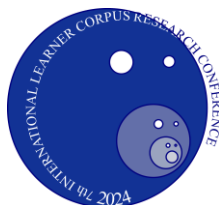




UNIVERSITY OF TARTU

Institute of Foreign Languages
and Cultures



Tartu, 26-28 September 2024

7th Learner Corpus Research Conference

Thursday 26 September 2024	
08:00 - 09:00	Registration (Foyer, Institute of Foreign Languages and Cultures, Lossi 3)
09:00 - 09:30	Opening (White Hall, University of Tartu Museum, Lossi 25)
09:30 - 10:30	Gaëtanelle Gilquin: Opening Pandora's box: Is learner corpus research inclusive enough? Chair: <i>Sylvie De Cock</i> (White Hall, University of Tartu Museum, Lossi 25)
10:30 - 11:30	Coffee break (Foyer, Institute of Foreign Languages and Cultures, Lossi 3)

Attendees should be aware that photography and video recording will occur during the conference. The resulting images and footage may be used on our social media platforms and the event website. If you have any concerns regarding this policy, please contact us.

	Learner corpus based SLA studies Chair: <i>Hilde Hasselgård</i> Lossi 3 (328)	Complexity Chair: <i>Hildegunn Dirdal</i> Lossi 3 (319)	Software for learner corpus analysis Chair: <i>Tomáš Gráf</i> Lossi 3 (117)
11:30 - 12:00	Akira Murakami, Masato Terai, Yu Tamura, Junya Fukuta Proficiency-dependent factors influencing L2 dative alternation	Sandra Götz, Philine Metzger and Fabian Kettenhofen Syntactic Complexity Development in Intermediate learner English: A longitudinal pilot study	Nicolas Ballier and Bernardo Stearns Exploring learner knowledge with Large Language Models fine-tuned with the EFCAMDAT
12:00 - 12:30	Roger W. Gee, M. Karen Jogan, Kathleen S. Jogan Productive use of derivational affixes across proficiency levels	Sylvi Rørvik Syntactic-complexity development in Norwegian learner English	Nathan Vandeweerd Compiling oral learner corpora: Is automatic transcription really worth it?
12:30 - 13:00	Jen-Yu Li and Thomas Gaillat Exploring Verb-Noun collocations across L2 English proficiency levels in L1 Mandarin and French learners	Paolo Brasolin and Arianna Bienati Complexity or complexities? A simulation study on lexical complexity in expert and learner texts through the lens of information theory	Ye Lei Wordless: An integrated corpus tool with multilingual support for the study of language acquisition, pedagogy, and assessment
13:00 - 14:30	Lunch (Ülikooli Kohvik, Ülikooli 20)		
	Learner corpus based SLA studies Chair: <i>Akira Murakami</i> Lossi 3 (328)	Complexity Chair: <i>Sylvi Rørvik</i> Lossi 3 (319)	Pedagogical applications Chair: <i>Nathan Vandeweerd</i> Lossi 3 (117)
14:30 - 15:00	Hilde Hasselgård Of-constructions in the interlanguage of Norwegian learners of English. A pseudo-longitudinal study	Blanka Pojslová The effect of different feedback strategies on lexical sophistication of ESP writing	Lynne Flowerdew The role of learner corpora in needs analysis for ESP (English for Specific Purposes)
15:00 - 15:30	Olga Lopopolo A Diasystematic Construction Grammar (DCxG) analysis of Progressive Aspectuality in Multilingual Learners of English as Additional Language	Andrea Ágnes Reményi Verb phrase versatility as a syntactic complexity indicator in L2-English written texts Cancelled due to unforeseen circumstances	Petter Hagen Karlsen and Susan Nacey Pedagogical applications of learner corpus research: How far have we come?
15:30 - 16:30	Poster session (Foyer, Institute of Foreign Languages and Cultures, Lossi 3)		

Attendees should be aware that photography and video recording will occur during the conference. The resulting images and footage may be used on our social media platforms and the event website. If you have any concerns regarding this policy, please contact us.

16:30 - 17:00	Coffee break (Foyer, Institute of Foreign Languages and Cultures, Lossi 3)		
	Learner corpus design and methodology Chair: <i>Kevin McManus</i> Lossi 3 (328)	Complexity Chair: <i>Sandra Götz</i> Lossi 3 (319)	Language for Academic Purposes Chair: <i>Lynne Flowerdew</i> Lossi 3 (117)
17:00 - 17:30	Lisa Lenort, Annette Portmann, Matthias Schwendemann and Josef Ruppenhofer A metadata scheme for cross-corpus analyses of L2 acquisition	Hildegunn Dirdal, Stine H. Johansen and Philip Durrant Syntactic complexity across L1, L2 and L3	Dámaso Izquierdo Alegría Comparing morphosyntactic features in undergraduate dissertations in Spanish by Estonian and Spanish students: a corpus-driven approach
17:30 - 18:00	Elisa Corino and Eugenio Goría Creating a multimodal corpus of intercomprehension: first reflections on annotation	Natalia Judith Laso Martín and María Belén Díez Bedmar Exploring Noun Lexical Diversity and Noun Phrase Complexity: A Learner Corpus-Based Study of B1 and C1 Spanish EFL Learners' Email Writing	Helena Wedig, Carola Strobl, Jim Ureel, Tanja Mortelmans The influence of L1 Dutch on cohesion in L2 German academic writing: A contrastive corpus-based analysis
18:00 - 18:30	Mahbod Ghaffari and Saeed Safari Cambridge-Belgrade (CamBel) Persian Learner Corpus: Design and Methodology	Henrik Kaatari, Tove Larsson, Ying Wang, Pia Sundqvist and Taehyeong Kim The effect of self-initiated L2 activities on intermediate-level students' lexical complexity	
18:45 - 20:30	Welcome reception at the White Hall of the University of Tartu Museum (Lossi 25) The museum will remain open for attendees to visit freely.		

Attendees should be aware that photography and video recording will occur during the conference. The resulting images and footage may be used on our social media platforms and the event website. If you have any concerns regarding this policy, please contact us.

Friday 27 September 2024

08:30 - 09:00	Registration (Foyer, Institute of Foreign Languages and Cultures, Lossi 3)			
09:00 - 10:00	Cristóbal Lozano: Designing and compiling a multi-L1 corpus of L2 Spanish: Lessons learnt from CEDEL2 Chair: <i>Susan Nacey</i> (White Hall, University of Tartu Museum, Lossi 25)			
10:00 - 10:45	Coffee break (Foyer, Institute of Foreign Languages and Cultures, Lossi 3)			
	Learner corpus design and methodology Chair: <i>Lisa Lenort</i> Lossi 3 (328)	Complexity Chair: <i>Sandra Reitbrecht</i> Lossi 3 (319)	Software demonstration session Chair: <i>Matthias Schwendemann</i> Lossi 3 (117)	Translation studies Chair: <i>Helena Wedig</i> Lossi 3 (227)
10:45 - 11:15	Amanda Huensch Considerations for planning and developing publicly-shared L2 speech corpora	Thomas Gaillat Assessing the validity of new structural complexity metrics as features of L2 proficiency	Mojca Stritar Kučuk KOST, the first learner corpus for Slovene as a second language	Agnieszka Leńko-Szymańska, Łucja Biel and Katarzyna Wasilewska The acquisition of the article system by Polish advanced learners of English: evidence from legal translations
11:15 - 11:45	Minjin Kim and Kevin McManus Design, measurement, and analysis in longitudinal corpus-based SLA research: A systematic review	Andreea Dinca, Ana-Maria Bucur and Madalina Chitez Tracking phraseological inter-complexity: a corpus based contrastive analysis of learner English and Romanian novice academic writing	Tomáš Gráf, Barbora Bulantová and Kryštof Buchal Introducing ETC, the first learner corpus of non-native teacher English	Zuriñe Sanz Villar The Use of Punctuation in a German-to-Basque LTC
11:45 - 12:15		Soroosh Akef, Detmar Meurers, Amália Mendes and Patrick Rebuschat Empirically testing the “Trade-Off Hypothesis”: A machine learning experiment	Kais Allkivi-Metsoja, Taavi Kamarik, Karina Kert and Silvia Maine ELLE Text Evaluator – a corpus-based tool for learning and teaching Estonian	Siri Fürst Skogmo and Susan Nacey Triangulation with learner translation corpora

Attendees should be aware that photography and video recording will occur during the conference. The resulting images and footage may be used on our social media platforms and the event website. If you have any concerns regarding this policy, please contact us.

12:15 - 13:45	Lunch (Ülikooli Kohvik, Ülikooli 20)			
	Learner corpus design and methodology Chair: <i>Amanda Huensch</i> Lossi 3 (328)	Complexity, accuracy and/or fluency Chair: <i>Thomas Gaillat</i> Lossi 3 (319)	Discourse analysis and pragmatics Chair: <i>Mari Kruse</i> Lossi 3 (117)	Assessment and testing Chair: <i>Agnieszka Leńko-Szymańska</i> Lossi 3 (227)
13:45 -14:15	Matthias Schwendemann, Katrin Wisniewski, Torsten Zesch, Lisa Lenort Compiling a multi-corpus database for automatically annotating developmental stages in L2 German	Sandra Reitbrecht Capitalisation and consonant doubling in German as a foreign language – an error analysis of learner texts at different CEFR proficiency levels	Maggie Charles and Karin Whiteside Learner corpus genre analysis: What can it reveal about learners' discursal weaknesses?	Nathan Vandeweerd, Peter Thwaites, Magali Paquot and Jiacheng Shen Using crowdsourced comparative judgement and rubric-based rating to grade texts in the ICLE corpus: a report on reliability and validity
14:15 - 14:45	Olga Lopopolo, Arianna Bienati, Jennifer-Carmen Frey, Aivars Glaznieks and Stefania Spina Discussing the categorization of speakers' language background: implicit assumptions and methodological challenges for Learner Corpus Research	Katarzyna Rokoszewska The development of complexity, accuracy and fluency in L2 English writing at secondary school–group and individual learning profiles	Mei-Hua Chen, Wei-Fan Chen, Garima Mudgal and Henning Wachsmuth A multidimensional comparison of argumentative writing of English learners from different language families	Andrea Hudoušková Collocations with the verbs have, make/do, give and get in L2 Czech: how corpus research can inform CEFR descriptions of Czech at levels A2-B2
14:45 - 15:15	Jelena Kallas, Kristjan Suluste, Raili Pool, Helen Kaljumäe Estonian L2 Learner Corpora: current state and perspectives		María Luisa Carrió-Pastor Teaching interpersonal devices with English learner corpora	
15:15 - 15:45	Coffee break (Foyer, Institute of Foreign Languages and Cultures, Lossi 3)			

Attendees should be aware that photography and video recording will occur during the conference. The resulting images and footage may be used on our social media platforms and the event website. If you have any concerns regarding this policy, please contact us.

	Work in progress session			
	WIP 1 Chair: <i>Valentin Werner</i> Lossi 3 (328)	WIP 2 Chair: <i>Dámaso Izquierdo Alegría</i> Lossi 3 (319)	WIP 3 Chair: <i>Johanna Wolf</i> Lossi 3 (117)	WIP 4 Chair: <i>Olga Lopopolo</i> Lossi 3 (227)
15:45 - 16:00	Lionel Nicolas, Egon W. Stemle, Magali Paquot, Hubert Naets and Alexander König A call for more data-oriented reporting in Learner Corpus research	Ana Abigail Flores Hernández, Hilda Hidalgo Avilés and Abigail Carretero Hernández DMorphC: A web tool for carrying out the identification, grouping and counting of morphologically complex words.	Marge Käsper and Anu Treikelder Creation and exploration perspectives of ScientEst, an Estonian Learner corpus of French Academic Discourse	Erik Castello and Katherine Ackerley EFL students' use of phrasal verbs of action and motion in spoken and written video-clip descriptions
16:00 - 16:15	Oxana Bejenari, Paola Cotta Ramusino, Claudio Macagno, and Tatsiana Maiko Russian Spoken Learner Corpus: Designing and collecting a spoken corpus of Italian university learners of Russian	Reeli Torn-Leesik and Liina Tammekänd The use of English articles in essays written as part of the Estonian National Examination of English	Andrea Lösel, Matthias Schwendemann and Franziska Wallner Quoting and Referencing Mastery: ExpoKo's Insights Into Student Academic Writing Challenges	Adriano Ferraresi and Silvia Bernardini Interacting to learn: Towards an inclusive corpus of EFL learners' interactions with AI-based chatbots
16:15 - 16:30	Kelly Lilles Bridging academic and technological domains. The new framework for developing the Estonian L1 and L2 preschool children's speech corpus.	Piotr Garbacz Structural Characteristics of L3 Norwegian as Spoken by Polish Native Speakers	Andrea Lösel Exploring Formulaic Language in Student Academic Writing in L1 and L2 German	
16:30 - 16:45	Ewa Badyda and Lucyna Warda-Radys The Polish Learner Corpus "FoKo" - main project assumptions	Mari Kruse To what extent do P-bursts resemble lexicogrammatically meaningful chunks? An exploratory study on beginner level foreign language writing processes	Josef Ruppenhofer, Torsten Zesch, Katrin Wisniewski and Anette Portmann Automatic generation of target hypotheses for learner language	
17:00 - 18:00	LCA General Assembly (Foyer, Institute of Foreign Languages and Cultures, Lossi 3)			
18:30 - 20:00	Guided tour of the city of Tartu			
20:00	Conference dinner (Dorpat restaurant, Soola 6)			

Attendees should be aware that photography and video recording will occur during the conference. The resulting images and footage may be used on our social media platforms and the event website. If you have any concerns regarding this policy, please contact us.

Saturday 28 September 2024

Saturday 28 September 2024			
	Learner corpus based SLA studies Chair: <i>Liina Tammekänd</i> Lossi 3 (328)	Accuracy and complexity Chair: <i>Reeli Torn-Leesik</i> Lossi 3 (319)	Lexis Chair: <i>Erik Castello</i> Lossi 3 (227)
09:00 - 09:30	Stefania Spina, Aivars Glaznieks and Andrea Abel <u>Intensification of adjectives in young L2 learners of German and Italian</u>	Valentin Werner and Robert Fuchs <u>Frequency vs. accuracy in learner Englishes: A study on tense and aspect</u>	Svatava Škodová and Melissa Shih-hui Lin <u>Motion Verbs in the Second/Foreign Language Acquisition of Czech: a corpus-based study on non-native speakers of Czech with Chinese L1</u>
09:30 - 10:00	Nozomi Miki and Akira Murakami <u>Navigating the complexity of causality annotation in learner language</u>	Patricia de Crignis and Johanna Wolf <u>What a L1-German Learner Corpus can tell us about the acquisition of Romance Languages: The MuLeCo Project</u>	Marine Makhatadze <u>Georgian English learner corpus and its lexicographic applicability</u>
10:00-10:30	LingLing Xu <u>L1 Influence on Chinese English Learners' Use of Individual Senses of "IN"</u>	Adrian Jan Zasina <u>How does the first language influence the shape of text? A Comparison of Korean and Polish non-native Czech texts</u>	Lingmin Huang <u>The impact of topic on the use of lexical bundles by EFL and ESL learners</u>
10:30 - 11:00		Olesya V. Kisselev, Mikhail Kopotev and Anton Vakhranov <u>Lexical complexity indices as markers of proficiency in L2 Russian</u>	Rita Juknevičienė <u>A Contrastive Analysis of Lithuanian Children's Language</u>
11:00 - 12:00	Brunch (Foyer, Institute of Foreign Languages and Cultures, Lossi 3)		

Attendees should be aware that photography and video recording will occur during the conference. The resulting images and footage may be used on our social media platforms and the event website. If you have any concerns regarding this policy, please contact us.

	Learner corpus based SLA studies Chair: <i>Abigahil Flores Hernández</i> Lossi 3 (328)	Discourse analysis and pragmatics Chair: <i>Svatava Škodová</i> Lossi 3 (319)	
12:00 - 12:30	Sibel Aybek To Collocate or not to Collocate: Exploring Verb-Noun Collocations of Turkish EFL Learners	Shujun Wan, Julián Moreno-Schneider and Georg Rehm Writing argumentative essays: Discourse strategies in L1- and L2-authored versus ChatGPT-generated text	
12:30 - 13:00	Tatiana Aleksandrova Thematic progression in written argumentative paragraphs of Chinese advanced learners of French		
13:15 - 14:15	Ilmari Ivaska: Cross-linguistic influences in different levels of granularity: How are they different and why does it matter for LCR? Chair: <i>Liina Lindström</i> (Assembly hall, Main building of Tartu University, Ülikooli 18)		
14:15 - 14:45	CLOSING CEREMONY (Assembly hall, Main building of Tartu University, Ülikooli 18)		

Attendees should be aware that photography and video recording will occur during the conference. The resulting images and footage may be used on our social media platforms and the event website. If you have any concerns regarding this policy, please contact us.

POSTERS

Aylin Braunewell, Julia Schlauch and Jana Gamper. [Levels and Modeling of Variability in L2 Learner Corpora. Insights from a Corpus of Newly Migrated Students in Germany](#)

Eugenio Gorla and Caterina Mauri. [DiverSIta: a new corpus for the documentation of L2 Italian](#)

Mingyan Gao. [Understanding Foundation-level ESL Students' Academic Writing: A Lexico-grammatical Analysis across Genres and Disciplines](#)

Petter Hagen Karlsen and Sylvi Rørvik (Inland Norway University of Applied Sciences). [Subject-Verb Agreement in English Learner Texts: A Pseudo-Longitudinal Perspective](#)

Senora Mustun, Mauselyn Pattikawa, Yosef Tanggu Solo, Vonny Juliana Ruhulestin, Yeremina Karolina Ngabalin, Theresia Astrie Nunumete. [Building the young learner corpus - A longitudinal, bilingual \(English-Indonesian\) corpus: Challenges of collecting data from different educational settings.](#)

Siyuan Liu. [Construction, transcription and annotation of a longitudinal multimodal interlanguage corpus: An ongoing study with L1 Italian and L2 Chinese](#)

Sylvie De Cock, Damien De Meyere, Liesbeth Degand, Thomas François, Gaëtanelle Gilquin, Sylviane Granger, Danqing Huang, Marie-Aude Lefer, Hubert Naets, Magali Paquot, Kristel Van Goethem, Patrick Watrin with the collaboration of Jennifer Carmen-Frey, Alex Koëning and Egon Stemle. [Introducing the CLARIN Knowledge Centre for Learner Corpora](#)

Sylvie De Cock and Jennifer Thewissen. [Getting down to business with the Apprentice Multiple Business GEnRes \(AMBER\) corpus](#)

Tobias Pauls. [SEEFLEX – The Corpus of Secondary School English as a Foreign Language \(EFL\) Exams](#)

Valentin Werner, Robert Fuchs, Anna Rosen, Lea Bracke and Bethany Stoddard. [The Corpus of Young German Learner English](#)

Attendees should be aware that photography and video recording will occur during the conference. The resulting images and footage may be used on our social media platforms and the event website. If you have any concerns regarding this policy, please contact us.