the Linking of a Sign Language Ontology with Lexical Data</i> Thierry Declerck	6
<i>Modelling Collocations in OntoLex-FrAC</i> Christian Chiarcos, Katerina Gkirtzou, Maxim Ionov, Besim Kabashi, Fahad Khan and Ciprian-Octavian Truică	10
<i>TIAD 2022: The Fifth Translation Inference Across Dictionaries Shared Task</i> Jorge Gracia, Besim Kabashi and Ilan Kernerman	19
<i>Creating Bilingual Dictionaries from Existing Ones by Means of Pivot-Oriented Translation Inference and Logistic Regression</i> Yves Bestgen	26
<i>Compiling a Highly Accurate Bilingual Lexicon by Combining Different Approaches</i> Steinþór Steingrímsson, Luke O’Brien, Finnur Ingimundarson, Hrafn Loftsson and Andy Way .	32
<i>A Category Theory Framework for Sense Systems</i> David Strohmaier and Gladys Tyen	42
<i>Converting a Database of Complex German Word Formation for Linked Data</i> Petra Steiner	52
<i>Resolving Inflectional Ambiguity of Macedonian Adjectives</i> Katerina Zdravkova	60
<i>MorphoLex Turkish: A Morphological Lexicon for Turkish</i> Bilge Arican, Aslı Kuzgun, Büşra Marşan, Deniz Baran Aslan, Ezgi Saniyar, Neslihan Cesur, Neslihan Kara, Oguzhan Kuyrukcu, Merve Ozcelik, Arife Betul Yenice, Merve Dogan, Ceren Oksal, Gökhan Ercan and Olcay Taner Yıldız	68
<i>Time Travel in Turkish: WordNets for Modern Turkish</i> Ceren Oksal, Hikmet N. Oguz, Mert Catal, Nurkay Erbay, Ozgecan Yuzer, Ipek B. Unsal, Oguzhan Kuyrukcu, Arife B. Yenice, Aslı Kuzgun, Büşra Marşan, Ezgi Saniyar, Bilge Arican, Merve Dogan, Özge Bakay and Olcay Taner Yıldız	75
<i>WordNet and Wikipedia Connection in Turkish WordNet KeNet</i> Merve Doğan, Ceren Oksal, Arife Betül Yenice, Fatih Beyhan, Reyhan Yeniterzi and Olcay Taner Yıldız	85
<i>Homonymy Information for English WordNet</i> Rowan Hall Maudslay and Simone Teufel	90

Workshop Program

Monday, June 20, 2022

- 9:00–9:05 *Welcome*
- 9:05–9:35 *Towards the Profiling of Linked Lexicographic Resources*
Lenka Bajcetic, Seung-bin Yim and Thierry Declerck
- 9:35–10:05 *Towards the Linking of a Sign Language Ontology with Lexical Data*
Thierry Declerck
- 10:05–10:35 *Modelling Collocations in OntoLex-FrAC*
Christian Chiarcos, Katerina Gkirtzou, Maxim Ionov, Besim Kabashi, Fahad Khan
and Ciprian-Octavian Truică
- 10:35–11:05 *Coffee Break*
- 11:05–11:20 **GLOBALEX – Overview and Update**
Ilan Kernerman
- 11:20–11:50 *Creating Bilingual Dictionaries from Existing Ones by Means of Pivot-Oriented
Translation Inference and Logistic Regression*
Yves Bestgen
- 11:50–12:20 *Compiling a Highly Accurate Bilingual Lexicon by Combining Different Ap-
proaches*
Steinþór Steingrímsson, Luke O’Brien, Finnur Ingimundarson, Hrafn Loftsson and
Andy Way
- 12:20–12:50 *A Category Theory Framework for Sense Systems*
David Strohmaier and Gladys Tyen
- 12:50–14:00 *Lunch Break*
- 14:00–14:30 **ELEXIS – Overview and Update**
Simon Krek

Monday, June 20, 2022 (continued)

- 14:30–15:00 *Converting a Database of Complex German Word Formation for Linked Data*
Petra Steiner
- 15:00–15:30 *Resolving Inflectional Ambiguity of Macedonian Adjectives*
Katerina Zdravkova
- 15:30–16:00 *MorphoLex Turkish: A Morphological Lexicon for Turkish*
Bilge Arican, Aslı Kuzgun, Büşra Marşan, Deniz Baran Aslan, Ezgi Saniyar, Neslihan Cesur, Neslihan Kara, Oguzhan Kuyrukcu, Merve Ozcelik, Arife Betül Yenice, Merve Dogan, Ceren Oksal, Gökhan Ercan and Olcay Taner Yıldız
- 16:00–16:30 *Coffee Break*
- 16:30–17:00 *Time Travel in Turkish: WordNets for Modern Turkish*
Ceren Oksal, Hikmet N. Oguz, Mert Catal, Nurkay Erbay, Ozgecan Yuzer, Ipek B. Unsal, Oguzhan Kuyrukcu, Arife B. Yenice, Aslı Kuzgun, Büşra Marşan, Ezgi Saniyar, Bilge Arican, Merve Dogan, Özge Bakay and Olcay Taner Yıldız
- 17:00–17:30 *WordNet and Wikipedia Connection in Turkish WordNet KeNet*
Merve Doğan, Ceren Oksal, Arife Betül Yenice, Fatih Beyhan, Reyhan Yeniterzi and Olcay Taner Yıldız
- 17:30–18:00 *Homonymy Information for English WordNet*
Rowan Hall Maudslay and Simone Teufel
- 18:00–18:05 *Closing*

