

THE 4TH

ASIA PACIFIC CORPUS LINGUISTICS CONFERENCE (APCLC 2018)

Takamatsu, Japan, September 17-19, 2018

Conference program



Plenary speech 1:

Corpus linguistics and the study of fiction – methodological and theoretical challenges

Michaela Mahlberg (University of Birmingham, UK)

Abstract:

Corpus research has traditionally focused on non-literary texts. Along with developments in Digital Humanities, there seems to be an increasing interest into the study of fictional texts, sometimes referred to under the umbrella term ‘corpus stylistics’ (Semino and Short 2004). In order to be able to account as fully as possible for features of literary texts we need to create new tools and develop methodologies that are tailored to the task at hand. In this paper, I will illustrate key functionalities of the web application CLiC (<http://clic.bham.ac.uk/>) that has been specifically designed for the corpus linguistic study of narrative fiction. The case studies that I will present look at textual patterns that contribute to the creation of fictional characters. The examples will be drawn from the CLiC corpora. The CLiC corpora comprise more than 130 books across four subcorpora: the corpus of Dickens’s Novels, the 19th Century Reference Corpus (19C), the Corpus of 19th Century Children’s Literature (ChiLit) and the Corpus of Additional Requested Texts (ArTs). For all CLiC texts, direct speech and specific places around speech have been marked up (Mahlberg et al. 2016). Hence, CLiC can run searches across defined textual subsets and support the analysis of features of narrative fiction. An important question is how a range of features and patterns in fiction can be brought together in a coherent theoretical framework. My suggestions towards such a framework focus on a lexically-driven approach to body language and raise more fundamental questions about how far corpus linguistics can change our theoretical perspective on fiction.

References

- Mahlberg, M., Stockwell, P., Joode, J. de, Smith, C., & O’Donnell, M. B. (2016). CLiC Dickens: novel uses of concordances for the integration of corpus stylistics and cognitive poetics. *Corpora*, 11(3), 433–463.
<https://www.eupublishing.com/doi/full/10.3366/cor.2016.0102>
- Semino, E., & Short, M. (2004). *Corpus Stylistics. Speech, Writing and Thought Presentation in a Corpus of English Writing*. London: Routledge.



Michaela Mahlberg is Professor of corpus linguistics at the University of Birmingham, UK, where she is also the Director of the Centre for Corpus Research and the Director of Research and Knowledge Transfer for the College of Arts and Law. Michaela is the editor of the *International Journal of Corpus Linguistics* (John Benjamins) and together with Wolfgang Teubert she edits the book series *Corpus and Discourse* (Bloomsbury). One of her main areas of research is Dickens’s fiction and the socio-cultural context of the 19th century. Her publications include *Corpus Stylistics and Dickens’s Fiction* (Routledge, 2013), *English General Nouns: a Corpus Theoretical*

Approach (John Benjamins, 2005) and *Text, Discourse and Corpora. Theory and Analysis* (Continuum, 2007, co-authored with Michael Hoey, Michael Stubbs and Wolfgang Teubert). Michaela was the Principal Investigator on the AHRC-funded project CLiC Dickens: Characterisation in the representation of speech and body language from a corpus linguistic perspective which led to the development of the CLiC web app.

Website:

<https://www.birmingham.ac.uk/staff/profiles/elal/mahlberg-michaela.aspx>

Plenary speech 2:

Understanding and Advancing the Data-Driven Learning (DDL) Approach

Laurence Anthony (Waseda University, JAPAN)

Abstract:

Data-Driven Learning (DDL) is an inductive, self-directed approach to language-learning in which learners interact directly or indirectly with a corpus under the guidance of the language instructor. Numerous empirical studies have shown DDL to be effective. DDL can empower learners to address their individual discipline-specific language needs in the classroom. It has also been shown to work with learners of different proficiency levels and ages, and in classrooms with and without computers installed. However, reports have revealed that both instructors and learners face major challenges when using the DDL approach. These include finding a suitable corpus, knowing what to search for when a corpus is available, reducing data overload when making observations, incorporating findings as part of learning, and contextualizing what is learned in the target field and discourse community. Other challenges relate to the download, installation and use of corpus analysis software. In this paper, I will first present a case for using DDL in the language learning classroom. I will then explain how the various challenges associated with DDL can be minimized or fully addressed through careful preparation, effective classroom practices, and the use of innovative support tools. Finally, I will discuss how language program administrators and instructors can contribute to future DDL tools and methods development.



Laurence Anthony is Professor of Applied Linguistics at the Faculty of Science and Engineering, Waseda University, Japan. He has a BSc degree (Mathematical Physics) from the University of Manchester, UK, and MA (TESL/TEFL) and PhD (Applied Linguistics) degrees from the University of Birmingham, UK. He is a former Director and the current coordinator of graduate school English in the Center for English Language Education in Science and Engineering (CELESE). His main research interests are in corpus linguistics, educational technology, and English for Specific Purposes (ESP) program design and teaching methodologies. He received the National Prize of the Japan Association for English Corpus Studies (JAECS) in 2012 for his work in corpus software tools design. He is the developer of various corpus tools including AntConc, AntCorGen, AntWordProfiler, FireAnt, ProtAnt, and TagAnt.

Website:

<http://www.laurenceanthony.net/>

Plenary speech 3:

Distributional Methods in Corpus Linguistics: Towards a Hermeneutic Cyborg

Stefan Evert (Friedrich-Alexander-Universität, Germany)

Abstract:

Distributional approaches in natural language processing are based on the assumption that the meaning of a linguistic item (typically a word or phrase) can be inferred from its distribution in language. More precisely, the distributional hypothesis stipulates that two items differ in meaning to the same degree as they differ in their distribution, which is operationalized in the form of a high-dimensional co-occurrence feature vector. Such distributional models (DM) can flexibly be applied to a wide range of tasks depending on how the target and feature items are chosen (texts, sentences, phrases, words, morphemes, or word pairs).

After giving an overview of the fundamental techniques and parameters of distributional approaches, my talk focuses on two applications in the field of corpus linguistics. I will use these examples to illustrate how DM can help overcome shortcomings of essential corpus analysis techniques.

At the level of texts, DM provide an unsupervised account of stylistic variation across authors, registers and domains. As a result, such effects can systematically be included in frequency comparisons and similar quantitative analyses, leading to much more accurate findings.

At the level of words, DM create a fine-grained semantic representation based on collocational profiles, which play a central role in computational lexicography and corpus-based discourse analysis. As a result, innovative corpus analysis tools can be developed that combine sophisticated computational techniques with human interpretation into an interactive hermeneutic procedure – merging man and machine into what I like to call the Hermeneutic Cyborg.



Stefan Evert holds the Chair of Computational Corpus Linguistics at the University of Erlangen-Nuremberg, Germany. After studying mathematics, physics and English linguistics, he received a PhD degree in computational linguistics from the University of Stuttgart, Germany. His research interests include the statistical analysis of corpus frequency data (significance tests in corpus linguistics, statistical association measures, Zipf's law and word frequency distributions), quantitative approaches to lexical semantics (collocations, multiword expressions and distributional semantics), multidimensional analysis (linguistic variation, language comparison, translation studies), as well as processing large text corpora (IMS Open Corpus Workbench, data models and query languages, Web as corpus, sentiment analysis, text clustering).

Web site:

www.stefan-evert.de
www.linguistik.fau.de

Plenary speech 4:

How Deep Learning Changes Natural Language Processing

Naoaki Okazaki (Tokyo Institute of Technology, Japan)

Abstract:

When I first saw the success of deep learning in other research areas, I did not expect that it would greatly influence the research of Natural Language Processing (NLP). Although we are still not sure whether deep learning truly advances the technologies to realize computers that can truly understand the language, it is now very common in the NLP community, achieving the state-of-the-art performance in most tasks. In this talk, I briefly overview the advantage of deep learning over the conventional methods of machine learning, e.g., automatic feature extraction, generic gradient-based learning, end-to-end learning, and versatile software framework. I then explain the key ideas of deep learning that have widely been accepted in NLP: distributed representations of words/phrases/sentences, encoder-decoder models, attention mechanisms, etc. Deep learning has not only provided an alternative approach to the statistical NLP, but also bridged NLP to other research areas and increased the ‘bravery’ of NLP research. I will explain the recent trends of NLP research including multi-modal processing and context modeling. I conclude this talk by summarizing the future prospect of NLP.



Naoaki Okazaki is a professor at School of Computing, Tokyo Institute of Technology. Prior to this position, he worked as a research fellow in National Centre for Text Mining (NaCTeM) (in 2005), as a post-doctoral researcher in University of Tokyo (in 2007-2011), and as an associate professor in Graduate School of Information Sciences, Tohoku University (2011-2017). He obtained his PhD from Graduate School of Information Science and Technology, University of Tokyo in 2007. He has served as a technical consultant in SmartNews Inc. since 2013. He is also a visiting research scholar of the Artificial Intelligence Research Center (AIRC), National Institute of Advanced Industrial Science and Technology (AIST). His research interests include natural language processing, text mining, and machine learning. He is also

known as a developer open-source software toolkits including: CRFSuite (a fast implementation of Conditional Random Fields), SimString (an implementation for approximate string matching), and libLBFGS (a C library of L-BFGS).

His research is highly regarded nationally and internationally. Dr. Naoaki Okazaki has published more than 120 refereed papers in journals and conference proceedings including ACL, NAACL, EMNLP, COLING, IJCNLP, and AACL. He is a recipient of various awards: the Young Scientists' Prize, the Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology (MEXT); Funai Research Award; Docomo Mobile Science Award; and several Best Paper Awards at PACLIC 29 and domestic NLP conferences in Japan. In addition, he has contributed prominently to the academic and scientific community. He has served as area chairs (ACL 2012 and 2016), workshop co-chairs (IJCNLP 2013), and publication chair (EMNLP-CoNLL 2012). He joined numerous program committees for international conferences including AACL, IJCAI, ACL, NAACL, EMNLP, EACL, and COLING. He has also been in the TACL elite standing reviewing pool.

Web site:

<http://www.chokkan.org/index.html.en>

Plenary speech 5:

The Individual and the Group in Corpus Linguistics

Michael Barlow (University of Auckland, New Zealand)

Abstract:

In this presentation I examine samples of the spoken output of five White House Press Secretaries in order to investigate the nature and extent of individual differences in production. While there are many idiosyncratic differences in grammar and style in the speech of individuals, the aim here is to investigate very frequent components of grammar. In this study, we can examine differences in the distribution of the most common bigrams in spoken usage. The results show that even with these common bigrams such as of the, there are major distinctions in usage by different press secretaries and we find that the inter-speaker variation for a range of bigrams is much greater than the intra-speaker variation. The results pose interesting questions for the relationship between the language of the individual and the language of the group and for the representation of grammar in relation to production and comprehension.



Michael Barlow received his PhD in Linguistics from Stanford University. He is currently Associate Professor in Applied Language Studies at the University of Auckland in New Zealand. Dr. Barlow has written books and articles on corpus linguistics and regularly gives presentations, courses and workshops at institutions and conferences around the world. He has created several text analysis programs including concordancers MonoConc and ParaConc and a collocation extraction program, Collocate. A recently developed program, WordSkew, is designed to apply corpus analysis techniques while at the same time taking note of the structure of texts. He is also one of the founding members of Asia Pacific Corpus Linguistics Association.

Website

<http://michaelbarlow.com/>

ASIA PACIFIC CORPUS LINGUISTICS CONFERENCE 2018

PROGRAM

Sunday, September 16th

10:00-13:00 Session WS1: Workshop 1 **Classroom B**

Pre-conference workshop 1:

Stefan Evert, **Multivariate analysis (with R)**

10:00-13:00 Session WS2: Workshop 2 **Classroom C**

Pre-conference workshop 2:

Michael Barlow, **WordSkew: Text and Corpus**

14:00-17:00 Session WS3: Workshop 3 **Classroom B**

Pre-conference workshop 3:

Mike Scott, **Exploring Features of WordSmith 7.0**

14:00-17:00 Session WS4: Workshop 4 **Classroom C**

Pre-conference workshop 4:

Stephen Jeaco, **The Prime Machine – a tool to help English language learners get started with the exploration of English language and literature**

Monday, September 17th

09:00-09:20 Conference Hall

Opening ceremony (CHAIRS: [Hitoshi Isahara](#) and [Yukio Tono](#))

09:20-10:20 Conference Hall (CHAIRS: [Michael Barlow](#))

Plenary speech 1:

Corpus Linguistics and the Study of Fiction – Methodological and Theoretical Challenges
Michaela Mahlberg (University of Birmingham, UK)

10:20-10:40 Coffee Break

10:40-12:40 Session 1A: Room 54 (CHAIR: [Shin'Ichiro Ishikawa](#))

10:40 [Martin Weisser](#)

Automatically Enhancing Tagging Accuracy and Readability for Common Freeware Taggers

11:10 [Sai Ma](#), [Guangsa Jin](#) and [Michael Barlow](#)

Using gestures in speaking a second language ---A study based on a multi-modal corpus

11:40 [Mike Scott](#)

Corpus Filtering and Clean-up Operations

12:10 [Hai Xu](#), [Xiaofei Lu](#) and [Vaclav Brezina](#)

Acquisition of the Chinese Particle *le* by L2 Learners: A Corpus-based Approach

10:40-12:40 Session 1B: Room 61-1 (CHAIR: [Yoshiyuki Notohara](#))

10:40 [Shiho Nobesawa](#) and [Tomohisa Sano](#)

A Study on the Use of Color Wheel for the Representation of Emotions

11:10 [Keiko Mochizuki](#), [Yaming Shen](#), [Zhang Zheng](#) and [Laurence Newbery-Payton](#)

Acquisition of Tense/Aspect markers in Learners Corpora of English/Chinese/Japanese

11:40 [Richard Davis](#) and [Yu-Chi Wang](#)

MWE in a small learner corpus: Lexical bundles, formulaic phrases, and light verbs across four levels of learners

12:10 [Huanhuan Wang](#) and [Suxing Zhao](#)

A Corpus-based Investigation of Chinese EFL Learners' Use of Modal Verbs in Writing for Practical Purposes

10:40-12:40 Session 1C: Room 61-2 (CHAIR: [Emi Izumi](#))

10:40 [Joe Geluso](#) and [Roz Hirsch](#)

A qualitative analysis of the effect of reference corpora and choice of statistic on keyword analysis

11:10 [Yoshiho Satake](#)

The effects of corpus use on correction of article- and preposition-omission errors

11:40 [Emi Izumi](#)

Errors and Beyond—A Corpus-based Stylistic Analysis of “Japanese English” Discourse

12:10 [John Blake](#)

Corpus-based Error Detector for Computer Scientists

10:40-12:40 Session 1D: Room 62 (CHAIR: [Yuka Ishikawa](#))

- 10:40 [Vandana Mishra](#), [Niladri Sekhar Dash](#) and [Jayshree Chakraborty](#)
Corpus based Critical Discourse Analysis: A study of Political Ideology in Hindi Newspapers
- 11:10 [Tehseen Zahra](#), [Omama Tanvir](#) and [Ms Khoula](#)
A Corpus-based Critical Discourse Analysis of Racial Stereotyping in American Newspapers
- 11:40 [Alaa Mamdouh Akef](#), [Yingying Wang](#), [Andalib Labna](#) and [Erhong Yang](#)
Chinese-Arabic parallel corpus-based study of Chinese body language chengyu

10:40-12:40 Session 1E: Room 63 (CHAIR: [Deny Kwary](#))

- 11:10 [Philipp Heinrich](#)
Stylistic Features in Corporate Disclosures and Their Predictive Power
- 11:40 [Katsumi Goto](#)
Is the Continuation of The Mystery of Edwin Drood a Posthumous Work of Charles Dickens?: A Multivariate Analysis
- 12:10 [Reiko Ikeo](#)
A corpus-driven analysis of present-tense fiction

10:40-12:40 Session 1F Room 64 (CHAIR: [Mariko Abe](#))

- 10:40 [Tehseen Zahra](#), [Akhtar Abbas](#) and [Wasima Shehzad](#)
Corpus-based Approaches to ELT: Current Situation and Future Implications in Pakistan
- 11:10 [Ali Al-Laith](#) and [Muhammad Shahbaz](#)
A Large-scale Annotated Corpus for Opinion Analysis in Arabic News from Twitter
- 11:40 [Katharine Levendis](#) and [Andreea S. Calude](#)
“When you don’t know what you know” – Use of Māori loanwords in a diachronic corpus of New Zealand English

12:40-14:00 Lunch Break

14:00-15:00 Conference Hall (CHAIRS: [Michael Barlow](#))

Plenary speech 2:

Understanding and Advancing the Data-Driven Learning (DDL) Approach
Laurence Anthony (Waseda University, JAPAN)

15:00-16:00 Poster session

Poster session

- 15:00 [Haihong Huang](#) and [Toshiyuki Kanamaru](#)
The construction of a Business Katakana Word List using an instant domain-specific web corpus and BCCWJ
- 15:00 [Kyoko Kanzaki](#) and [Hitohsi Isahara](#)
Corpus based analysis of semantic relations between two verbal constituents in lexical compound verbs
- 15:00 [Noriaki Katagiri](#) and [Yukiko Ohashi](#)
Visualizing English Classroom Spoken Data on Multi-modal Interface to Create Versatile Linguistic Resources
- 15:00 [Yuichiro Kobayashi](#), [Yusuke Kondo](#) and [Mariko Abe](#)
Predicting EFL learners' oral proficiency levels in monologue tasks
- 15:00 [Yingying Lao](#), [Takumi Hosokawa](#) and [Dongli Han](#)
Elucidation of the influence of the structure of lyrics on the ease of understanding lyrics
- 15:00 [Minoru Sasaki](#)
Multi-Domain Word Embeddings for Semantic Relation Analysis among Domains
- 15:00 [Naoki Sugimori](#)
Functional Classification of Lexical Bundles in TED Talks
- 15:00 [Hiroki Tanioka](#), [Kaoru Kimura](#), [Kazuma Takaoka](#), [Ryohei Nakatani](#) and [Yoshitaka Uchida](#)
Automatic Generation of Japanese Question-Answering Pairs
- 15:00 [Dax Thomas](#)
Exploring word-formation in science fiction using a small corpus
- 15:00 [Mayumi Uchida](#) and [Katsunori Kotani](#)
Comment Corpus for Study Abroad Program Evaluation
- 15:00 [Eiko Yamamoto](#) and [Hitoshi Isahara](#)
Study on tour guide support integrating real examples and Web resources
- 15:00 [Jiawei Yong](#), [Katsumi Kanasaki](#), [Kiyohiko Shinomiya](#), [Shintaro Kawamura](#) and [Shohichi Naitoh](#)
A Semi-Supervised Clustering Method for Dialogue Identification in Meeting Minutes Corpora
- 15:00 [Yo Sato](#) and [Kevin Heffernan](#)
How distinctive is your corpus? A metric for the degree of deviation from the norm

16:00-16:20 Coffee Break

16:20-17:50 Session 2A: Room 54 (CHAIR: [Shin'Ichiro Ishikawa](#))

- 16:20 [Hitoshi Isahara](#), [Kyoko Kanzaki](#), [Shigeru Nemoto](#), [Akio Yoshida](#) and [Kozo Moriguchi](#)
Can Neural Machine Translation System Create Training Data?
- 16:50 [Stephen Jeaco](#)
Mixing and Matching DIY and Ready-made Corpora: Introducing the DIY Text Tools built into The Prime Machine

16:20-17:50 Session 2B: Room 61-1 (CHAIR: [Yoshiyuki Notohara](#))

- 16:20 [Fu-Ying Lin](#)
A case study on the features of Spoken Academic English: “a/an+N+of construction”
- 16:50 [Muhammad Din](#)
A Corpus-based Study on English Adjectives Ending with Suffix -ly
- 17:20 [Kayo Yamamoto](#), [Tamao Araki](#) and [Richard S. Lavin](#)
Lexical characterization of semi-popularization articles on agricultural topics

16:20-17:50 Session 2C: Room 61-2 (CHAIR: [Emi Izumi](#))

16:20 [Ping-Yu Huang](#)

Using Corpus-derived Co-occurrence Information to Extract Terminology

16:50 [Takashi Yamamura](#), [Masato Hino](#) and [Kazutaka Shimada](#)

Dialogue Act Annotation and Identification in a Japanese Multi-party Conversation Corpus

17:20 [John Blake](#)

Inter-annotator agreement: By hook or by crook

16:20-17:50 Session 2D: Room 62 (CHAIR: [Yuka Ishikawa](#))

16:20 [Hisae Akihiro](#)

Discourse function of après in French informal conversations

16:50 [Mahul Bhattacharyya](#), [Atanu Saha](#) and [Niladri Sekhar Dash](#)

A Usage Based Analysis of Bangla Discourse Markers in Mass Media Texts

17:20 [Philipp Heinrich](#) and [Fabian Schäfer](#)

Extending Corpus-Based Discourse Analysis for Exploring Japanese Social Media

16:20-17:50 Session 2E: Room 63 (CHAIR: [Deny Kwary](#))

16:20 [Da-Chen Lian](#) and [Shu-Kai Hsieh](#)

A Topic-aware Comparable Corpus of Chinese Variations

16:50 [Deny Kwary](#)

Towards the First Online Indonesian National Corpus

17:20 [Yukio Tono](#)

Developing multilingual language learning resources using the CEFR-J

16:20-17:50 Session 2F: Room 64 (CHAIR: [Mariko Abe](#))

16:20 [Guoxiao Tao](#), [Fan Pan](#) and [Alex Chengyu Fang](#)

A Developmental Study of Prepositional Phrases in English L1 Students' Academic Writing

16:50 [Nobuko Tahara](#)

Deviant Uses of Metadiscursive Nouns in NNS Essays

Tuesday, September 18th

09:00-10:00 Conference Hall (CHAIR: [Yukio Tono](#))

Plenary speech 3:

Distributional Methods in Corpus Linguistics: Towards a Hermeneutic Cyborg
[Stefan Evert](#) (Friedrich-Alexander-Universität, Germany)

10:00-10:20 Coffee Break

10:20-12:20 Session 3A: Room 54 (CHAIR: [Satoru Uchida](#))

10:20 [Hyeon-Okh Kim](#) and [Hye-Kyung Lee](#)

Identifying the vocabulary size for Korean EFL students in linguistics

10:50 [Masatoshi Sugiura](#), [Yoshito Nishimura](#) and [Daisuke Abe](#)

Distinguishing L1 and L2 Using Three Linguistic Aspects: A Logistic Regression Model Study

11:20 [Shin'Ichiro Ishikawa](#)

S-genitives and Of-genitives Seen in L2 English Learners' Essays: A Study Based on the ICNALE Written Essays

11:50 [Chikako Nishigaki](#), [Kiyomi Chujo](#) and [Shiro Akasegawa](#)

A Japanese/English DDL Tool for Primary School CEFR Pre-A1 EFL Learners

10:20-12:20 Session 3B: Room 61-1 (CHAIR: [Aika Miura](#))

10:20 [Hikaru Misumi](#) and [Yukio Tono](#)

A corpus-based analysis of the use of subordinators by Japanese learners of English

10:50 [Takehiro Teraoka](#)

Analysis of Associative Information for Second Language Learning of Japanese

11:20 [Midori Tanimura](#) and [Masataka Yamaguchi](#)

Politeness and consensus building strategies in a task-based corpus of English learners

11:50 [Aika Miura](#)

Annotating the functions of learner utterances from a spoken corpus and assigning the degrees of their grammatical accuracy and discursal acceptability

10:20-12:20 Session 3C: Room 61-2 (CHAIR: [CK Jung](#))

10:20 [Alexandra Garcia](#)

Dead of a lesser God: Victims' voice and representation in the Colombian press

10:50 [Totok Suhardijanto](#), [Arawinda Dinakaramani](#) and [Rahmad Mahendra](#)

Developing Language Resources for Indigenous Languages in Indonesia: Annotated Indonesian and Javanese Corpus Building

11:20 [Pilar Valverde](#)

Design principles and data collection for CELEN: a corpus of Learner Spanish in Japan.

11:50 [Sun-Hee Lee](#), [Beomil Kang](#) and [Hye-Ryung Yu](#)

A Corpus-based Analysis of Formulaic Expressions in Korean Academic Prose

10:20-12:20 Session 3D: Room 62 (CHAIR: [Yasutake Ishii](#))

- 10:20 [Sylviane Cardey](#) and [Peter Greenfield](#)
Scientific Findings and their Markers
- 10:50 [Hongbo Zheng](#) and [Xuemei Zhang](#)
Compiling Chinese Learner Dictionary with Corpora: practices and challenges
- 11:20 [Xuhua Zhang](#)
The phraseological treatment of non-high-frequency words by online dictionaries for English learners
- 11:50 [Naho Kawamoto](#)
What determines the quality of “good examples” for CEFR-levels?

10:20-12:20 Session 3E: Room 63 (CHAIR: [Yuichiro Kobayashi](#))

- 10:20 [Mei Yung Vanliza Chow](#)
The representation of nurses in oncology and paediatric nursing research articles - who are they and what do they do?
- 10:50 [Ge Gao](#)
Is the relationship between semantic prosodies and transitivity reliable? A corpus-based cross-linguistic investigation
- 11:20 [Chie Urawa](#)
Corporate Cautionary Statements (Disclaimers): A Critical Genre Analysis of Professional Communication
- 11:50 [Helen Wan](#)
Combining Corpus Linguistics and Critical Discourse Analysis to study the “Leftover Woman” phenomenon in Hong Kong press

12:20-13:40 Lunch Break

13:40-15:10 Session 4A: Room 54 (CHAIR: [Satoru Uchida](#))

- 13:40 [Hiroko Usami](#)
CEFR Receptive and Productive Vocabulary Knowledge of Japanese English Learners
- 14:10 [Yasutake Ishii](#) and [Yukio Tono](#)
Investigating Japanese EFL learners’ overuse/underuse of English grammar categories and their relevance to CEFR levels
- 14:40 [Satoru Uchida](#) and [Masashi Negishi](#)
Assigning CEFR-J levels to English texts based on textual features

13:40-15:10 Session 4B: Room 61-1 (CHAIR: [Aika Miura](#))

- 13:40 [Ying Lin](#)
A Corpus-based Study on the Use of Linking Adverbials in Master Theses in Psychology by Taiwanese EFL Learners
- 14:10 [Cui Xu](#)
A Corpus-based Study of Stancetaking as Seen from Stance Adverbials in Interpreter-mediated Political Discourse
- 14:40 [Kason Chien](#)
A Corpus-based Study on Linking Adverbials in Academic Writing by Taiwanese EFL Students

13:40-15:10 Session 4C: Room 61-2 (CHAIR: [Martin Weisser](#))

13:40 [Raksangob Wijitsopon](#)

The textual functions of lexis: A case study of "nowadays" in learner and English – Thai general corpora

14:10 [Cheng-Hao Chang](#)

Examining the usage of phrasal verbs in theses of Taiwanese TESOL MA students versus TESOL papers in foreign journals

14:40 [Jian Yu](#) and [Moongee Jeon](#)

Specific Syntactic Complexity Measures Across Two Corpora of EFL Writing

13:40-15:10 Session 4D: Room 62 (CHAIR: [Michael Barlow](#))

13:40 [Chen Chuan Lo](#)

Lexical Bundles and Disciplinary Variation in EFL Master's Thesis

14:10 [Andy Seto](#)

Speech Acts in the Hong Kong Corpus of Spoken English (HKCSE)

14:40 [Yuka Ishikawa](#)

A Corpus-based Study on Frequent Noun Phrases in Engineering Academic Texts

13:40-15:10 Session 4E: Room 63 (CHAIR: [Yuichiro Kobayashi](#))

13:40 [Ying Yang](#)

A Corpus-based Cognitive Study on the C-E Translation of Chinese Political Publicity Texts

14:10 [Ta-Ya Lin](#)

Comparing the Hedging Devices Used in Two NNS Academic Writing Corpora

14:40 [Huichen S. Hsiao](#) and [Lestari Mahastuti](#)

Alternative Look: A Collostructional Analysis of Ditransitive Constructions in Mandarin

15:10-18:00 Pre-banquet tour to Ritsurin Garden

18:00-20:00 Banquet at Ritsurin Garden

Wednesday, September 19th

09:00-10:00 Conference Hall (CHAIR: [Hitoshi Isahara](#))

Plenary speech 4:

How Deep Learning Changes Natural Language Processing
[Naoaki Okazaki \(Tokyo Institute of Technology, Japan\)](#)

10:00-10:20 Coffee Break

10:20-11:50 Session 5A: Room 61-1 (CHAIR: [Naoki Sugimori](#))

10:20 [Howard Hao-Jan Chen](#) and [Nina Yiting Tsai](#)

Using a Large Social Science Corpus to Build an Automatic Writing Suggestion System

10:50 [Kunihiko Miura](#)

The Use of Active Reading to Integrate Language Skills: A Data-Driven Learning Approach

11:20 [Mitsuhiro Morita](#), [Satoru Uchida](#) and [Yuka Takahashi](#)

Prefixes and suffixes in Japanese junior high school English textbooks

10:20-11:50 Session 5B: Room 61-2 (CHAIR: [Hiroko Usami](#))

10:20 [Yoshiyuki Notohara](#)

A Usage-based Approach to Criterial Canonical Construction in an L2 Learner Corpus

10:50 [Saeiko Komori](#), [Masatoshi Sugiura](#) and [Wenping Li](#)

Examining the applicability of the mean dependency distance (MDD) for SLA: A case study of Chinese learners of Japanese as a second language

10:20-11:50 Session 5C: Room 62 (CHAIR: [CK Jung](#))

10:20 [Sun-Hee Lee](#) and [Chanyoung Lee](#)

A Corpus-based Analysis of Hedging in Korean Academic Texts

10:50 [Chae Kwan Jung](#) and [Seok-Chae Rhee](#)

New English Education in North Korea: A Corpus Approach

11:20 [Seonmin Park](#)

Trends of Interdisciplinary Corpus Research Conducted by Korean Undergraduate Students Majoring in Sciences

10:20-11:50 Session 5D: Room 63 (CHAIR: [Atsuko Umesaki](#))

10:20 [Somjit Jiranantiporn](#)

Is this "problem" giving you "trouble"? A corpus-based examination of the differences between the nouns "problem" and "trouble"

10:50 [Atsuko Umesaki](#)

A phraseological account of structural patterns: "it is no NN that" and "there is no NN that"

11:20 [Wei Lin](#) and [Alex Chengyu Fang](#)

The Transitivity Continuum in English Non-finite Verbal Complement Clauses: A Corpus-based Study

11:50-13:00 Lunch Break

13:00-14:00 Conference Hall (CHAIR: [Yukio Tono](#))

Plenary speech 5:

The Individual and the Group in Corpus Linguistics
Michael Barlow (University of Auckland, New Zealand)

14:00-14:20 Coffee Break

14:20-15:50 Session 6A: Room 61-1 (CHAIR: [Naoki Sugimori](#))

14:20 [Hao-Jan Howard Chen](#)

Analyzing English Verb-Noun Miscollocations Extracted from a Large Chinese/Taiwanese EFL Learner Corpus

14:50 [Shuji Hasegawa](#)

An Analysis of the Vocabulary Used in Oxford Reading Tree: As a Reference for Early English Education in Japan

15:20 [Amma Kazuo](#)

Extracting Patterns from Transition of Occurrence Frequency of Grammar Items in a Junior High School Textbook in Japan

14:20-15:50 Session 6B: Room 61-2 (CHAIR: [Hiroko Usami](#))

14:20 [Simon Tuchais](#)

Corpus-based study of the spontaneous forms omowareru and omoeru in Japanese

14:50 [Kelvin Lee](#)

The many faces of Lelouch Lamperouge: A corpus approach to language in Japanese animation

15:20 [Kevin Heffernan](#) and [Yo Sato](#)

The varying complexity of the syntactic role of nouns: Evidence from Japanese corpora

14:20-15:50 Session 6C: Room 62 (CHAIR: [Yoko Iyeiri](#))

14:20 [Yoon Jung Oh](#)

Structural approaches for understanding diachronic culture corpus

14:50 [Akira Okada](#)

Historical Analyses of English Negative prefixed words from ME to PE - '-able' adjectives -

15:20 [Yoko Iyeiri](#)

Negation in Benjamin Franklin's Writings: A Stylistic Analysis of his Autobiography and Letters

14:20-15:50 Session 6D: Room 63 (CHAIR: [Yasutake Ishii](#))

14:20 [Daehyeon Nam](#)

Lexical bundles in ESL science and technology perspective majors' writing

14:50 [Kevin Jiang](#)

"This is because...": authorial practice of (un)attending this in academic prose across disciplines

15:20 [Fei Lian](#)

Genre-based comparative study of German word length

16:00-16:30 Conference Hall

Closing ceremony (CHAIRS: [Hitoshi Isahara](#) and [Yukio Tono](#))