

# 英語コーパス学会

## 第46回大会資料

日時：2020年10月3日（土） - 10月4日（日）

# 英語コーパス学会 第46回大会プログラム

October 3 (Sat), 2020

## Day 1 Schedule

Time	Event
08:00-09:00	Registration/Setup/Getting to know each other
09:00-09:30	Opening ceremony
09:30-10:30	Plenary 1: Jesse Egbert (Northern Arizona University). 'Corpus Linguistics' or 'Linguistics with a Corpus'?
10:30-12:00	Site exploration/Video viewing/Meeting others
12:00-13:00	Lunch break
13:00-13:25	<b>Presentation Q&amp;A session 1 (JAECS members)</b> <small>Dax THOMAS (Meiji Gakuin University). Using the USAS Semantic Tagset To Explore Persuasive Language in Jeremy Taylor's Holy Living and Holy Dying, 1650-1651 Tatsuya ISHII (Kobe City College of Technology), Takeshi KAWAMOTO (Hiroshima University). Patterns Related To the Functions of Moves/Steps in the Introductions of Experimental Medical Research Articles: Combining Move Analysis and N-Grams Approach Laurence NEWBERY-PAYTON (Tokyo University of Foreign Studies), Sho FUKUDA (Tohama University), Keiko MOCHIZUKI (Tokyo University of Foreign Studies). Learnability of English Verb-Particle Combinations and the Effect of Linguistically Motivated Instruction 神澤 克徳 (京都工芸繊維大学)、小林 雄一郎 (日本大学)、田中 悠介 (京都大学) . 大学生を対象とする英語スピーキングテストの回答音声に基づくコーパス構築</small>
13:30-13:55	<b>Presentation Q&amp;A session 2 (JAECS members)</b> <small>Akira Moriya and Yoko Iyeyri (Kyoto University). The Be/Have-Perfect in 19th-Century American English: A Corpus-Based Analysis of Some Missionary Documents in Hawaii Risako AZEMOTO (Kyushu University). The Learners' Development of Using Polysemous Words: The Case of 'Over' 仁科 恭徳 (神戸学院大学) . ムーブと形容詞の振る舞いから見た航空会社プロフィールのディスコース分析 投野由紀夫 (東京外国語大学) ・川原田将之 (東京工業大学) ・渡辺亮嗣 (ネットアドバンス) ・星野守 (ネットアドバンス) ・奥村学 (東京工業大学) . CEFR-J 準拠英語教育用 web コーパスの開発</small>
13:55-14:30	Break

Time	Event
14:30-15:30	<b>Presentation Q&amp;A session 3 (JAECs non-members)</b> Elen LE FOLL (Osnabrück University). Issues in Compiling and Exploiting Textbook Corpora Lukasz GRABOWSKI (University of Opole). Phrase Frames as an Exploratory Tool for Studying Translation Patterns: A Corpus-Based Descriptive Study Niall CURRY (Coventry University), Robbie LOVE (Aston University), Olivia GOODMAN (Cambridge University Press). Investigating Publisher Application of Corpus Research on Recent Language Change To ELT Coursebook Development Jamie WILLIAMS (Nottingham Trent University), David WRIGHT (Nottingham Trent University). Pronominal Ambiguity and Ascriptions of Responsibility in the UK Daily Coronavirus Briefings Ibrahim BASHIR (Jubail Industrial College), Kamariah YUNUS (Universiti Sultan Zainal Abidin). A Corpus Analysis of Prepositional Colligations in Nigerian Legal Discourse Barrios LEYRE (Universitat de Lleida), Vázquez GLÓRIA (Universitat de Lleida). Factuality and Conditional Sentences With Indicative Mode: A Corpus-Based Study Rosana VILLARES (University of Zaragoza). Corpus Linguistics Tools for the Creation of Linguistic Resources That Support the Internationalisation of Tertiary Education Hakan CANGIR (Ankara University), Taner CAN (TED University). Speaking of Extinction: A Comparative Corpus-assisted Analysis of the Environmental Framing in Climate Fictions and the News on the Net Yaogang CHEN (The University of Edinburgh). Exploring the Use of Hedges in Academic Writing: A Corpus-Based Analysis Between Chinese TESOL Students and Expert Writers Xiaojing CHEN (The University of Edinburgh). A Comparative Study of Transition Markers in Thesis Abstracts From Chinese Undergraduates and RA Abstracts From Prestigious Journals Mat RAWSTHORNE (The University of Nottingham). Shared Experience: From I-Illness To We-Illness? Narrative Informed Corpus Linguistic Analysis of a Moderated Online Mood Disorders Forum Melissa KEMBLE (The University of Sydney). As Good as the Men? A Corpus-Based Analysis of Media Representations of Athletes Competing in the New Women's Australian Rules Football League
15:30-16:00	Break
16:00-17:00	General Assembly
17:00-18:30	Welcome Reception (online)

## October 4 (Sun), 2020

### Day 2 Schedule

Time	Event
08:00-08:45	Registration/Setup/Getting to know each other
08:45-09:00	Housekeeping/Notices
09:00-10:00	Plenary 2: Sowmya Vajjala (National Research Council Canada). NLP Beyond NLPers – the many faces of NLP in academia and real-world
10:00-11:00	Site exploration/Video viewing/Meeting others
11:00-12:00	<b>Presentation Q&amp;A session 4 (JAECs non-members)</b> Martin SCHWEINBERGER (The University of Queensland). A Corpus-Based Analysis of Ongoing Change in the Adjective Amplifier Systems of Hong Kong and Philippine English Marine Laisa MATTE (Univates University), Larissa GOULART (Northern Arizona University), Simone SARMENTO (Federal University of Rio Grande do Sul), Rozane Rodrigues REBECHI (Federal University of Rio Grande do Sul). Becoming a President: A Diachronic Study on the Language of Brazilian President Jair Bolsonaro Yating YU (The Hong Kong Polytechnic University). Media Representations of 'Leftover Women' in China: A Corpus-Assisted Critical Discourse Analysis Michael HENSHAW (Hokkaido University). A Novel Approach To the ESP Keyword List: 2815 Entries With Frequent Lexical Bundles for Data-Driven Learning Will LINGLE (University of Aizu). Contrasting Narratives: The Greek Financial Crisis in Newspaper Editorials Wilfred Gabriel A. GAPAS (University of Santo Tomas), Rachele BALLESTEROS-LINTAO (University of Santo Tomas). The Discursive News Values of the 2017 Marawi City Crisis: A Corpus-Assisted Multimodal Discourse Analysis of Selected Newspaper Reports Ashleigh COX and Eric FRIGNAL (Georgia State University). Comparing Measures of Directness in Corpora of Essays Written by Iraqi EFL Learners, Native English-Speakers, and Advanced ESL College Students Andrew SCHNEIDER (Embry-Riddle Aeronautical University), Rachele UDELL (Georgia State University), Eric FRIGNAL (Georgia State University). Sky High: Building a Corpus of English for Flight Training Muchamad Sholakhuddin AI FAJRI (Universitas Airlangga). The Construction of Coronavirus in English-Language Indonesian Newspapers: A Corpus-Assisted Discourse Analysis 寺田 里紗(東京外国語大学) . NICT JLE コーパスを用いた日本人英語学習者のイラスト描写における習得レベル別特徴分析

Time	Event
12:00-13:00	Lunch break
13:00-13:25	<b>Presentation Q&amp;A session 5 (JAECS members)</b> Reiko IKEO (Senshu University), Masayuki NAKAO (Tottori University), Eri SHIGEMATSU (Hiroshima University). A Corpus Stylistic Comparison of Speech Presentation of 21st-Century Present-Tense Fiction and 20th-Century Past-Tense Fiction Wei-Tung WANG (Tokyo University of Foreign Studies) and Yukio TONO (Tokyo University of Foreign Studies). Performance Evaluation of Automated CEFR Level Classification Tools Laurence ANTHONY (Waseda University), Natalie FINLAYSON, Emma MARSDEN, Rachel HAWKES, and Nick AVERY (National Centre of Excellence for Language Pedagogy, University of York). Extending Vocabulary Profiling To Languages Other Than English 松田 佑治 (立命館大学). 現代英語における he/she of NP の振る舞い-COCA での調査から-
13:30-13:55	<b>Presentation Q&amp;A session 6 (JAECS members)</b> Hu XIAOLIN (Tokyo University of Foreign Studies). A Validation Study of the Accuracy of Lexical Diversity Tools Satoru UCHIDA, Takehiko SHIMIZU and Saaya KIMURA (Kyushu University). A Corpus Based Approach To Creating an Advanced Wordbook for University Students 山崎 のぞみ (関西外国語大学). 付加疑問との連鎖関係からみた右方転位構造 (テイル) の機能 中谷安男 (法政大学). 経済学国際ジャーナルにおける Method の章のコーパス分析
13:55-14:30	Break
14:30-16:00	<b>Conference symposium/Discussion</b> <b>小・中・高における DDL 普及への挑戦 —DDL ツールの開発, 授業実践, 分野横断的考察—</b> <b>[Meeting the Challenges of Introducing DDL To Pre-Tertiary Students in Japan: Tool Development, Class Instruction and a Cross-Disciplinary Discussion]</b> Part 1: Data-driven learning for younger learners: Current issues, future directions. Peter Crosthwaite (University of Queensland) Part 2: 小・中・高校生のための3種のウェブ DDL ツールと教材:eDDL、hDDL、BES Search 西垣知佳子 (千葉大学)、赤瀬川史郎 (Lago 言語研究所)、石井雄隆 (千葉大学)、神谷昇 (千葉大学)。 西垣が発表 Part 3: 小・中・高校における DDL 実践の分野横断的考察 ① 発見学習から DDL の効果を考える 小山義徳 (千葉大学) ② 小学校英語教育の分野から 物井尚子, 星野由子 (千葉大学) ③ 小学校国語科教育の分野から 安部朋世 (千葉大学) Part 4: DDL の指導と評価 水本篤 (関西大学) Part 5: Communicative English Classroom using DDL in Thailand Pichinart Kumpawan (Surasakmontree School, Thailand)
16:00-16:15	Closing ceremony

# Plenary Talk 1: ‘Corpus Linguistics’ or ‘Linguistics with a Corpus’?

Jesse Egbert

Associate Professor of Applied linguistics (Northern Arizona University)

## Biography

Jesse Egbert is Associate Professor of Applied Linguistics at Northern Arizona University. Jesse specializes in register variation, quantitative methods in linguistics, and corpus linguistic approaches to legal interpretation. He is General Editor of the international peer reviewed journal *Register Studies*, and Technical Strand Editor for the series *Cambridge Elements in Corpus Linguistics*. He has published more than 60 peer-reviewed papers. Recent books include *Register Variation Online* (Cambridge, 2018), *Using Corpus Methods to Triangulate Linguistic Analysis* (Routledge, 2019), and *Doing Linguistics with a Corpus: Methodological Considerations for the Everyday User* (Cambridge, forthcoming).

## Abstract

Corpus linguistics encompasses a vast array of empirical research. It appears that there is only one characteristic that holds true for all corpus linguistic studies: the use of a corpus. I propose that it is not only possible to use a corpus without doing linguistics; it is actually quite common, even in mainstream research publications. So what do I mean by using a corpus without doing linguistics? Linguistics can be defined simply as the scientific study of language. I propose that linguistics requires at least three conditions: (1) linguistically meaningful variables, (2) linguistically valid units of observation, and (3) linguistic description. In the absence of any one of these three conditions, (quantitative) corpus linguistics can quickly become nothing more than pattern hunting and number crunching. In essence, a computer can perform quantitative corpus analysis, but only a linguist can do linguistics with a corpus. Drawing on the research traditions of syntactic complexity, lexical dispersion, and keyword analysis, I will illustrate the importance of these three conditions by comparing studies that satisfy them to studies that don't.

# Plenary Talk 2: NLP Beyond NLPers – the many faces of NLP in academia and real-world

**Sowmya Vajjala**

Data scientist and NLP Research Officer (National Research Council Canada)

## Biography

Sowmya Vajjala currently works as a researcher in Digital Technologies at National Research Council, Canada's largest federal research and development organization. She has worked in the area of Natural Language Processing (NLP) over the past decade in various roles – as a software developer, researcher, educator, and a senior data scientist. Her research interests lie in multilingual computing and educational applications of NLP. She recently co-authored a book: "Practical Natural Language Processing: A Comprehensive Guide to Building Real World NLP Systems", published by O'Reilly Media (June, 2020). She is interested in learning more about the relevance of NLP beyond research both in industry practice as well as in other disciplines, through inter-disciplinary research.

## Abstract

Natural Language Processing is an active area of research and its impact is also seen in many day-to-day applications we use, from generic tools such as email software to specialized ones such as language learning apps. Apart from being an active area of enquiry in itself, NLP methods are widely used in many disciplines, from linguistics to economics, from psychology to plant science. In this talk, I will introduce some common NLP practices, show how NLP differs between academia and industry, and discuss areas where NLP is useful beyond its home turf, including its use in corpus linguistics research. Drawing on my experiences as an NLP researcher and instructor, I will also touch upon what NLP can learn from corpus linguistics and other areas of study, and what is needed to train diverse groups interested in using NLP methods in their work.

# A corpus stylistic comparison of speech presentation of 21st-century present-tense fiction and 20th-century past-tense fiction

IKEO, Reiko (Senshu University, Japan); NAKAO, Masayuki (Tottori University, Japan); SHIGEMATSU, Eri (Hiroshima University, Japan)

**Keywords:** corpus stylistics, direct speech, past-tense fiction, present-tense fiction, speech presentation

## Abstract

In contemporary fiction, the use of the present tense for narration is becoming more common; almost 30% of the titles shortlisted for the Man Booker prize in this century are written in the present tense instead of the past tense. This paper compares direct speech presentation in present-tense fiction of the 21st-century with that in 20th-century past-tense fiction and identifies its stylistic features. This is part of our research project which compares speech, writing and thought presentation in these two textual divisions. The data of this research consists of two corpora. One is a corpus of present-tense narratives (the PREST corpus) and the other is of past-tense narratives (the PAST corpus). The total word count of the PREST corpus is 87,901 while the PAST is 88,319. Both the corpora were annotated using two annotation systems: an online automatic part-of-speech tagging device Sketch Engine (Kilgarriff et al. 2014) and an XML-conformant mark-up system, categorizing speech, writing and thought presentation, which derives from the model developed by Semino and Short (2004).

Speech presentation in the PREST and the PAST corpora respectively accounts for 45.6% and 48.2% of all the tags in each corpus, which suggests that almost half of the texts depict characters' speech. In both corpora, the direct forms account for approximately 29% of all the tags while indirect forms approximately 6%. However, an examination of the sub-corpus of direct speech in each corpus reveals that speech presentation in present-tense narrative and that in past-tense narrative are qualitatively different. In direct speech in present-tense narrative, present-tense verbs are overused while past-tense verbs are underused. The frequent use of present progressives in characters' direct speech seems to reflect their concerns about the immediate present. In addition, characters' speech styles seem to be more interactive, more responsive to their addressees.

## References

- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1(1): 7-36.
- Semino, E. and Short, M. (2004). *Corpus Stylistics: Speech, Writing and Thought Presentation in a Corpus of English Writing*. London: Routledge.

# Learnability of English Verb-Particle Combinations and the Effect of Linguistically Motivated Instruction

NEWBERY-PAYTON, Laurence (Tokyo University of Foreign Studies, Japan); FUKUDA, Sho (The University of Toyama, Japan ); MOCHIZUKI, Keiko (Tokyo University of Foreign Studies, Japan)

Keywords: phrasal verb, verb-particle combination, learner corpus, SLA

## Abstract

Verb-particle combinations (VPCs) have been shown to present difficulties for L2 learners (Liao & Fukuya 2004, Yoshitomi 2006, Siyanova & Schmitt 2007). Corpus studies have revealed that the difficulty of particular VPCs (Negishi, Tono & Fujita 2012) and their over- or under-use (Uchida 2012, Iio 2013, Ishii 2018) differ between L1 Japanese learners and native speakers or other learners. Acquisition studies (Yasuda 2010, Spring 2018, Nakagawa 2019) have assumed, explicitly or implicitly, that Japanese learners struggle to acquire VPCs since Japanese lacks such forms. Researchers have not attempted to explain VPCs with reference to forms in Japanese. This study examines the validity of such assumptions and the potential benefits of linguistically motivated instruction for Japanese learners. The research questions are as follows: 1. Is linguistically motivated instruction beneficial for VPC acquisition? 2. Does instruction making explicit reference to learners' L1 have particular benefits for VPC acquisition? Three sets of audio-visual resources were created to explain VPCs with a. insights from cognitive linguistics; b. insights from contrastive linguistics; c. a mix of the two. Four groups of 1st year non-English majors participated in the study. Three groups received one of the three audio-visual resources, while the fourth served as a control group. All groups then memorized a list of VPCs. Results are reported for pre-, post- and delayed post-test scores. The group using the "mixed" resources showed the only significant increase in average score between pre- and delayed post-test results, suggesting the resources were effective at aiding VPC acquisition. This increase in average score was observed for 80% of learners, suggesting the wide applicability of the resources. Furthermore, the increase in scores was not restricted to items explicitly covered in the resources, suggesting learners became better able to infer the meanings of VPCs more generally.

## References

- Iio, Y. (2013). A corpus-based study of Japanese EFL learners' use of English phrasal verbs. *Kumamoto University Studies in Social and Cultural Sciences*, 11, 35–53. (in Japanese)
- Ishii, Y. (2018). An analysis of phrasal verbs used by Japanese EFL learners: Based on spoken learner corpora and authorized English textbooks. *Learner Corpus Studies in Asia and the World*, 3, 101–119. (in Japanese)
- Liao, Y., & Fukuya, J. (2004). Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning*, 54(2), 193–226.
- 中川右也 (2019). 「句動詞における teachability と learnability の検証 イメージ・スキーマを用いた認知言語学的アプローチ」白畑智彦・須田孝司 (編) 『言語習得研究の応用可能性—理論から指導・脳科学へ—』 169–207.
- Negishi, M., Tono, Y., & Fujita, Y. (2012). A validation study of the CEFR levels of phrasal verbs in the English Vocabulary Profile. *English Profile Journal*, 3(1), 1–16.
- Siyanova, A., & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word verbs. *International Review of Applied Linguistics in Language Teaching*, 45, 119–139.
- Spring, R. (2018). Teaching phrasal verbs more efficiently: Using corpus studies and cognitive linguistics to create a particle list. *Advances in language and literary studies*, 9(5) 121–135.
- Uchida, T. (2012). Use of multiword verbs by non-advanced EFL learners: Focusing on common verb + particle combinations. 東京外国語大学「コーパスに基づく言語学教育研究拠点(CbLLE)」『コーパスに基づく言語学教育研究報告』, 8, 303–323.
- Yasuda, S. (2010). Learning phrasal verbs through conceptual metaphors: A case of Japanese EFL learners. *TESOL Quarterly*, 44(2), 250–273.
- Yoshitomi, A. (2006). The use of phrasal verbs by Japanese learners of English – Implications from story telling data –. In Yoshitomi, A., Umino, T., & Negishi, M. (Eds.), *Readings in second language pedagogy and second language acquisition In Japanese context* (pp 201–225). John Benjamins.

# Patterns Related To the Functions of Moves/Steps in the Introductions of Experimental Medical Research Articles: Combining Move Analysis and N-Grams Approach

Tatsuya Ishii (Kobe City College of Technology, Japan) Takeshi Kawamoto (Hiroshima University, Japan)

Keywords: move analysis, n-grams approach, keyword analysis, experimental medical research articles, introductions

## Abstract

The aim of this presentation is to describe the patterns of prototypical three-grams to five-grams in the Introductions of experimental medical research articles (RAs) with the IMRD structure. Saber (2012), a pioneering study in the field of medical RAs, combined move analysis, and the n-grams approach. He built four subcorpora of moves based on IMRD sections to produce keyword lists. Moreover, he reported three-grams to five-grams with the keywords and identified the function of steps in each move. In this study, however, we generated 12 subcorpora based on the move analysis of IMRD to produce n-grams with stronger relevance to the functions of steps in each move. We collected 300 experimental medical RAs from 30 journals published in 2014 (approximately 1.5 million words in total). Three moves that Nwogu (1997) described were found in Introductions: (Move1) Presenting the Background Information (40,412 words), (Move2) Reviewing the Related Research (80,875 words), and (Move3) Presenting the New Research (32,025 words). To identify the keywords in the three moves, we used CasualConc (Imao 2019) and calculated the scores of the Log-Likelihood Ratio. CasualConc produced three-grams to five-grams with the first keyword “is”, “been”, and “we” in Move 1, Move 2, and Move 3, respectively. We categorized these patterns into steps strongly associated with the functions of each move. The results of these patterns with the keyword show that Move 1 contains five patterns as Step 1 (established knowledge) and one pattern as Step 2 (problems). Move 2 includes three patterns as Step 1 (previous research) and two patterns as Step 2 (limitation). Move 3 involves three patterns as Step 1 (purpose), one pattern as Step 2 (procedure), and another single pattern as Step 3 (outcomes). We propose that corpora generation based on move analysis would lead to better pedagogically useful lists of academic phrases.

# 大学生を対象とする英語スピーキングテストの回答音声に基づくコーパス構築

神澤克徳（京都工芸繊維大学）、小林雄一郎（日本大学）、田中悠介（京都大学）

Keywords: コーパス構築、学習者話し言葉コーパス、英語スピーキングテスト、大学生

## Abstract

本発表の目的は、発表者らが現在構築中のコーパスである KIT Speaking Test Corpus の概要を述べ、それを使用した分析例を紹介することである。KIT Speaking Test Corpus は、京都工芸繊維大学で1年次生全員を対象に実施した英語スピーキングテスト (KIT Speaking Test) の回答音声を書き起こしたものである。書き起こしの対象となる回答音声は74時間16分（1名あたり7分45秒×575名）である。書き起こしデータには「フィラー」や「自己訂正」など17種類のタグを挿入している。また、ヘッダー情報として、受験者の属性のほか、スピーキングテストのスコアや直前に受験したTOEICのスコアを付与している。このような仕様によって、学習者の習熟度（テストスコア）とパフォーマンスの関係、スピーキングテストのタスクが受験者のパフォーマンスに与える影響などの分析が可能となり、さまざまな角度から日本語を母語とする大学生の英語スピーキング能力の実情を解明することができると期待される。本発表では、4段階の習熟度別の産出語彙量、特徴的な語句を報告する。具体的には、3人称代名詞、副詞、時制、接続表現、前置詞句、that節などの発達パターンに注目する。また、フィラーや自己訂正の頻度と生起位置に注目することで、学習者が言いどみやすい箇所を光を当てる。

# ムーブと形容詞の振る舞いから見た航空会社プロフィールのディスコース分析

仁科恭徳（神戸学院大学、日本）

Keywords: ディスコース分析、形容詞、ムーブ

## Abstract

本発表では、タグ付けしたDIYコーパスを用いて量・質の観点から航空会社のプロフィールを分析した結果を報告する。特に、当該プロフィールのムーブ（構造）と形容詞の振る舞いに焦点を置き、各航空会社が属している3つのアライアンス（スターアライアンス、ワンワールド、スカイチーム）間での量的・質的な違いや、航空会社プロフィール全般に通底している共通性に関して調査した結果を発表する。

ムーブに関しては、ムーブの種類、各ムーブの重要性、典型的なムーブ構造に関して検証する。形容詞の振る舞いに関しては、形容詞＋名詞の連辞的結合(colligation)や best＋AWARD といった優先的意味選択(semantic preference)に注目して分析を進める。

ビジネス・コミュニティーにおいて、プロフィールは代表的なジャンルの一つである。よって、このジャンルにおいて共有され慣習化された言語実態を暴くことで、ビジネス・コミュニティーに通底している知識や文化の理解を深めるだけでなく、English for Business Purposes(ビジネスに特化した英語(教育))における教材開発やシラバスデザインの元資料としても有効に活用することができる。

本研究におけるリサーチ・クエスションは以下のとおりである。

1. コーパス・データに基づく各種プロフィールの基本的な異なり
2. 当該プロフィールで使用されているムーブの特定（種類と数）
3. 当該プロフィールにおける各ムーブの必要性の数値化
4. 当該プロフィールにおける典型的なムーブ構造の解明
5. アライアンス間における形容詞の振る舞いの類似性と相違性

1~5のリサーチ・クエスションをもって、当該プロフィールの実態を精緻に暴く。1は当該プロフィールと他のプロフィールとの違いの可視化、2~4は当該プロフィールが担う伝達機能の解明、5はよりマイクロな視点から分析した当該プロフィールの特徴の解明を目的とする。例えば、リサーチ・クエスション2に関しては18種のムーブが認められ、リサーチ・クエスション3に関しては2種のムーブが義務的、3種のムーブが慣習的に使われていたが、発表時には質的な考察も含めて詳細を報告したい。

# The Be/have-perfect in 19th-century American English: A Corpus-based Analysis of Some Missionary Documents in Hawaii

MORIYA, Akira (Kyoto University, Japan); IYEIRI, Yoko, (Kyoto University, Japan)

Keywords: historical linguistics, English in Hawaii, perfect auxiliaries, Late Modern English

## Abstract

The present study discusses the be- and have-perfect in the ABCFM Hawaii Corpus (the Hawaii Corpus), a corpus of approximately 653,100 words compiled by our research team, using mainly journals written by eight members of the American Board of Commissioners for Foreign Missions (ABCFM), who migrated to Hawaii in the 19th century (cf. Iyeiri & Fukunaga, 2020, for some details of the corpus).

Both be- and have-auxiliaries are known to have been used in the perfect construction of mutative intransitive verbs such as come and go since the Old English period (e.g. ‘they are gone’ (be-perfect) and ‘they have gone’ (have-perfect), cf. Denison, 1998). In the Late Modern English period, however, the have-perfect began to replace the be-perfect, and became dominant as in contemporary English (Rydén & Brorström, 1987). This study aims to investigate the uses of perfect forms in the Hawaii Corpus, to compare them with the general trend in 19th-century English and to account for the characteristics of the be/have-perfect usage in different texts included in the Hawaii Corpus. Kytö (1997) suggests that American English (AmE) experienced an earlier increase in the rate of the have-perfect usage compared to British English, and the analysis of the Hawaii Corpus reveals indeed that the be-perfect is already quite rare, except for a few verbs such as improve and change, confirming the general trend in AmE. Other findings in the Hawaii Corpus include: the verb improve, showing a relatively frequent use of the be-perfect, prominently appears in certain contexts regarding health (e.g. her health is very much improved); the be-perfect is less frequent in journals than other text genres; and there are considerable differences in the preference of perfect auxiliaries even between a married couple.

## References

- Denison, D. (1998). Syntax. In S. Romaine (Ed.), *The Cambridge history of the English language: Volume IV 1776-1997* (pp. 92-329). Cambridge University Press.
- Iyeiri, Y. & Fukunaga, M. (2020, May 20-23). A Corpus-based analysis of negation in some 19th-century American missionary documents in Honolulu. . ICAME 41, Heidelberg, Germany.
- Iyeiri, Y., Fukunaga, M. & Moriya, A. (2019). ABCFM Hawaii Corpus. Ver. 1.2.
- Kytö, M. (1997). Be/have + past participle: The choice of the auxiliary with intransitive from Late Middle to Modern English. In M. Rissanen, M. Kytö & K. Heikkonen (Eds.), *English in transition: Corpus-based studies in linguistic variation and genre styles* (pp. 17-85). Walter de Gruyter.
- Rydén, M. & Brorström, S. (1987). *The be/have variation with intransitives in English: With special reference to the Late Modern Period*. Almqvist & Wiksell International.

# The Learners' Development of Using Polysemous Words: The Case of 'Over'

AZEMOTO, Risako (Kyushu University, Japan)

Keywords: second language acquisition, polysemous words, learner corpus

## Abstract

Acquiring multiple senses of a polysemous word perfectly is difficult for language learners. Although some researchers have tried to find the effective way of teaching a polysemous word, few studies have examined how learners develop their knowledge of it. Therefore, this presentation aims to show insights of how Japanese learners of English develop their use of a polysemous word "over". To achieve this purpose, this study set two research questions;

RQ1: Do Japanese learners of English use "over" the same way as native speakers do?

RQ2: What is the difference between "easy" senses and "difficult" senses?

To answer these two research questions, both quantitative and qualitative research was undertaken. First, a learner corpus, Longman Learners' Corpus, and a native corpus, British National Corpus, were used to collect sufficient examples of "over" from each proficiency level of the learners and English native speakers. After that, semantic tags of "over" from Tyler and Evans (2001) were used to annotate each example. Then, a correspondence analysis was generated based on the cross-tabulation which organizes the results of the annotation. Finally, the difference between easy senses and difficult senses for Japanese language learners was discussed based on the result of CA and the frequency of each sense.

Correspondence analysis shows clearly that Japanese language learners use "over" differently from English native speakers in terms of its senses. There are some senses which have a strong connection with the learners, such as covering, completion, and repetition senses, while control, focus-of-attention, transfer, preference, the-other-side-of senses have a strong connection with native speakers. In addition, by focusing on the senses which are no occurrence at the beginner level and which are produced little even at the advanced level, this study has succeeded in explaining which senses are difficult for language learners to produce. This study revealed that the protoscene and the senses which are in cluster 4, 5, and 6 in the polysemy network of over by Tyler and Evans (2001) are especially difficult for Japanese language learners. These findings might help teachers when they teach the word "over" to Japanese language learners of English, and help researchers when they assess the learners' production of English.

## Reference

Tyler, A., & Evans, V. (2001). Reconsidering prepositional polysemy networks: The case of over. *Language*, 724-765.

# CEFR-J 準拠英語教育用 web コーパスの開発

投野由紀夫（東京外国語大学）；川原田将之（東京工業大学）；渡辺亮嗣（ネットアドバンス）；星野守（ネットアドバンス）；奥村学（東京工業大学）

Keywords: CEFR, web corpus, pedagogical corpus

## Abstract

本研究は、ヨーロッパ言語共通参照枠（CEFR）を日本の英語教育に適用した CEFR-J（投野，2013）を用いて、英語教育を側面から支援する教育用コーパスを構築するプロジェクトの中間報告である。

手順としては、まず基本分野の小中高生向けの英語サイトを分野別に調査閲覧し、内容的に適切と思われるサイトの URL リストを作成した。次いで、URL を手がかりに web 上から HTML ファイルを自動取得し、収集した HTML ファイルからヘッダやフッタ等の不要な情報を取り除くため、ブロック単位と BootCat の 2 方法でテキストを抽出した。センテンス数やタグ情報などの指標をもとに、これらのファイルをスクリーニングした。この時点で、4 億 5500 万語（586,660 ファイル）のテキスト・データが取得できた。

次にこれらのテキストを、CEFR-J プロジェクトの資源を用いて東京工業大学奥村研究室で作成した CEFR レベル判定ツール（English Level Checker）を用いて、CEFR レベル（A1～C2）に自動分類した。この判定ツールは、テキスト中の CEFR-J Wordlist に基づく使用語彙レベル、主要文法事項の出現率、などをもとに機械学習されたものである。CEFR レベル判定ツールの評価を、A1～B2 各レベル 20 サンプルを抽出し 2 名の訓練された評価者により人手で CEFR 評価を行い（Cohen's Kappa=0.72）、機械判定と比較した。結果は A1-A2, A2-B1 などの隣接レベルでの誤判定がまだ多かったが、大きく A レベルと B レベルの 2 クラス判定では 70% 程度の一致度であった。

最後に小学館コーパスネットワーク(<https://scnweb.jpanknowledge.com/>)へ搭載され、新学習指導要領に基づく CAN-DO ベースの英語教育を側面から支援する言語教育資源としての意義、CEFR-J レベルの細分化への対応など、将来的な課題にも言及する。

## 参考文献

投野由紀夫（編著）（2013）『CAN-DO リスト作成・活用 英語到達度指標 CEFR-J ガイドブック』大修館書店。

# **A Comparative Study of Transition Markers in Thesis Abstracts From Chinese Undergraduates and RA Abstracts From Prestigious Journals**

**CHEN, Xiaojing, the University of Edinburgh**

**Keywords: genre analysis, metadiscourse, transition markers, thesis abstracts, research article abstracts**

## **Abstract**

The researcher conducts a comparative study between thirty thesis abstracts written by Chinese undergraduates and thirty research article abstracts collected from top-ranked journals in applied linguistics. Abstracts become an independent genre which can reflect language proficiency and critical thinking of the researchers. This study focus on the frequency of using transitions (e.g. “but” and “therefore”), which are the metadiscourse markers particularly used to structure the text by explicit connection between the main clauses. Five types of transition markers are identified and log-likelihood scores are used to investigate whether there are significant differences between the two kinds of abstracts. Subsequently, two interviews are also conducted to explore the in-depth reasons for the LL scores. The overall frequency of using transitions is not significantly different. However, according to the interviews, Chinese undergraduates sometimes overuse the transitions because it is easier to structure their articles which reveals their abundant vocabulary but deficiencies in text organization. The low frequency in the use of comparison may be attributed to the lack of questioning competence in examination-oriented education. Based on the results, the researcher believes that the focus for Chinese EFL learners in academic writing should be shifted from grammatical knowledge to research logic and critical thinking.

# A Corpus Analysis of Prepositional Colligations in Nigerian Legal Discourse

Ibrahim Bashir, Jubail Industrial College, Saudi Arabia and Kamariah Yunus, Universiti Sultan Zainal Abidin, Malaysia

**Keywords:** British National Corpus (BNC), corpus linguistics, colligation, prepositional colligations, legal phraseology, Nigerian law corpus (NLC)

## Abstract

English is recognised as a national lingua franca of Nigeria. It is adopted as a medium of instruction at almost all levels of the Nigerian educational system, in bureaucracy, the main mass media, national politics, science and technology, corporate companies, two legislative chambers, and the judiciary. The empirical studies using corpus approaches had shown that prepositions are essential linguistic features of legal phraseology. That area was under explored in the Nigerian context. As Nigeria is using English in legal affairs, the present study purports to investigate the representativeness and essentiality of prepositional colligations in Nigerian legal texts. Using a corpus methodology, a new corpus was compiled namely Nigerian Law Corpus (NLC) which contained 546313 word-tokens. The list of prepositional colligations generated from NLC were compared with two reference corpora – British National Corpus of Law (BNCL – 2.2 million word-tokens), and British National Corpus of General Written English (BNCW – 1 million word-tokens) explored using Lextutor (v.8.3). The analysis was aided with the AntConc (v.3.5.7), Lextutor (v.8.3), and Rayson's Log-likelihood Calculator. This paper reports the quantitative parts of a large study. The findings revealed that prepositional colligations were prevalent linguistic features in Nigerian legal discourse. The findings showed that prepositional colligations were under-represented in the NLC compared with BNCL (NLC observed frequency = 3619, relative frequency = 0.66, BNC observed frequency = 17923, relative frequency = 0.81, Over-/under-representation (-), Log likelihood = 134.82). On the other hand, prepositional colligations were over-represented in NLC compared with BNCW (NLC observed frequency = 3619, relative frequency = 0.66, BNCW = 3939, relative frequency = 0.39, Over-/under-representation (+), Log likelihood = 500.43). The study concludes that prepositional colligations were important linguistic elements that formed the phraseological profiles of Nigerian legal discourse due to their prevalent distributions and multifarious semantic and pragmatic functions.

# Corpus Linguistics tools for the creation of linguistic resources that support the internationalisation of tertiary education

VILLARES, Rosana (University of Zaragoza, Spain)

Keywords: corpus linguistics, teaching guides, glossary, translation, internationalisation

## Abstract

In the last decades, universities have developed different internationalisation strategies based on the English language, such as English-medium instruction courses or the translation of institutional websites, with the purpose of improving their international visibility and attracting international students (Dearden, 2014; Ferguson, 2007). However, the literature (e.g. Ferguson et al., 2011; Pérez-Llantada et al. 2011) has consistently reported the multiple difficulties faced by non-English native scholars when using English for their academic tasks. As part of the internationalisation strategy of a medium-sized Spanish university, lecturers must translate to English their modules' teaching guides. Given this situation, this study proposes the use of corpus linguistics tools to create a series of support resources that facilitate the translation of those documents. A corpus of 113 institutional reports written by the university and verified by the Spanish National Agency for Quality and Accreditation. These documents were gathered because they included all the information that later on the lecturers have to use in their teaching guides of both undergraduate and graduate degrees. In this way it was possible to familiarise with the official terminology that lecturers should use. The AntConc v3.5.7 software was used to extract frequency lists and concordances to identify the lecturers' recurrent terminology and phraseology. Based on the quantitative results, three resources were created. Firstly, a glossary of the teaching activities with their translation and definition so that lecturers could use them accurately. Secondly, a glossary with the most frequently used expressions written in Spanish, their equivalent in English, and a concordance line that illustrates the use of that specific term in context. Lastly, several teaching guide templates were created to homogenise the teaching guide genre while simultaneously allowing variation according to disciplinary differences. By sharing these resources with the lecturers, it was expected that some of the translation challenges they may face (terminology inconsistencies, literal translation, target readership) (Hurtado, 2001; Parra Galiano, 2005) would be overcome.

## References

- Dearden, J. (2014). English as a medium of instruction– a growing global phenomenon. British Council.
- Ferguson, G. (2007). The global spread of English, scientific communication and ESP: questions of equity, access and domain loss. *Ibérica*, 13, 7-38.
- Ferguson, G.; Pérez-Llantada, C. and Plo, R. (2011). English as an international language of scientific publication: a study of attitudes. *World Englishes*, 30(1), 41–59.
- Hurtado Albir, A. (2001). Traducción y traductología: introducción a la traductología. Madrid: Cátedra.
- Parra Galiano, S. (2005). La revisión de traducciones en la Traductología: aproximación a la práctica de la revisión en el ámbito profesional mediante el estudio de casos y propuestas de investigación (Doctoral thesis). Granada: Universidad de Granada.
- Pérez-Llantada, C.; Plo, R. and G. Ferguson. (2011). “You don't say what you know, only what you can”: The perceptions and practices of senior Spanish academics regarding research dissemination in English. *English for Specific Purposes*, 30, 18–30.

# Exploring the Use of Hedges in Academic Writing: A Corpus-Based Analysis Between Chinese TESOL Students and Expert Writers

Chen, Yagang (University of Edinburgh, United Kingdom)

Keywords: hedges, academic writing, Chinese TESOL postgraduates, published research articles

## Abstract

The ability to write academic papers in English has been acknowledged as an important criteria to evaluate academic profession around the world (Tang, 2012). In academic writing, appropriate use of hedges is critical. However, the ability to express doubt and certainty appropriately in English is generally acknowledged as a difficult task for language learners. The present study investigates how Chinese TESOL postgraduates employ lexical hedges in their academic writing in comparison to professional writers, attempting to help improve their writing skills. Using Hyland (1998a)'s taxonomy, a quantitative corpus-based study was conducted to examine hedges based on grammatical class and function. The findings demonstrate that : 1) Chinese TESOL postgraduates use hedges extensively in their academic writing. 2) Chinese TESOL students display similarities in overall frequency of hedges and choice of some hedge items. 3) Both groups rely on content-oriented hedges more than reader-oriented hedges. Differences are also detected. Chinese TESOL students employ significantly more hedges than expert writers, particularly with modal verbs and adverbial hedges. Different preferences of individual hedge items by the two groups are also presented. Pedagogical implications generated from the main findings are concerned with material design and teaching instructions for EAP and academic writing courses at tertiary level.

# **Factuality and conditional sentences with Indicative mood in subordinate clauses: a corpus-based study**

**Barrios, Leyre (Universitat de Lleida, Spain); Vázquez, Glòria (Universitat de Lleida, Spain)**

**Keywords:** factuality, if-conditionals, Indicative, Spanish, annotation.

## **Abstract**

Factuality (or certainty) is the notion that refers to the speaker's commitment to the truthfulness of a situation, which is expressed through different linguistic markers. This study belongs to TAGFACT project whose aim is to analyze and automate the interpretation of factuality in Spanish in journalistic texts.

A conditional sentence consists of two clauses, a main clause (apodosis) and a subordinate clause (protasis), which are linked by a conjunction typically used for hypothetical situations (Quirk et al., 1985; Sweetser 1990, Dancygier 1993; Montolío, 1999; RAE 2009).

One of our aims is to analyze the factuality of both clauses of some conditional sentences in Spanish, particularly those conditionals with protasis with Indicative and introduced by the conjunction *si*, which in English is *if*. Also, we pretend to propose linguistic clues for the automatic annotation of factuality for both clauses.

The protasis of a conditional sentence typically do not describe factual situations, but this is just in some types of conditionals. In addition, the factuality of apodosis has hardly been studied, which is very important, as we will see later.

For Spanish, there are studies that analyze the interpretation of the protasis of these sentences in terms of factuality, but not of the apodosis. Different bipartite (Veiga and Mosteiro, 2006) or tripartite (Montolío 1999) factual classifications have been made. Our study is based on the latter, since it goes deeper into formal issues and distinguishes between real, potential and unreal conditionals from the point of view of the protasis.

we have analyzed 241 if-conditionals from 3 different corpus: the corpus Now (Mark Davies, 2019), CORPES XXI (Real Academia de la lengua Española, 2020) and CREA (Real Academia de la lengua Española, 2015) whose protasis include indicative tenses: presente de Indicativo, pretérito imperfecto de Indicativo, pretérito perfecto simple and pretérito perfecto compuesto. In total, 25 temporal combinations have been analyzed.

Firstly, the interpretation of the factuality of the protasis has been revised, extending in some cases the current bibliographical proposal, and secondly, a factual interpretation has been proposed for the apodosis.

The result of this study is the proposal of a series of linguistic rules to formalize the interpretation of the factuality of the two clauses of the conditional sentences that present protasis with indicative in Spanish, thus completing the previous study done on the subjunctive (Barrios and Vázquez, 2020).

## **Bibliography**

Barrios, L. and Vázquez, G. (2020). Factuality and conditional sentences with Subjunctive mood in subordinate clauses: a corpus-based study, in Congreso Internacional de Lingüística y de corpus (21-23 October, 2020), Colombia: Antioquía.

Dancygier, B. (1993): "Interpreting conditionals. Time, knowledge and causation", *Journal of Pragmatics*, 19, pp. 403-434.

Davis, Mark (2018). Now. Corpus del español. Disponible en: <https://www.corpusdelespanol.org/>

Montolío, E. (1999) "Las construcciones condicionales", in I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (vol. 3, pp. 3805-3878). Madrid: Espasa-Calpe.

Rodríguez Rosique, S. (2008). *Pragmática y Gramática: condicionales concesivas en español*. In *Studien Zur Romanischen Sprachwissenschaft Und Interkulturellen Kommunikation* (Vol. 47). Frankfurt am Main: Peter Lang.

Schwenter, S. (2001): Expectations and (in)sufficiency: Spanish como conditionals. *Linguistics* 39-4, pp. 733-760

Sweetser, E. (1990). *From etymology to pragmatics. Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge University Press / Peking University Press.

Veiga, A. y Mosteiro, M. (2006). *El modo verbal en cláusulas condicionales, causales, consecutivas, concesivas, finales y adverbiales de lugar, tiempo y modo*. Salamanca: Ediciones Universidad de Salamanca.

# Speaking of Extinction: A Comparative Corpus-assisted Analysis of the Environmental Framing in Climate Fictions and the News on the Net

Hakan CANGIR (Ankara University), Taner CAN (TED University)

Keywords: Cli-fi, Semantic Domains, Collocations, Framing, Metaphors

## Abstract

The cognitive linguist George Lakoff developed an alternative view of metaphor which has long been relegated to the context of creative writing. In his view, metaphors are part of cognitive frames we use in our everyday lives to understand abstract concepts in terms of concrete objects and activities. Straddling contemporary cognitive theories and corpus linguistics, the present study seeks to compare the (metaphorical) frames in the climate fictions (aka cli-fi) and the news on the Internet by identifying the key semantic domains, collocations and key words in context. For this comparative analysis, we compiled a corpus of 26 novels by 26 novelists and exploited the Climate Change in the News Corpus provided by Lancsbox. We used the tools Wmatrix for semantic tagging and Lancsbox and Voyant tools to report collocational patterns and key words in context. The theoretical and methodological approach of the study was mainly based on Lakoff's framing theory and ecocriticism. As the initial stage of the research, we designated two key semantic domains, namely "Green Issues" and "Universe," and explored semantic collocations and key words in context in the two corpora. When the two corpora are compared, the preliminary results indicate that cli-fi writers employ a more powerful language (e.g. "ecological terror") to bring up a new perspective on the environmental crisis we are going through. These writers' darker and emotionally disturbing tone is in stark contrast to the neutral, disinterested tone we find in the news corpus. News writers seem to use a more business-oriented language where we find the words with strong political and economic connotations. Subsequently, the preliminary results drawn from the research is partly in line with Lakoff's concerns about the cognitive frames we create about environmental problems, and his emphasis on the role of storytelling in changing these false frames.

# Investigating Publisher Application of Corpus Research on Recent Language Change To ELT Coursebook Development

CURRY, Niall (Coventry University, UK); LOVE, Robbie (Aston University, UK); GOODMAN, Olivia (Cambridge University Press, UK)

Keywords: ELT, materials development, corpus linguistics, adverbs, spoken British English, language change

## Abstract

Corpus linguistics (CL) has come to occupy an important space in the language teaching industry (see e.g. Römer, 2011), but the specific ways in which English Language Teaching (ELT) publishers use CL research to inform materials development are under-studied, meaning that it is not known whether CL is being used by publishers to its full potential. This study investigates the use of CL research by a major international ELT publisher – Cambridge University Press – by (a) conducting a case study into recent change in adverbs in casual spoken British English; (b) sharing the findings with editors from the publisher; and (c) investigating how the editors use the corpus-informed findings in developing coursebooks. We address two main research questions:

1. How has the use of adverbs changed in casual spoken British English between the 1990s and 2010s?
2. What is the role of the editorial process in developing corpus-informed coursebooks?

Through our analysis of the Spoken BNC1994 (BNC Consortium, 2007) and the Spoken BNC2014 (Love et al., 2017), we find evidence of major recent changes in the usage of frequent adverbs. Overall, adverbs are significantly more frequent in the more recent dataset. Furthermore, there is evidence of noteworthy functional and syntactic shifts in a number of adverbs including ‘like’, ‘so’, ‘just’, ‘well’ and ‘literally’, which, we argue, warrant discussion and presence in the English language teaching classroom.

Following the corpus analysis, we conducted in-depth interviews with the editors and a review of the materials they subsequently produced using the corpus findings (Cowan et al., 2018; Kilbey et al., 2018; Goldstein & Jones, 2019). We evaluate how our findings are represented in the materials, finding some evidence of effective use of corpora in materials development but revealing limitations in current corpus research approaches which prevent editors from employing CL research more effectively.

## References

- BNC Consortium. (2007). The British National Corpus, version 3 (BNC XML Edition). Distributed by Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk/>
- Cowan, A., Ward, C., & Ting, T. (2018). Talent Level 2. Cambridge: Cambridge University Press.
- Goldstein, B., & Jones, C. (2019). Evolve Level 6. Cambridge: Cambridge University Press.
- Kilbey, L., Ward, C., & Ting, T. (2018). Talent Level 3. Cambridge: Cambridge University Press.
- Love, R., Dembry, C., Hardie, A., Brezina, V., & McEnery, T. (2017). The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics*, 22(3), 319-344. DOI: 10.1075/ijcl.22.3.02lov
- Römer, U. (2011) ‘Corpus Research Applications in Second Language Teaching’, *Annual Review of Applied Linguistics*, 31, 205-225.

# Issues in Compiling and Exploiting Textbook Corpora

LE FOLL, Elen (Osnabrück University, Germany)

Keywords: Corpus Design, Coursebooks, Pedagogical Materials, English as a Foreign Language, Classroom English

## Abstract

Textbooks are known to be “one of the most important educational inputs” (Pingel 2010: 7) and, as such, have long been cherished objects of research, especially in the social and political sciences. In applied linguistics, page-by-page analysis of textbook language was once a difficult, time-consuming process. However, the development of digital data storage and retrieval enabled Mindt (1987; 1992) to pioneer a new approach to language textbook analysis using computer-readable textbook corpora. Since, multiple studies have highlighted the potential of corpus-based textbook analysis both for materials development and evaluation, and as a means of capturing learner language input and better understanding learner language production (cf. Römer 2004; Meunier & Gouverneur 2009).

However, in practice, compiling textbook corpora can be an arduous task, fraught with potential pitfalls. In this paper, I highlight some of the issues specific to compiling a corpus of textbooks and propose possible solutions. The compiling process begins with the selection of the materials to be included in the corpus, a process driven by the research questions. I discuss questions pertaining to the design of the sampling frame, corpus size, representativeness, and balance (cf. Biber 1993). Due to the many different font types, colours, and often complex page layouts, automatic OCR (optical character recognition) is usually much more complex for textbooks.

For many corpus linguistic measures, the basic unit is the text. However, textbooks are essentially collections of many different texts. Moreover, textbooks also pose problems due to some texts being extremely short (e.g. instructions), whilst others can span several pages (short story). In addition, what constitutes a textbook must also be defined since educational publishers now offer multimedia packages that include coursebooks, workbooks, teacher’s guides, audio and video materials, etc.

These issues in compiling and exploiting textbook corpora are illustrated with examples from the compilation and analysis of a manually annotated textbook corpus of nine series (43 volumes) of secondary school EFL textbooks.

## References

- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243–257.
- Meunier, F., & Gouverneur, C. (2009). New types of corpora for new educational challenges: Collecting, annotating and exploiting a corpus of textbook material. In K. Aijmer (Ed.), *Studies in Corpus Linguistics* (Vol. 33, pp. 179–201). John Benjamins.
- Mindt, D. (1987). *Sprache, Grammatik, Unterrichtsgrammatik: Futurischer Zeitbezug im Englischen I* (1st ed.). Diesterweg.
- Mindt, D. (1992). *Zeitbezug Im Englischen: Eine Didaktische Grammatik Des Englischen Futurs*. Gunter Narr Verlag.
- Römer, U. (2004). Comparing real and ideal language learner input: The use of an EFL textbook corpus in corpus linguistics and language teaching. In G. Aston, S. Bernardini, & D. Stewart (Eds.), *Studies in Corpus Linguistics* (Vol. 17, pp. 151–168). John Benjamins.

# Phrase frames as an exploratory tool for studying translation patterns: a corpus-based descriptive study

Lukasz Grabowski (University of Opole, Poland)

Keywords: phrase frames; phraseology; parallel corpora; English-to-Polish translation; corpus-based study

## Abstract

Designed as a proof-of-concept, this descriptive corpus-based study focuses on the construct of phrase frames, originally defined as a contiguous sequence of  $n$  words identical except for one (Fletcher 2002). Although phrase frames were already used as a means of exploring pattern variability across and within different text types or registers written in English (Römer 2010; Gray & Biber 2013; Fuster-Marquez & Pennock-Speck 2015; Grabowski 2015; Forsyth & Grabowski 2015; Cunningham 2017; Juknevičienė & Grabowski 2018; Lu et al. 2018), there has been no attempt so far to employ that construct as a unit of analysis in descriptive research on translation. More precisely, we aim to verify whether phrase frames found in English source-language texts reveal similar generalizable syntagmatic patterns in Polish translations. Also, we aim to investigate whether the observed translation patterns help highlight lexical and stylistic peculiarities of translations which would otherwise be difficult to capture.

In this study, we use the English-Polish parallel corpus Paralela (Pęzik 2016), notably its European Parliament proceedings sub-corpus (Koehn 2005), to identify and describe translation patterns that emerge from one functionally-defined English phrase frames expressing attitudinal stance (it is \* clear that) used as a starting point for the analysis. The findings provided insights into English-to-Polish translation patterns, which revealed that the Polish equivalents are realized with a high degree of regularity and can be generalized into syntagmatic patterns similar to phrase frames. We also obtained valuable cross-linguistic insights into corresponding syntagmatic structures in English and Polish. All in all, the study design is potentially applicable beyond English-Polish translation.

# Pronominal Ambiguity and Ascriptions of Responsibility in the UK Daily Coronavirus Briefings

WILLIAMS, Jamie (Nottingham Trent University, UK); WRIGHT, David (Nottingham Trent University, UK)

Keywords: discourse, corpus-based discourse analysis, political discourse, COVID, coronavirus

## Abstract

Within political discourse, pronouns have been highlighted as important linguistic features due to their inherent ambiguity, and their roles in creating distance or closeness and accepting or denying responsibility for actions (Fetzer and Bull, 2008; Mulderrig, 2012). These issues are pertinent to the COVID-19 pandemic, as governments attempt to clearly communicate guidance to the general public, as well as describe steps being taken to slow the virus' spread. Within the context of the United Kingdom, one of the worst affected countries globally, we investigated how pronouns were used by governmental speakers to administer responsibility and whether they contributed to reported criticisms of ambiguity in the government's communications (Oliver, 2020).

A corpus of 92 political speeches, totalling 117,779 words, was constructed based upon official transcripts from the UK government's website. Focussing on the use of the first person plural (1PL) pronouns, 3,045 concordance lines were analysed to identify (1) their referent – particularly if they carried an exclusive (we – the government) or inclusive (we – the country) reading, and (2) the transitivity patterns these pronouns act as Participants in (Halliday and Matthiessen, 2014).

We argue that the UK government uses the inherent ambiguity between the exclusive and inclusive readings of this pronoun to mitigate their own portrayed responsibility for controlling the spread of the virus. We argue that they do so through at least two means. Firstly, when using 1PL pronouns in an exclusive manner, although they represent themselves overwhelmingly as Actors, they obscure the precise details about the measures they are taking. Secondly, when using the 1PL pronouns in an inclusive manner, they represent the British public as co-Actors in processes they have no control over and indeed are usually considered to responsibility of the government themselves.

## References

- Fetzer, A. and Bull, P. (2008) 'Well, I answer it by simply inviting you to look at the evidence': The strategic use of pronouns in political interviews. *Journal of Language and Politics*, 7(2): 271–289.
- Halliday, M.A.K. and Matthiessen, C.I.I.M. (2014). *Halliday's Introduction to Functional Grammar*, 4th Edition. London: Routledge.
- Mulderrig, J. (2012) The hegemony of inclusion: A corpus-based critical discourse analysis of deixis in education policy. *Discourse and Society*, 23(6): 701–728.
- Oliver, D. (2020) Covid-19 highlights the need for effective government communications. *BMJ*, 369: m1863, DOI: <https://doi.org/10.1136/bmj.m1863>.

# **Shared Experience: From I-Illness To We-Illness? Narrative Informed Corpus Linguistic Analysis of a Moderated Online Mood Disorders Forum**

**RAWSTHORNE, Mat (University of Nottingham, UK)**

**Keywords:** peer support, corpus linguistics, narrative

## **Abstract**

Analysis of interactions in a supervised online mutual aid platform serving people with stress and low mood. Key questions: Are the forum participants sharing their stories, a common strategy for soliciting social support in online communities and does this sharing process assist in recovery by destabilising their narratives? Methods: A combination of computer-assisted corpus linguistic analyses were employed to examine the structure and nuances of the exchanges, attempting a ‘big qual’ study of broad social processes while also attending to important detail of word use. Results: Although people turn to peer support to help and be helped through sharing stories, there does not seem to be much surveillance for the innovative moments that would shift from being stuck in the ‘Same Old Story’. While platforms like Big White Wall can provide a pause in experience that facilitates separation between the person and the problem, there is no evidence of them stepping back to frame new perspectives to make room for counter narrative. Pennebaker et al (2003) proposed that healing could be detected by shifts from singular to plural personal pronouns, and verb tense changes (Tausczik & Pennebaker, 2010). Closer scrutiny of the interactions points to people talking at each other rather than to or with each other.

# A corpus-based analysis of ongoing change in the adjective amplifier systems of Hong Kong, Indian, and Philippine English

Martin Schweinberger, The University of Queensland, Australia

**Keywords:** Adjective Amplification, Hong Kong English Philippine English, Indian English, Conditional Inference Trees, Variationist Sociolinguistics, Language Change

## Abstract

This study focuses on one of ongoing change in adjective amplifiers (very, really, so, etc.) in Hong Kong (HKE), Indian (IndE), and Philippine English (PhiE) based on data from the International Corpus of English.

While previous research on changes in amplifier systems has successfully applied multivariate methods and unearthed intricate interdependencies and highly systematic trajectories of change in inner circle varieties of English (e.g. D’Arcy 2015; Tagliamonte & Denis 2014), only few studies (e.g. Fuchs & Gut 2016) have analysed ongoing change in adjective amplification in Asian varieties of English. The current study adds to existing research in focusing on change in adjective amplification in HKE, IndE, and PhiE.

One of the most consistent findings in previous research on changes in adjective amplification in informal spoken discourse has been the replacement of very by really. The present study uses Conditional Inference Trees to assess if this trend holds true for HKE, IndE, and PhiE and if the underlying factors that drive this change in inner circle varieties are also at work in these Asian English varieties.

The analysis shows that the amplifier systems of HKE and IndE are very stable and amplifier choice in these varieties is determined predominately by intra-linguistics factors (adjective type, syntactic context, semantic category). In contrast, the amplifier system of PhiE shows notable signs of ongoing change which is driven predominately by social factors (age and gender of speakers).

The results indicate that during stasis and initial stages of change, language-internal factors determine amplifier choice while social factors become more important once changes have reached mid-range. The paper argues that once certain variants (really and so) gain social meaning, it is this social meaning (the association with specific social groups) drives and accelerates change as speakers want to associate with social groups that carry covert prestige.

## References

- D’Arcy, Alexandra F. 2015. Stability, stasis and change – the longue duree of intensification. *Diachronica* 32(4): 449–493.
- Fuchs, Robert & Ulrike Gut. 2016. Register variation in intensifier usage across Asian Englishes. In Heike Pichler (ed.), *Discourse-pragmatic variation and change: Insights from English*, 185–210. Cambridge: Cambridge University Press.
- Tagliamonte, Sali. A. & Derek Denis. 2014. Expanding the transmission/diffusion dichotomy: Evidence from Canada. *Language* 90(1): 90–136.

# A novel approach to the ESP keyword list: 2800 entries with frequent lexical bundles for data-driven learning

HENSHAW, Michael (Hokkaido University, Japan)

Keywords: AntConc, DDL, lexical bundles, n-grams, ESP

## Abstract

Here I present and describe the creation of a wordlist designed to go in the hands of students. Many English for Specific Purposes (ESP) wordlists are constructed with coverage metrics and efficiency in mind, often for use by instructors or materials developers, meaning these lists and associated corpora are scarcely seen by students. Furthermore, while general academic multiword lists exist, those designed for specific fields are lacking. Thus, as a model to fill these gaps, I present the Keywords of One Health Biomedical Sciences (KOBS), a 2800-lemma database derived from the One Health English Corpus, a 2.8 million-token corpus of 651 research articles. Both materials were created for and distributed to L2 first-year PhD students of veterinary medicine, a multidisciplinary field guided by the One Health approach whose interests overlap with medicine, ecology, and other health sciences. To my knowledge, this is the first large vocabulary list or corpus developed for veterinary medicine. Additionally, from inception, KOBS was designed for the student-as-user; it is an easily navigable 2 MB Excel file with over 10 tabs of curated sublists. Each entry includes: most-used lexical bundles both left and right of the node, thematic category (e.g. lab technique, comparison, biochemistry), highest correlated subcorpus (e.g. Introduction, Infectious Diseases), and more. KOBS is a data-driven learning tool for writing research papers whereby students may, for instance, filter the Discussion sublist for category 'transition' to return 31 words disproportionately used in the discussion section: #1) also. freq./text=3.38: "it has also been shown to", "not only in X, but also in Y"; #31) unfortunately. freq./text=0.04: "Unfortunately, this (method does not)". In short, this enhanced keyword list offers a method for writers to bypass direct interaction with corpora and achieve quick solutions.

# As Good as the Men? A Corpus-Based Discourse Analysis of Media Representations of Athletes Competing in the New Women's Australian Football League

KEMBLE, Melissa (The University of Sydney, Australia)

Keywords: gender bias, sports news, discourse analysis, Australian Rules Football (AFL), evaluation, corpus linguistics

## Abstract

This study investigates how players in the newly established women's Australian Football League (AFLW) are represented in the print news media, with respect to patriarchal discourses. Australian Football not only has one of the largest sporting communities across the country, but also boasts the fourth highest live crowd attendance globally (Gullen, 2015). As with many other physical contact sports, the professional league originated exclusively as a male sport, until the new women's league officially commenced in 2017. This research is based on a 87,500 word corpus of AFLW news articles published in the Herald Sun newspaper, one of the most widely-read papers across Australia. The data are taken for the 12-month period commencing June 2016, excluding articles published during the nine-week season. A reference corpus is also used. Combining corpus linguistics with text analysis of appraisal (i.e. evaluative language), I analyse the AFLW corpus to identify whether there is evidence of patriarchal discourses (i.e. othering, objectification, trivialisation, stereotyping). I also compare the media coverage before the season to the coverage after the season in order to identify whether there are any significant differences in the way players are represented. The keyword analysis reveals a potentially positive shift away from previously documented patriarchal discourses; however, the text analysis reveals a disproportionate focus on gender stereotypes and othering. The (female) athletes are frequently negatively appraised in direct comparison to their male counterparts, which thus positions the professional Australian Football sporting domain as preferably male. This research contributes to existing linguistic research on gender bias by exploring how female athletes entering a male-dominated 'masculine' sport (Koivula, 2001) are portrayed in the media, with respect to patriarchal discourses. It provides a foundation for future linguistic research into representations of athletes in sports reporting using a mixed-method approach combining corpus-linguistics with CDA.

## References

- Gullen, S. (2015). Footy punches high as crowd pleaser. Herald Sun. . Last accessed March 2020.  
Koivula, N. (2001). Perceived characteristics of sports categorized as gender-neutral, feminine and masculine. *Journal of Sports Behaviour* 24(4), 377-393.

# Becoming a President: A Diachronic Study on the Language of Brazilian President Jair Bolsonaro

MATTE, Marine (UNIVATES, BR); GOULART, Larissa (NAU, US); SARMENTO, Simone (UFRGS, BR); REBECCHI, Rozane (UFRGS, BR)

Keywords: Discourse analysis, political speech, diachronic variation.

## Abstract

Since the election of far-right candidate Jair Bolsonaro as president of Brazil, several newspapers have written about the president's discourse and his use of language (cf. Beirão, 2019; Lago, 2018). Some claim Bolsonaro speaks the "language of the people" or "the language of social media", among other coinages. Nevertheless, none of these reports have attempted to conduct a systematic analysis of Bolsonaro's speech. Following corpus-assisted discourse analysis studies (Baker, 2006), in this presentation, we conduct a diachronic investigation of Bolsonaro's political speeches during his seven terms in the House of Representatives, which started in 1991. Our goal was to identify the themes that became popular as his political career progressed. The corpus was compiled from the official records of the government and divided into seven subcorpora representing each mandate. Multi-word units (MWU) were extracted using Sketch Engine. Then, key-MWUs were identified for each of the seven mandates using the speeches of other representatives as a reference corpus. Overall, 1,641 MWU composed of 2 to 4 words were identified as key, with 279 of those occurring in all seven subcorpora. These MWUs were then classified according to semantic categories that emerged from the data, such as legal matters, armed forces, money and natives' rights. Preliminary results show that over time references to money and communism have declined, while references to education, punishment, and to electoral matters have steadily increased. The results of this quantitative and qualitative study can help us better understand the construction of the far-right discourse in Brazilian Portuguese as Bolsonaro is a representation of right-wing politics, and encourage future comparisons across languages.

# Comparing measures of directness in corpora of essays written by Iraqi EFL learners, native English-speakers, and advanced ESL college students

COX, Ashleigh (Georgia State University, USA); FRIGINAL, Eric (Georgia State University, USA); MUSTAFA, Sabah S. (University of Baghdad)

Keywords: learner corpus, comparative corpora, EFL writing

## Abstract

Learner corpora offer valuable insight on variation in EFL writing. This study compares markers of directness and indirectness in corpora of argumentative essays written by Iraqi students learning English, native English-speaking undergraduate students in the United States, and advanced nonnative English-speaking undergraduate students in the United States. The advanced nonnative English-speaking group is included as a successful nonnative model for comparison. There are many previous comparative corpus studies, but Iraqi EFL writing remains under-researched. Ortega (2017) argues that there is a need to include more under-researched populations of language learners in SLA studies, and a learner-corpus approach is one way to address this gap by expanding our knowledge of inter-language writing patterns (Gilquin & Granger, 2015). One area of writing that can be studied using learner corpora is cultural variation in argument styles. Comparing learner writing with successful nonnative English writing can also help inform teachers and researchers about how to help English learners advance. This can be particularly useful for research on helping students with TOEFL essays.

In this study, corpora of argumentative essays written by 100 Iraqi students learning EFL, 100 native English-speaking undergraduate students in the United States, and 100 advanced ESL undergraduate students in the United States were compared. Each student participating wrote two essays, one on the prompt “Write an essay based on this statement: Nowadays, people put too much emphasis on personal appearance and fashion” and another on the prompt “Write an essay based on this statement: The best way to ensure a good future for yourself is to plan carefully while you are still young.” The essays were scored using TOEFL rubrics, and linguistic features were tagged using the Multidimensional Analysis Tagger (Nini, 2014), which use a tagging system modeled after the Biber Tagger. The use of hedges, possibility modals, downtoners, emphatics, and amplifiers were compared for essays written by each group of students, taking into consideration the effect of the prompt. The three groups differed significantly in their use of hedges, possibility modals, downtoners, and emphatics, and they differed moderately significantly in their use of amplifiers. Prompt had a significant impact on the use of hedges and downtoners. Implications for teaching EFL writing, with a special focus on learners from Iraq, are discussed.

## References

- Gilquin, G., & Granger, S. (2015). Learner language. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp.418-435). Cambridge University Press.  
<http://hdl.handle.net/2078.1/145508>
- Nini, A. (2014). Multidimensional Analysis Tagger (Version 1.3.1) . Available from:  
<http://sites.google.com/site/multidimensionaltagger>.
- Ortega, L. (2017). New CALL-SLA research interfaces for the 21st century: Toward equitable multilingualism. *Calico Journal* 34(3), 285-316.

# Contrasting Narratives: The Greek Financial Crisis in Newspaper Editorials

LINGLE, Will (University of Aizu, Japan)

Keywords: Critical Discourse Analysis, Corpus Assisted Discourse Studies, News Discourse

## Abstract

Critical discourse analysis (CDA) provides approaches to explicitly political analyses of texts in their social contexts, though CDA has attracted criticism for overinterpretation of results from analyses of small numbers of texts (Widdowson, 2004). Central to the critical interest in mass media representations is the claim that media texts exert ideological effects on readers in cumulative fashion through repetition of similar characterizations of events and social actors. Corpora offer means of investigating CDA claims, which may be combined in an approach known broadly as corpus-assisted discourse studies (CADS) (Baker, 2006; Partington, 2003). This study began with a Systemic Functional Linguistics (SFL) transitivity analysis (Halliday and Matthiessen 2014) of two US newspaper editorials on the Greek financial crisis, which revealed differences in how the two texts expressed sympathy for key social actors and attributed blame for the crisis. To test whether the portrayals in these two texts represented consistent patterns in the editorial pages, corpora of over half a million words of editorial texts from each paper from 2013-2015 were compiled. Texts focusing on the Greek crisis were retrieved from these using AntConc (Anthony 2011) to produce smaller specialized corpora. Using these, the transitivity analysis was scaled up to focus on all clauses representing one of the three main groups of social actors in the crisis. In each corpus, dominant patterns of representation that could be called 'narratives' or schema emerged which revealed that the two newspapers expressed sympathy and attributed blame for the crisis in ways consistent with the two sample texts. The New York Times expressed sympathy for the Greek people and blamed the EU creditors, while the Washington Post blamed the Greek government and praised the EU creditors. Corpora proved valuable informants into broader patterns of discourse, and as a means of reducing researcher bias.

## References

- Anthony, L. (2011). AntConc (Version 3.2.2.1). . Tokyo, Japan: Waseda University. Available at: <https://www.laurenceanthony.net/software>
- Baker, P. (2006). Using corpora in discourse analysis. Bloomsbury.
- Halliday, M. and Matthiessen, C. (2014) Halliday's Introduction to Functional Grammar (4th ed.). London: Routledge.
- Partington, A. (2003). The linguistics of political argument: The spin-doctor and the wolf-pack at the White House. Routledge.
- Widdowson, H. G. (2004). Text, context, pretext: Critical issues in critical discourse analysis. Blackwell.

# Media Representations of ‘Leftover Women’ in China: A Corpus-assisted Critical Discourse Analysis

Yating YU, Department of English, The Hong Kong Polytechnic University

**Keywords:** Leftover Women; Media Representations; Corpus-assisted Critical Discourse Analysis; Gender Ideologies; Chinese English-language News Media

## Abstract

The term ‘leftover women’, commonly referring to single women older than 27, has been in popular use in Chinese media since 2007. This study investigates how leftover women are linguistically represented in the English-language news media in China by employing a corpus-assisted approach to critical discourse analysis (CDA). A specialised corpus of 303 English news articles (i.e., 236,254 words), covering the years between 2007 and 2017, was built for this purpose. This study adopts a three-step procedure (i.e. identification, interpretation, and explanation) to examine the immediate co-texts of the lemma LEFTOVER WOMAN by combining the corpus linguistics (CL) concept (i.e. Sinclair’s Meaning Shift Units) and techniques (i.e. collocates and concordances) and CDA approaches (van Leeuwen’s sociosemantic approach, Charteris-Black’s critical metaphor analysis, and Lazar’s feminist critical discourse analysis). These findings shed light on media representations of leftover women, the contested ideologies emerging from these representations, and how shifting gender politics and identity shapes and is shaped by media in the world’s most populous nation. Additionally, this study contributes to the growing literature of corpus-assisted CDA in the domain of gender representations by proposing an analytical framework that is potentially applicable to future research.

## References

- Charteris-Black, J. (2004). *Corpus approaches to critical metaphor analysis*. Basingstoke: Palgrave Macmillan.
- Lazar, M. M. (2007). Feminist critical discourse analysis: Articulating a feminist discourse praxis. *Critical Discourse Studies*, 4(2), 141–164. <https://doi.org/10.1080/17405900701464816>
- Sinclair, J. M. (1996). The search for Units of Meaning. *Textus*, IX(1), 75–106.
- Sinclair, J. M. (2004). *Trust the text*. London: Routledge.
- van Leeuwen, T. (2008). *Discourse and practice: New tools for critical discourse analysis*. Oxford: Oxford University Press.

# Sky High: Building a Corpus of English for Flight Training

Schneider, Andrew (Embry-Riddle Aeronautical University, United States); Rachele Udell (Georgia State University, United States); Eric Friginal (Georgia State University, United States)

**Keywords:** Aviation English, English for Specific Purposes, ESP, Corpus development, flight training, policy

## Abstract

Corpus research in the field of Aviation English has largely focused on communication between professional pilots and air traffic controllers (see Bieswanger, 2016; Borowska, 2017). However, there has yet to be a thorough corpus analysis of discourse between student pilots and their flight instructors. Furthermore, it is currently unknown to what extent the high-stress, high-stakes environment of in-air flight training influences multilingual student pilots' L2 development or if factors such as stress and time pressure impact learners' ability to successfully communicate in this specialized register of English with both familiar and unfamiliar interlocutors. Guided by Biber and Conrad's (2009) framework, a situational analysis of flight training discourse was conducted at a prominent aviation school in the United States. This analysis revealed that flight training operations involve oral, simulated flight, and in-air flight activities featuring a diverse range of communicative interactions between student pilots, instructors, school support staff, and local air traffic control. The resulting Corpus of Flight Training (CFT) is the only representative corpus of authentic discourse used in flight training to date. It is comprised of audio and video recordings of one-on-one, instructional communication between ESL student pilots from a variety of L1 backgrounds paired with both mono- and multilingual English-speaking flight instructors. The CFT currently guides the development of pre-training assessments and language support programs for accepted students (classroom instruction, one-on-one tutoring, etc.) and will potentially inform macro language policy decisions for all FAA-certified flight training programs in the United States. The unique nature of this context has presented an equally distinctive set of challenges in the development of the CFT. Nevertheless, collecting corpus data in high-stress environments is both important and beneficial. This presentation explores these challenges and benefits and details the necessary foundations of policy and procedure needed for data collection in high-stakes occupational training scenarios.

## References

- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge University Press.
- Bieswanger, M. (2016). Aviation English: Two distinct specialised registers?. *Variational Text Linguistics*
- Revisiting Register in English, De Gruyter, 67-86.
- Borowska, A. (2017). *Avialinguistics: The study of language for aviation purposes*. Peter Lang GmbH

# The Construction of Coronavirus in English-Language Indonesian Newspapers: A Corpus-Assisted Discourse Analysis

AL FAJRI, Muchamad Sholakhuddin (Universitas Airlangga, Indonesia)

Keywords: Media representation, Coronavirus, Covid-19, discourse analysis, corpus linguistics

## Abstract

This study aims to examine the construction of Covid-19 in English-language Indonesian newspapers by using corpus-assisted discourse analysis. The data were collated from news articles containing the word “Coronavirus” or “Covid-19” published in two English-language Indonesian news media from January to July 2020. This resulted in 3055 texts with 1,233,822 words. Collocation analysis was employed to reveal discourses by focusing on verbal and adjectival collocates of “Coronavirus” and “Covid-19”. A span of five words on either side of the search term was set and MI score was used as a measure to calculate the collocates. A collocational pair that occurs less than 10 times and has an MI score less than 3 was excluded. The top 100 resulted collocation lists were then grouped thematically on the basis of their semantic meaning, adopting UCREL Semantic Analysis System (USAS) (Piao et al., 2015). Concordance analysis was also carried out to better interpret the collocates and to examine them more qualitatively. My preliminary findings suggest that Covid-19 is constructed as an enemy in war/conflict by using WAR metaphors (e.g. fight, combat, and battle). Portraying this pandemic as a war may force people to be obedient to the government Covid-19 policies as evidence of their patriotism. However, this may also lead to an authoritarian government. For example, the Indonesian government imposes Regulation in Lieu of Law (Perppu) No. 1/2020 on the response to the COVID-19 pandemic that includes the impunity (both civil and criminal law) for policy-making officials in taking any extraordinary measures, which can be misused by the government. Coronavirus is also represented as a deadly disease (example of collocates: kill, threaten, deadly, contagious, and fast-spreading). This portrayal may help make people more aware of the danger of Covid-19. However, it also can create unnecessary fear and stigma. For instance, some people reject the burial of Covid-19 dead in the cemetery near to their residence, in spite of using Covid-19 burial protocols, because they are afraid of being infected.

# The Discursive News Values of the 2017 Marawi City Crisis: A Corpus-Assisted Multimodal Discourse Analysis of Selected Newspaper Reports

GAPAS, Wilfred Gabriel (University of Santo Tomas, Manila, the Philippines)

Keywords: Marawi City, terrorism, news values, corpus-assisted multimodal discourse analysis (CAMDA), press photography

## Abstract

The reporting of both terrorism and conflicts is extensively investigated in past media discourse studies across contexts. However, a dearth in the literature exists regarding newsworthiness and its establishment in various forms of news discourse. Based on this niche, this thesis sought to closely examine how the 2017 Marawi City siege is constructed in print news reports as newsworthy. Specifically, it was aimed at determining (a) the linguistic and visual resources used to construe newsworthiness, (b) the news values used to construe the event's newsworthiness, and (c) the existing multimodal relationship between language and images. Sixty news reports from three Philippine broadsheets were analyzed using Bednarek and Caple's (2017) discursive news values framework. This multimodal corpus was examined with multiple methods, namely: (a) collocation analysis of MARAWI and its top three first order collocates, (b) manual analysis of verbal and visual devices utilized for each news value, and (c) the verbal and visual relationship in terms of each discursive news value. A scrutiny of the corpus reveals how local broadsheet journalists constructed the crisis as newsworthy. Firstly, MARAWI is associated closely with collocates that establish Proximity, Negativity, Eliteness, Personalization, Timeliness, Positivity, and Consonance; through collocates city, said, and Maute, recurring second order collocates of Negativity and Eliteness were largely identified. Regarding the manual analysis of texts and images, a range of verbal and visual devices were identified to construct news values except for Aesthetic Appeal. Secondly, an intrasemiotic assessment of news value trends in both verbal and visual data reveal that while the former constructs a concrete set of news values (i.e., Proximity, Timeliness, Superlativeness, and Eliteness), images appear to lack this feature. Lastly, when assessed for their multimodal relationship, texts and images appear to exhibit harmony, where texts and images complement news values with each other. How the crisis' newsworthiness is constructed in the corpus is believed to have implications on news reporting principles and the role of the media in shaping public awareness of events.

## Reference

Bednarek, M., & Caple, H. (2017). *The discourse of news values: How news organizations create newsworthiness*. Oxford University Press.

# NICT JLE コーパスを用いた日本人英語学習者のイラスト描写における習得レベル別特徴分析

寺田里紗 東京外国語大学大学院生

Keywords: 学習者コーパス, スピーキング評価, 発話分析, 話し言葉コーパス

## Abstract

NICT JLE コーパスを用いた  
日本人英語学習者のイラスト描写における習得レベル別特徴分析

### 発表要旨

近年、学習指導要領では英語力の育成に「思考力・判断力・表現力」が強調され、単に知識・技能を身につけるだけでなくそれを場面・状況に合わせてどのように使うか、に関して注目が集まっている。本研究ではこのような観点から、日本人英語学習者のスピーキングにおけるイラスト描写タスクに関して、英語力レベルの異なる学習者の発話を、レベル別の言語的・談話構成的特徴に焦点を置いて分析した。

NICT JLE コーパス（和泉他 2004）のイラスト描写タスク中のレストラン・シーンが描かれているイラストを分析対象として選択した。まず、ほとんどの学習者が言及する中心的場面に焦点を置いて、その場面に含まれている主要な描写要素（人物・服装・位置・行為・表情・背景など）のリスト化を行った。次にこれらの関連部分の描写に関して、SST レベル別学習者集団の産出データを分析し、総語数、文数などの全体量、個別の描写要素の言及割合および発話分量、使用言語表現などについて観察・分析を行った。また基準として母語話者データも同様の分析をし、比較を行った。

結果として、学習段階が上がるにつれて、描写量、描写要素の割合、使用言語表現の多様性、構造の複雑化などが増加傾向になることが明らかになった。発表では、それらの詳しい特徴と SST レベルとの対応関係を質的分析も含めて報告する予定である。

### 参考文献

Halliday, M.A.K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Halliday, M.A.K. (1994). *An introduction to FUNCTIONAL GRAMMAR*, Second edition. London: Edward Arnold.

和泉絵美・井佐原均・内元清貴（編著）（2004）. 『日本人 1200 人の英語スピーキングコーパス』 東京：アルク.

和泉絵美（2013）. 日本人学習者の英語発話における論理展開力. 投野由紀夫・

金子朝子・杉浦正利・和泉絵美（編著）. 『英語学習者コーパス活用ハンドブック』.

東京：大修館書店出版, pp178-186

# Extending vocabulary profiling to languages other than English

Laurence ANTHONY (Waseda University), Natalie FINLAYSON, Emma MARSDEN, Rachel HAWKES, and Nick AVERY (National Centre of Excellence for Language Pedagogy, University of York)

Keywords: NCELP, vocabulary, profiling, non-English, tools

## Abstract

Vocabulary profiles of corpora are often created as a step towards creating and/or modifying pedagogic materials for a target learner audience. Two of the most used desktop vocabulary profiling tools are Range and its more modern equivalent AntWordProfiler. For online vocabulary profiling, Web VP tool, which is part of Compleat Lexical Tutor is a popular alternative. All these tools can in theory be used to profile texts of any language. However, they rely on levelled vocabulary lists where each item in the list is grouped according to its “word family”, “flemma”, or lemma category. They also rely on each item in a list being a single string of characters (i.e words). These limitations introduce problems when attempting to profile languages such as English and French (and almost all other languages) which are composed of both single- and multi-word units. They also hugely complicate the process of vocabulary profiling for languages with a high degree of declension such as German and Spanish. In this presentation, we will first discuss the problems of vocabulary profiling in English and languages other than English. Next, we will explain how an existing desktop profiling tool was adapted for use at the National Centre for Excellence for Language Pedagogy (NCELP), UK to assist researchers in the creation of curricula specifications for the teaching of Spanish, French, and German vocabulary and also teachers hoping to implement these specifications. Then, we will explain the next stage of the project, which is to develop an open-access, online version of the tool.

# Performance Evaluation of Automated CEFR Level Classification Tools

WANG, Wei-Tung (TUFS, Japan); TONO, Yukio (TUFS, Japan)

Keywords: tools evaluation, automated CEFR level scoring, spoken language assessment

## Abstract

There is a growing interest in using the Common European Framework of Reference for Languages (CEFR) for classifying teaching points (vocabulary, grammar, and texts) into different proficiency levels. Recently, several automated tools are available for judging the CEFR levels based on text characteristics (e.g. Text Inspector, CEFR-based Vocabulary Level Analyzer, among others). The present study aims to critically evaluate the performance of different automated tools for CEFR-level classification of texts.

To this end, the TUFS-ELC Spoken Corpus was used, which is a collection of oral interview transcripts with CEFR gradings judged by two professionally trained raters. According to the procedure of the Cambridge English Assessment, when two raters have different ratings, a third rater will be consulted for further evaluation to reach an agreement. We basically followed the same procedure. This spoken corpus can serve as a gold standard for comparing different tools' performance. For the performance evaluation, 184 samples were selected from the major CEFR levels, A2 to B2, in order to be processed by three classification tools: (1) Text Inspector, (2) CEFR-based Vocabulary Level Analyzer (CVLA, Uchida & Negishi, 2018), and (3) English Level Checker (Hayashi, et al. 2017). The results were compared against the human raters' judgements as the gold standard.

The results show that the overall CEFR level classifications turned out fairly accurate across the tools, although each tool showed strengths and weaknesses in dealing with different aspects of text characteristics such as how to deal with texts with different length, the proportion of CEFR level vocabulary, the robustness against off-list words, among others. Also, qualitative analyses will be made in order to determine what textual properties affect the inaccurate estimation of CEFR levels compared to human ratings. Possible implications and future directions of automated CEFR-level text classification will be discussed.

## References

Hayashi, M., Sasano, R., Takamura, H., & Okumura, M. (2017). Judging CEFR levels of English learners' essays based on error-type identification and text quality measures. *Proceedings of the 18th International Conference on Intelligent Text Processing and Computational Linguistics, 2017*. Uchida, S. and M. Negishi (2018) Assigning CEFR-J levels to English texts based on textual features. In Y. Tono and H. Isahara (eds.) *Proceedings of the 4th Asia Pacific Corpus Linguistics Conference (APCLC 2018)*, pp. 463-467.

# Using the USAS Semantic Tagset to Explore Persuasive Language in Jeremy Taylor’s Holy Living and Holy Dying, 1650-1651

THOMAS, Dax (Meiji Gakuin University, Japan)

Keywords: semantic tagging, persuasive language, Wmatrix

## Abstract

This presentation reports on the initial stages of a study on persuasive language in two texts, Holy Living and Holy Dying, written by Jeremy Taylor in 1650 and 1651. The purpose of the study is two-fold: 1) to explore the persuasive language Taylor uses in his writing; and 2) to explore the usefulness of semantic tagging in this type of investigation. The corpus, consisting of the two Taylor texts (226,035 tokens), was first tagged with the USAS Semantic Tagset using the Wmatrix interface. A list of key concepts (semantic keyness) was generated using the Wmatrix interface and concordance lines were consulted for finer detail on the nature of the persuasive technique being used. Several persuasion techniques (such as “emotional appeal”, “attack”, “inclusive/exclusive language”) were selected and semantic tags were identified from the USAS tagset that related to each of these persuasion techniques. When exploring “emotional appeal”, for example, the “E” tag (Emotion) was used as a search item. While not all semantic tags resulted in useful search results, it was found that Taylor seemed to prefer negative persuasion techniques, such as appeals to fear and sadness in his writing. This is illustrated well by four out of the top five most frequent Emotion-related items being E4.1- (repentance), E5- (fear), E4.1- (sorrow), and E4.1- (sad). By working with a general-to-specific approach — that is, from persuasion technique, to general semantic concept category, to specific lexical item — elements of persuasion in the text could be readily identified.

# 現代英語における he/she of NP の振る舞い—COCA での調査から

MATSUDA, Yuji (Ritsumeikan University, Japan)

Keywords: 人称代名詞（主格）, of 句による後置修飾, he/she of NP, 同格用法, 人称代名詞（主格）の削除

## Abstract

Quirk et al. (1985: 352) は、現代英語において、人称代名詞を修飾する際には厳しい制限があるが、いくつかのタイプは容認されると指摘している。しかし、Quirk et al. (1985: 352) の指摘がどのように我々に影響するのかわ、より具体的に調査する必要がある。実際に、“Jack Sparrow, he of the gold teeth, hoop earring and (The Guardian, 2011/5/19)”などの事例は観察されるからである。そこで、本発表は、人称代名詞（主格）が of 句により後置修飾される事例に焦点を絞り、現代英語における he/she of NP の特徴を提示する。まず、The Guardian (2011/5/19) の事例のように、固有名詞（人名）の直後に、he/she of NP が同格用法として生起する事例については、COCA を用いた調査の結果、やや Newspaper のジャンルで用いられるという結果が得られた。さらに、同格用法で用いられる he/she of NP の he/she が削除される事例を指摘し、その分析を行った(e.g., Jack Sparrow, of NP)。その理由は、人称代名詞の後置修飾を極力避けたいという発信者の意識によるものであると指摘した。そして、この言語事実は前述した Quirk et al. (1985: 352) の指摘と矛盾しないものである。

# A Corpus Based Approach to Creating an Advanced Wordbook for University Students

UCHIDA, Satoru (Kyushu University, Japan), SHIMIZU Takehiko (Kyushu University, Japan), KIMURA Saaya (Kyushu University, Japan)

Keywords: advanced wordlist, vocabulary learning, university education

## Abstract

The purpose of the present study is to report on the creation of an advanced vocabulary wordbook for university students. The book, titled “Word Quest” (Kyushu University Press), aims to cover advanced vocabulary that students are likely to encounter when developing their expertise. It contains about 1500 words and phrases in total with three main sections of “Academic”, “SDGs” (Sustainable Development Goals) and “Studying abroad.” The words in the former two sections are selected based on our original corpus consisting mainly of academic books and papers. Some examples are “agitate”, “ephemeral”, and “viable” from the academic section and “eradicate”, “infectious”, and “sewage” from the SDGs section. To the best of our knowledge, this is the first attempt to group words along the framework of SDGs, and it will be reported how we achieved this objective. It will be also shown how we created this book collaboratively with university students as an attempt to cover their actual needs in their current and future studies.

To examine the usefulness of our wordlist, a quantitative analysis was carried out using online news articles. It can be assumed that news passages contain advanced vocabularies that are essential for understanding the current issues around the world, hence important for university students who wish to get continuous updates on the world trend. For this purpose, 408 articles published in a week were collected from the BBC website with average length of 681.6 words. It will be shown that 70.6 percent of the articles contain at least three words from the book and 96.1 percent contains at least one. Even though the words in our list are mostly specialist terms and are not high in frequency, the results show that the book successfully covers what university students encounter when they research

# A Validation Study of the Accuracy of Lexical Diversity Tools

HU XIAOLIN, Tokyo University of Foreign Studies

Keywords: lexical diversity, tools evaluation, accuracy

## Abstract

In recent years, a variety of lexical diversity tools are available for easy access to the analysis of a large amount of texts or corpora (Stuart Webb, 2019). However, few studies focus on the accuracy assessment for these tools.

This research aims to evaluate the accuracy of the lexical diversity tools and to explore the possible factors affecting the outputs of different tools. To this end, six widely used tools were tested on 24 transcriptions from NICT JLE Corpus (Izumi et al. 2004). The tools were divided into two groups according to their lemmatization principles. Group 1 (no lemmatization, all forms treated as different wordforms, i.e. types) contains Text Inspector (Bax, 2012), VocabProfile Program (Cobb, 2002), Coh-Metrix (Graesser, 2004) and CLAN (MacWhinney, 2000); Group 2 (with lemmatization, all inflected forms treated as one type) contains CLAN (MacWhinney, 2000), Lexical Complexity Analyzer (Lu, 2012) and TAALED (Kyle & Crossley, 2015). The type/token ratio (TTR) and Guiraud's index obtained by manually counted types and tokens were used as gold standards to evaluate the accuracy. This research only focuses on TTR and Guiraud, but the accurate counting and recognition of types and tokens will affect the other metrics as well.

The results show that in Group 1, Text Inspector has the most accurate mean scores in TTR against the gold standard compared with other tools (Error percentages: Text Inspector 0.33%, Coh-Metrix 1.00%, VocabProfile 2.22%, CLAN 4.20%). In CLAN(unlemmatized mode), a positive correlation between the proportion of contracted forms in texts and the error rates of TTR was observed ( $r = 0.95$ ). For Group 2, CLAN(lemmatized mode) had the most robust performance in TTR (CLAN 0.70%, Lexical Complexity Analyzer 2.43%, TAALED 13.74%). Although a smaller difference was observed in Guiraud's index, Text Inspector (Group 1) and CLAN (Group 2) outperformed the other tools. Therefore, tools for lexical diversity should be selected carefully according to different data types and research aims.

## References

- Bax, S. (2012). Text Inspector. Online text analysis tool. Available from: <https://textinspector.com/>.
- Cobb, T. (2002). The Web Vocabulary Profile. [http://www.er.uqam.ca/nobel/r21270/texttools/web\\_vp.html](http://www.er.uqam.ca/nobel/r21270/texttools/web_vp.html)
- Graesser, A., McNamara, D., & Louwse, M. (2004). Coh-Metrix, available from <http://www.cohmetrix.com>
- Izumi, E., Uchimoto, K. and Isahara, H. (2004). The NICT JLE corpus: Exploiting the language learners' speech database for research and education. *International Journal of the Computer, the Internet and Management*, 12(2): 119–125.
- Kyle, K. & Crossley, S (2015) TAALED, available from <http://www.linguisticanalysistools.rog>
- Lu, Xiaofei (2012). The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives. *The Modern Language Journal*, 96(2), 190-208.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk: Volume 1: Transcription Format and Programs. Volume 2: The Database*. Lawrence Erlbaum Associates.
- Webb, S. (2019). *The Routledge Handbook of Vocabulary Studies*, London: Routledge

# 付加疑問との連鎖関係からみた右方転位構造（テイル）の機能

YAMASAKI, Nozomi (Kansai Gaidai University, Japan)

Keywords: Spoken BNC2014, 右方転位構造（テイル）, 付加疑問, 発話末, 連鎖

## Abstract

本研究は、話し言葉に特徴的な右方転位構造—「テイル(tail)」と呼ぶ—が発話末で付加疑問(TQ)と共起する場合、その連鎖関係がテイルの機能にどのように影響するかという点について分析する。

テイルは、節中の要素と同一指示の要素が再度、名詞句(1)や代名詞(2)の形式で発話末に現れる構造であり、しばしば付加疑問と共起する。

(1)so but it's a much better place now isn't it? The park? (Spoken BNC2014: SC7H)

(2)so it's a nice circuit that isn't it? (Spoken BNC2014: SGAN)

テイルには、指示対象の確認・明確化や強調機能、話者の評価や感情を聞き手と共有する対人関係機能があることが指摘されている (Aijmer, 1989; Biber et al., 1999; Durham, 2011; McCarthy & Carter, 1997; Rühlemann, 2007; Timmis, 2009, 2015)。

本調査では、Spoken BNC2014 から付加疑問と共起するテイルを収集した。TQ との連鎖順序とテイルの機能の関係を、発話が評価的か否かの別、テイルの構造や形式、さらにテイルの意味カテゴリーを組み込んで分析した。

その結果、名詞句テイルの場合、「TQ+テイル」の順のテイルは指示対象を明確化・確認する機能が大きく、「テイル+TQ」は、話者の態度表明や感情表出とより関わっていること示した。一方、代名詞テイルの場合、all of that など 2 語以上のテイルのほとんどが「TQ+テイル」の順序で現れ、確認・強調機能との関連性を示したが、that など 1 語のテイルは「テイル+TQ」の順序で現れる傾向が高かった。これらのことから、テイルが付加疑問と共起する時、共起する順序によって、テイルが果たす主な機能に違いが見られることや、名詞句テイルと代名詞テイルは、付加疑問との共起パターンが同じではないことが分かった。

## 参考文献

- Aijmer, K. (1989). Themes and tails: The discourse functions of dislocated elements. *Nordic Journal of Linguistics*, 12, 137-154.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education.
- Durham, M. (2011). Right dislocation in Northern England: Frequency and use — perception meets reality. *English World-Wide*, 32(3), 257-279.
- McCarthy, M. & Carter, R. (1997). Grammar, tails and affect: Constructing expressive choices in discourse. *Text*, 17(3), 405-429.
- Rühlemann, C. (2007). *Conversation in context: A corpus-driven approach*. Continuum.
- Timmis, I. (2009). 'Tails' of linguistic survival. *Applied Linguistics*, 31(3), 325-345.
- Timmis, I. (2015). Tails. In K. Aijmer, & C. Rühlemann (Eds.), *Corpus pragmatics: A Handbook* (pp. 304-327). Cambridge University Press.

# 経済学国際ジャーナルにおける Method の章のコーパス分析

中谷安男 法政大学経済学部 日本

Keywords: Academic Writing, Economics Journal Corpus, Research Methods, Applicability of Corpus Linguistics

## Abstract

競争力の高い国際ジャーナルに論文が掲載されるためには、研究の信頼性と妥当性を確立し、論文の中で示した実験手法などを再現できるように客観的に書く必要がある (Cohen, 1994)。これらを明示するのがメソッド (Method) と呼ばれる研究計画や研究手法を記載する章である。査読者は、この章を精査し、研究手法が明確で、適切に結果を導きだしているか判断し採択を決める (中谷, 2016)。しかしながら Swales(2004)が指摘しているように、これまで Method の適切な書き方に関する論文は多くない。また、規模の大きなコーパスに基づき、信頼性の高い手法で特徴的な語彙やクラスター表現を分析した研究は少ない。特に、経済学の国際的論文のコーパスデータに基づく Method の検証はほとんどない (中谷, 2020)。

本論はこの点に注目して、インパクトファクターの高い以下の4つの学術雑誌から37本の論文を集め、約65万語のEJC (Economics Journal Corpus)を構築した。American Economic Review、Econometrica、Quarterly Journal of Economics、International Economicである。今回は、EJCの中からMethodの章を抽出しコーパスを構築し、論文の他の章であるIntroduction, Result, Discussionの章のコーパスと比較検証した。AntConk Windows (3.5.8)を活用し2つのコーパスのKeyword分析を行った。結果としてMethodにおける名詞や動詞の特徴語やクラスター表現を抽出できた。さらに、これらを活用したMoveの構成が明らかになった。

# Symposium: 小・中・高における DDL 普及への挑戦 —DDL ツールの開発, 授業実践, 分野横断的考察—

西垣知佳子 (千葉大学) 赤瀬川史郎 (Lago 言語研究所) 水本篤 (関西大学) 石井雄隆 (千葉大学)  
Peter Crosthwaite (University of Queensland) 安部朋世 (千葉大学) 物井尚子 (千葉大学)  
小山義徳 (千葉大学) 星野由子 (千葉大学) 神谷昇 (千葉大学) Pichinart Kumpawan (Surasakmontree School)

Keywords: 発見学習, eDDL, hDDL, BES Search, 文法指導

## Abstract

本シンポジウムは DDL SIG の企画である。DDL (data-driven learning) は、主に大学生を対象として世界的に利用が広がっている。本 SIG 企画は、日本の小・中・高校に DDL を導入するための試みと成果を、以下の 4 点から報告、検討する。

1) 小・中・高校生のための 3 種のウェブ DDL ツールと教材: eDDL, hDDL, BES Search  
まず、これまでに学校現場で実施してきた DDL 実践の結果明らかになった DDL の成果を報告する。続いて、小・中・高校の英語授業で DDL を普及させるために開発した 3 種類のウェブ DDL ツールを紹介する。これらのツールは、学習者の英語力と認知レベルに合致する学習用コーパスと、使い易い検索ツールを搭載しており、小学生用 DDL ツール (eDDL), ならびに中・高生用 DDL ツール (hDDL) を開発した。併せて、教師が DDL 教材作成に利用できる入門・初級レベルの英文を収集した検索ツール (BES Search) を開発した。全て登録不要、無料で利用できる。

## 2) DDL 実践の分野横断的考察

教育心理学の観点から、DDL における発見学習の効果を検討する。

また、発表者らの行う DDL では、学習者は帰納的に英語の文法規則を発見し、日本語で言語化する。学習者の発見内容を見てみると、国語科で学んだ知識を使って、英語の文法規則を記述していることがわかった。そこで、文法に関連する事柄が、外国語科と国語科の検定教科書で、どのようなことが、どのように学ばれているかを調査した。その結果を報告する。

## 3) DDL の指導と評価

既存の大学生用 DDL 評価テストと質問紙調査を検討し、小・中・高校生用の評価テストと質問紙の開発に向けた構想を述べる。

## 4) 国際的視点から考える DDL

本プロジェクト独自の DDL をタイ・バンコクの中学校英語教師が実践した事例を紹介する。また、pre-tertiary learners の DDL 研究の先駆者である University of Queensland の Crosthwaite 先生に、海外での初等・中等教育における DDL についてお話しをいただく。

以上を踏まえて、最後に、日本の初等・中等教育における DDL の活用について検討する。

## 参考資料

小学生用 DDL ツール (eDDL) <https://e.ddl-study.org/>

中・高生用 DDL ツール (hDDL) <https://h.ddl-study.org/>

検索ツール (BES Search) <https://bessearch.ddl-study.org/>

## **Timetable of symposium**

### **Part 1: Data-driven learning for younger learners: Current issues, future directions.**

Peter Crosthwaite (University of Queensland)

### **Part 2: 小・中・高校生のための3種のウェブDDLツールと教材:eDDL、hDDL、BES Search**

西垣知佳子 (千葉大学)、赤瀬川史郎 (Lago 言語研究所)、石井雄隆 (千葉大学)、神谷昇 (千葉大学)

西垣が発表

### **Part 3: 小・中・高校におけるDDL実践の分野横断的考察**

① 発見学習からDDLの効果を考える 小山義徳 (千葉大学)

② 小学校英語教育の分野から 物井尚子, 星野由子 (千葉大学)

③ 小学校国語科教育の分野から 安部朋世 (千葉大学)

### **Part 4: DDLの指導と評価**

水本篤 (関西大学)

### **Part 5: Communicative English Classroom using DDL in Thailand**

Pichinart Kumpawan (Surasakmontree School, Thailand)

---

英語コーパス学会 (Japan Association for English Corpus Studies)

会長 石川慎一郎

事務局 〒819-0395 福岡県福岡市西区元岡 744 九州大学大学院言語文化研究院 内田諭

e-mail: [jaecs.hq@gmail.com](mailto:jaecs.hq@gmail.com) 郵便振替口座:00930-3-195373 URL: <http://jaecs.com>

---