

英語コーパス学会
第47回大会
アブストラクトブック

**JAECS 47TH
CONFERENCE
ABSTRACT BOOK**

2021.10.2

JAECS
Japan Association for English Corpus Studies

2021/10/2 英語コーパス学会第47回大会 (JAECS 47th Conference)
プログラム(8/20現在)



0950-1000	開会式					
1000-1100	基調講演1 松本曜先生「認知意味論研究におけるコーパスと実験の利点と限界」(司会・指定討論者:森下裕三)					
	Session 1 英語学・英文学(発表室1)【司会:杉森】		Session 2 英語教育(発表室2)【司会:堀家】		Session 3 言語資源開発(発表室3)【司会:佐々木】	
1105-1125	村岡	【1】知覚動詞補文に出現する受身表現の容認可否について	澤口	【2】日本人大学生EFL学習者の make+名詞のコロケーション使用について	石川(慎)	【3】「1961-2021日本語小説コーパス」の構築:日英マインドスケープ対照研究の新しい可能性
1130-1150	泉類	【4】中英語期の補部と2人称代名詞の構文関係:ICAMETによる分析	チホネンコ・望月	【5】Verification of the effectiveness of 20 months of speaking lessons for high school learners: An analysis of fluency on the Aptis speaking test	仁科・赤瀬川	【6】日英・英日/パラレルコーパスオンライン検索ツール『(仮称)パラレルリンク』(Ver.1.0)の開発に向けて
1155-1215	黒田	【7】多様な指標を組み込んだトピックモデル可視化ツールの開発とテキスト分析への応用	ニューベリ・ペイトン	【8】A Learner Corpus-Based Study of L1 Effects on L2 English Auxiliary Verb Use: The Case of "Will"	大橋・片桐・押切	【9】授業コーパス構築のための自動タグ付けツール "Classroom Corpus Tagger" の開発
1215-1240	会員総会					
	Session 4 英語教育(発表室1)【司会:和泉】		Session 5 英語教育・英語学(発表室2)【司会:梶山】		Session 6 ESP(発表室3)【司会:村岡】	
1330-1350	佐々木	【10】日本人学習者の英語原因表現使用:ICNALEに基づく量的概観	小島・金田	【11】ライティング評価とテキストの特徴との相関関係:メタ分析による研究成果の統合	石川(有)	【12】工学系大学院生のための教材開発-日英コーパスの分析
1355-1415	畔元	【13】CEFR準拠教科書における英語コロケーションの難易度変化要因の特定	石井・河本	【14】N-grams at the Beginning of the Moves in the Results Section of Experimental Medical Research Articles	清水・村田	【15】生化学英語学術論文のための学術語彙リスト
1420-1440	堀家	【16】高校英語指導における句動詞の扱い-教科書とセンター試験の分析から-	中谷	【17】オックスフォード・ユニオンにおけるリーダーシップ育成の示唆:英語圏のリーダーの発話コーパス分析	小谷・佐野	【18】金融関連辞典と実務資料コーパスを用いた経済・金融分野の英語語彙リスト研究
	Session 7 英語学・英文学(発表室1)【司会:田畑】		Session 8 文法・統語(発表室2)【司会:泉類】		Session 9 DDL(発表室3)【司会:畔元】	
1450-1510	藤田	【19】LDA Topic Modelling of Tennyson's Poetry	松田	【20】have long V-ed 構文の典型例	佐竹	【21】英語の動詞-名詞コロケーション学習に対するDDLの効果
1515-1535	内田	【22】チャンクのコロケーション:spaCyを用いた共起分析の試み	安間	【23】Distribution of repeated appearance of grammar items in junior high school textbooks through nonlinear regression	西垣他	【24】中学生のためのWeb版DDL支援ツールの開発と活用
1540-1600	木山・渋谷	【25】動詞の意味はトピックから推測できるのか:英語の動詞 run を例に	小林・佐野	【26】現代スペイン語における主語後置の数理モデル化	アントニ	【27】Introducing AntConc 4.00: A fast, powerful, and easy-to-use corpus analysis tool for small and large-scale corpus analysis and data-driven learning
1610-1710	基調講演2 Vaclav BREZINA先生「Statistics and data visualization in corpus linguistics with #LancsBox」(司会・指定討論者:宇佐美裕子)					
1710-1730	閉会式・学生優秀発表賞授賞式					

ごあいさつ

ようこそ英語コーパス学会第47回大会へ。

コロナ禍が依然猛威を振るうなか、本年の英語コーパス学会大会は、本会にとって、2度目のオンライン大会となります。2020年に開催された第46回大会は、オンラインの良さをいかし、広く海外の非会員からの発表も受けつけ、盛会裏に終了しました。

第47回大会は、これまでの対面学会とできるだけ近い形での開催を目指し、発表資格を会員に限定いたしました。十分な数の発表応募をいただけるか不安もありましたが、おかげさまで、非常に多くのお申し込みをいただきました。厳正な審査の結果、本会の大会としてはおそらく過去最多となる27本の発表が選ばれました。

採択された発表は、英語学・英文学・英語教育学、また、各種のコーパスやツールにかかわるものなど、内容も多彩で、コーパス研究の裾野の広さと、潜在的な可能性の大きさを改めて感じさせるものです。会員のみなさまには、ぜひ、これらの発表を通して、ご自身の研究の新しいヒントをつかんでいただき、今後の研究や教育実践に活かしてもらえればと存じます。

第47回大会では、昨年度に引き続き、「学生優秀発表賞」の審査を行います。これは、すぐれた研究発表を行った学生会員を顕彰するもので、選考結果は閉会式で公表されます。どうぞご期待ください。

また、第47回大会では、大会の活性化を目指し、新しい工夫を3つ試みました。1つめは、予稿集の刊行です。予稿集の刊行により、発表者は、自身の研究発表の内容を正式な記録にとどめることができます。また、聴講者は、予稿を読むことで、発表の内容をよりよく理解することができ、当日の質疑応答の深まりも期待されます。2つめは、大会をサポートする学生スタッフ制度の新設です。この制度は、若い会員に早い段階から学会運営に関わってもらうことを念頭に作られました。今回の大会には、6名の大学院生が大会スタッフとして協力を申し出てくれました。スタッフは、当日の参加者のお世話のほか、一部セッションの司会も担当いたします。本学会の未来を担う若い力を感じていただければ幸いです。最後に3つ目は、前夜祭企画として、2本の講演の司会者によるプレトークを企画したことです。オンラインの学会では気軽な雑談を楽しめないという制約がありますが、うちとけた雰囲気で開催される前夜祭が会員の皆様の懇親の場になればと思っています。

学会の大会と言えば、堅苦しく、窮屈な場というイメージが浮かびます。しかし、それは本来、真摯な「知の競演」の場であるとともに、祝祭的な「知の饗宴」の場でもあるべきでしょう。本日が、ご参加の会員のみなさまにとって、知的な刺激と興奮にあふれる楽しくも充実した一日となりますよう！

2021年10月2日
英語コーパス学会会長
石川 慎一郎

英語コーパス学会 第47回大会実行委員会

JA ECS 47th Conference, Organizing Committee

委員長 石川慎一郎(神戸大)

委員 和泉絵美(京都大) 内田諭(九州大) 杉森直樹(立命館大)

学生スタッフ 畔元里沙子(九州大院) 梶山達也(同志社大院) 佐々木恭子(神戸大院)

泉類尚貴(慶応大院) 堀家利沙(神戸大院) 村岡宗一郎(日本大院)

アブストラクトブック改訂記録

2021/8/18 アントニ氏発表概要を修正

2021/8/21 佐野・小林両氏による発表の第1発表者を変更

2021/9/16 講演者の講演概要(2本)を追加

招待講演 (Invited Talks)

招待講演 I

松本曜先生 (国立国語研究所教授)

「認知意味論研究におけるコーパスと実験の利点と限界」

司会・指定討論者 Chair/ Discussant

森下裕三 (環太平洋大学准教授)

Dr. Yuzo MORISHITA (Associate Professor, International Pacific University)

【概要】

認知言語学の量的研究においては、コーパスと実験の両方が用いられるが、それぞれどのような利点と限界があるのだろうか。保田 (2011)は、「犬」などの具体名詞の意味を特定する際のコーパスと実験の有効性を検討し、動作に関する意味要素を抽出するにはコーパスが向いているが、外見的特徴を抽出するにはコーパスよりも描画実験が有効であるとしている。

また、松本 (2021)は、移動事象の言語表現の研究におけるコーパスと発話実験の有効性を比較し、総合的な言語使用の実態を知るためにはコーパスに強みがあるが、そこに含まれる文がどのような事象を描いているのか確定できないという限界があるとしている。一方、発話実験は特定の移動事象の描写を引き出せる点で有効だが、その結果がどこまで言語使用の代表的なものかについて検討が必要であるとしている。

一般に、コーパスに基づく研究では、言語化されていない情報を扱いにくいという課題がある。コーパスと実験の利点と限界を認識した上で研究を進める必要がある。

【参考文献】

松本 曜 2021. 「移動表現の研究におけるコーパスと実験」 篠原和子・宇野良子 (編) 『実験認知言語学の深化』 287-309. ひつじ書房

保田 祥 2011. 『名詞の百科事典的意味の抽出方法とその有用性: 内省・描画実験・コーパス調査』博士論文、神戸大学.

講演プレトーク (2021/10/1)

森下裕三講師 「はじめての認知言語学」

招待講演2

Dr. Vaclav BREZINA (Senior Lecturer, Lancaster University, UK)

“Statistics and data visualization in corpus linguistics with #LancsBox”

司会・指定討論者 Chair/ Discussant

宇佐美裕子 (東海大学准教授)

Dr. Hiroko USAMI (Associate Professor, Tokai University)

Abstract

In an ideal world, theory and practice would go together hand in hand. Strong theoretical and methodological grounding of corpus linguistic research leads to robust results, which can be meaningfully applied in practice. In reality, however, we can often see a disconnect between high theoretical requirements of current research and what corpus linguists can actually do in their studies using existing tools. For example, we can see a limited range of statistical measures used in corpus linguistics, which until fairly recently typically relied on the log likelihood measure and a couple of collocation statistics (t-score and MI score) precisely because these were easily available in existing tools.

Corpus linguistics as a versatile methodology of language analysis (McEnery & Hardie 2011) thus requires access to appropriate software tools. These need to be able to cope with increasing demands on the sophistication of the analysis and increasing size of the data. In recent years, researchers have been critically re-evaluating the existing procedures in the field and have proposed more rigorous approaches to data analysis (e.g. Kilgarriff 2005, 2012; Gries 2006, 2013; Lijffijt et al. 2014; Brezina & Meyerhoff 2014; Brezina et al. 2015; Gablasova et al. 2017; Brezina 2018). Reflecting on this debate and combining statistical sophistication and accessibility is the main challenge that needs to be met by corpus linguists today; the analyses should encourage a multi-dimensional view of data, easy comparison, and effective visualization.

In this lecture, I will deal with key questions of corpus methodology and statistics and the implementation of statistical solutions in the #LancsBox software (Brezina et al. 2015). #LancsBox is a free multi-platform tool, which can analyse any language. It can be used by linguists, language teachers, translators, historians, sociologists, educators and anyone interested in quantitative language analysis. Extensive documentation about #LancsBox is available, also in Japanese: http://corpora.lancs.ac.uk/lancsbox/docs/pdf/LancsBox_5.1_manualJP.pdf

References:

- Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge University Press.
- Brezina, V., & Meyerhoff, M. (2014). Significant or random. A critical review of sociolinguistic generalisations based on large corpora. *International Journal of Corpus Linguistics*, 19(1), 1-28.
- Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics*, 20(2), 139-173.
- Gries, S. T. (2013). *Statistics for linguistics with R: a practical introduction*. Berlin: Walter de Gruyter.
- Gries, S. Th. (2006). Some proposals towards a more rigorous corpus linguistics. *Zeitschrift für Anglistik und Amerikanistik*, 54(2), 191-202.
- Kilgarriff, A. (2005). Language is never, ever, ever, random. *Corpus linguistics and linguistic theory*, 1(2), 263-276.
- Kilgarriff, A. (2012). Getting to know your corpus. In Sojka, P., Horák, A., Kopecek, I. & Pala, K. *Text, Speech and Dialogue* (pp. 3-15). Berlin: Springer.
- Lijffijt, J., Nevalainen, T., Säily, T., Papapetrou, P., Puolamäki, K., & Mannila, H. (2014). Significance testing of word frequencies in corpora. *Digital Scholarship in the Humanities*, advanced access.
- McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.

講演プレートーク(2021/10/1)

宇佐美裕子講師 「はじめての#LancsBox」

時間帯別 研究発表内容一覧

室	No.	発表者	キーワード
11:05~11:25			
I	1	村岡	知覚動詞, 補文, 準動詞, 受身表現
II	2	澤口	学習者コーパス, コロケーション, 英語基本動詞, 日英対照意味論
III	3	石川(慎)	小説コーパス, 機械翻訳, 経年的言語変化, 日英小説比較
11:30~11:50			
I	4	泉類	Complementation, Directives, Performatives, ICAMET
II	5	チホネンコ・望月	【In English】Longitudinal Speaking Learner Corpus, the Aptis Speaking Test, Fluency, Complexity
III	6	仁科・赤瀬川	パラレルコーパス, ツール開発, オンライン, 日英・英日翻訳
11:55~12:15			
I	7	黒田	トピックモデル, 文学作品分析, Diagnostics, ビジュアルライゼーション
II	8	ニューベリー・ペイトン	【In English】Learner Corpus, Auxiliary verb, Modality, L1 influence
III	9	大橋他	授業コーパス, Tagger, Reflection, 発話分析
13:30~13:50			
I	10	佐々木	英語原因理由表現, because 過剰使用, 日本人学習者, 英作文
II	11	小島・金田	第二言語ライティング, ライティング評価, CAF, 結束性
III	12	石川(有)	工学系論文コーパス, ライティング教材開発, ESP, JSP
13:55~14:15			
I	13	畔元	CEFR, 教科書コーパス, コロケーション, Deep Learning
II	14	石井・河本	【In English】Move Corpus. Results Section. N-grams
III	15	清水・村田	学術語彙リスト, 基礎単語, ESAP, EGAP
14:20~14:40			
I	16	堀家	句動詞, 日本人高校生, 検定教科書, 大学入試センター試験
II	17	中谷	リーダーコーパス, 発話, オックスフォードユニオン, コミュニケーション戦略
III	18	小谷・佐野	ESP, 経済, 金融, 語彙
14:50~15:10			
I	19	藤田	【In English】Alfred Tennyson, LDA, Topic Modelling, Poem
II	20	松田	have long V-ed 構文, have long V-ed 構文の典型例, have long V-ed 構文の V-ed, have long V-ed 構文の主語
III	21	佐竹	DDL, コロケーション, COCA
15:15~15:35			
I	22	内田	コロケーション, チャンク, spaCy, 共起分析
II	23	安間	【In English】Grammar item, Junior high school textbook, Nonlinear regression, Factor analysis
III	24	西垣他	DDL, Teaching-oriented corpus, 気づき, 入門期 英文法
15:40~16:00			
I	25	木山・渋谷	多義語, Biterm topic model, トピック
II	26	小林・佐野	主語後置, 情報構造, 表層特徴, ロジスティック回帰
III	27	アントニ	【In English】AntConc, Corpus tools, DDL, KWIC

発表概要

・第 47 回大会では、以下の 3 つのタイプの発表があります。発表タイプは、発表概要集のタイトルの末尾に略号で記載しています。なお、発表タイプは、発表者自身による申告に基づきます。

- Presentation A

新規の学術的知見の報告・議論 / Report of new research findings

- Presentation B

コーパス・ツール・手法・プロジェクト等の紹介 / Introduction of a corpus, tool, method, and project

- Presentation C

実施中の分析の中間報告 / "Work in Progress" Report

Presentation 01

村岡宗一郎(日本大学大学院生)

【題目】

知覚動詞補文に出現する受身表現の容認可否について【C】

【題目(英語)】

On the Acceptability of Passive Expressions Appearing in the Complement of Perception Verbs

【キーワード】

知覚動詞, 補文, 準動詞, 受身表現

【概要】

現代英語の知覚動詞は補文に原形不定詞や現在・過去分詞をとるが、Bolinger(1974)などの先行研究によれば、I saw the children be beaten. のような受身表現は一般的に容認されない。しかし、村上春樹の『騎士団長殺し』の英訳には、I couldn't stand to see her be cremated. という例が確認される。Palmer(1968)はこのような例は稀であると分析をするが、Palmer(1974)や(1987)からその記述は削除されている。また、多くの先行研究では、I saw the children get beaten. は容認されるというが、BNC と COCA を用いて調査を行った結果、このような例はアメリカ英語にのみ用いられていることが明らかになった。本研究では、BNC や COCA を用いて、英米における使用頻度から知覚動詞補文に出現する受身表現の容認可否について分析していく。

【主要参考文献】

- Bolinger, D. 1974. "Concept and percept: Two infinitive constructions and their vicissitude." *World Papers in Phonetics Festschrift for Dr. Onishi's Kiju*, 65-91. The Phonetic Society of Japan.
- Palmer, F, R. 1968. *Linguistic Study on the English Verb*, London: Longman.
- Palmer, F, R. 1987. *The English Verb*, 2nd Edition, London: Longman.

Presentation 2

澤口遼(甲南高等学校・中学校(非))

【題目】

日本人大学生 EFL 学習者の make+名詞のコロケーション使用について【A】

【題目(英語)】

Exploring the Use of Make + Noun Collocations by Japanese EFL University Students

【キーワード】

学習者コーパス, コロケーション, 英語基本動詞, 日英対照意味論

【概要】

本研究の目的は, 学習者の同トピックについての英日のエッセイを集めた学習者コーパスである KUBEC を用い, 日本人大学生 EFL 学習者の英語動詞 make + 名詞のコロケーション使用において, 日本語「つくる」の意味体系が与える影響が習熟度と共にどのように変化するか調査することにある。

母語の影響を受けていると考えられるコロケーション使用率の頻度を学習者間で比較した結果, 習熟度別では差が見られず, 全習熟度の学習者が「関係」や「環境」など, make よりも抽象的な名詞と共起する「つくる」の多義性の影響を受けたコロケーションを使用していることが明らかになった。

これらの結果から, 学習者へのコロケーション指導への示唆を考察する。

【主要参考文献】

山西博之・水本篤・染谷泰正. (2013). 「関西大学バイリンガルエッセイコーパスプロジェクト—その概要と教育研究への応用に関する展望—」. 『関西大学外国語学部紀要』9, 117-139.

Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, 22(2), 173-195. doi: <https://doi.org/10.1093/applin/22.2.173>.

Anthony, L. (2017). AntPConc (Version 1.2.1) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>

Presentation 3

石川慎一郎(神戸大学)

【題目】

「1961-2021 日本語小説コーパス」の構築:日英マインドスケープ対照研究の新しい可能性【B】

【題目(英語)】

"1961-2021 Japanese Fiction Corpus" : For a New Comparative Study of Japanese/English Fictions

【キーワード】

小説コーパス, 機械翻訳, 経年的言語変化, 日英小説比較

【概要】

構築中の日本語小説コーパスについて報告する。これは Brown Corpus の標本抽出年である 1961 年を起点として, 2021 年まで, 10 年ごとの間隔で 3 大文芸誌(「新潮」「文學界」「群像」)に掲載された日本語小説とその英訳(機械翻訳 2 種)を収集するものである。本コーパスを用いることで, ジャンル要因を統制した上で, 現代日本語の経年変化を調査することができる。また, 付随する英訳データを Brown/ LOB(1961 年), Frown/ FLOB(1991-92 年), Crown/ CLOB(2009 年)等の小説データと対照することで, 時代要因を統制した上で, 日英小説の言語・文体・マインドスケープの比較を行うこともできる。発表では本コーパスの開発理念と手順, また, 収集済みのデータから明らかになった知見の一部を報告する。

【主要参考文献】

- 石川慎一郎(2015)「FROWN/FLOB Corpus および BCCWJ データの再構成に基づく英日対照言語研究用小説テキストデータセットの構築の試み: English-Japanese Modern Fiction Corpus (EJ-MoFic)の概要」『統計数理研究所共同研究レポート』340, 1-18.
- 石川慎一郎(2021)『ベーシックコーパス言語学』第2版. ひつじ書房.
- 三竹保宏(2018)「Deep Learning による AI 機械翻訳のイノベーション」『ビジネスコミュニケーション』55(8), 11.

Presentation 4

泉類尚貴(慶應義塾大学大学院生)

【題目】

中英語期の補部と2人称代名詞の構文関係:ICAMETによる分析【C】

【題目(英語)】

Constructions with Complementation and the Second-Person Pronouns in Middle English: With the Analysis of ICAMET

【キーワード】

Complementation, Directives, Performatives, ICAMET

【概要】

英語史における変化の一つに、補部の変化がある。定型節から非定型節への変化は、中英語期にその萌芽が見られる(Los, 2005)。補部と意味の関係について、Rohdenburg(1995)によれば、非定形節のほうが coercive force が強いことが示されている。Coercive force の強さは、command をはじめとする指令動詞の分析から提示された。一方で、同一の動詞が異なる補部を従える例も見られる。本研究では、補文の変化が起こりだした時代である中英語期に焦点をあてて、指令動詞の現れる構文の一つである、明示的遂行文における2人称代名詞の用法(thou系かye系か)と補部のthat節、不定詞の構文関係について、Innsbruck Computer Archive of Machine-Readable English Texts(ICAMET)を中心に収集したデータを示し、分析の可能性を示す。

【主要参考文献】

Los, Bettelou. *The Rise of the To-Infinitive*. Oxford: OUP, 2005.

Manabe, Kazumi. *The Syntactic and Stylistic Development of the Infinitive in Middle English*. Kyushu University Press, 1989.

Rohdenburg, G. 'On the Replacement of Finite Complement Clauses by Infinitives in English.' *English Studies*. 76:4 (1995), 367-388.

Presentation 5

TIKHONENKO Maxim (東京外国語大学大学院生) / MOCHIZUKI Keiko (東京外国語大学)
チホネンコ・マキシム / 望月圭子

【題目】

Verification of the effectiveness of 20 months of speaking lessons for high school learners: An analysis of fluency on the Aptis speaking test 【C】

【キーワード】

Longitudinal Speaking Learner Corpus, the Aptis Speaking Test, Fluency, Complexity

【概要】

We present a 20-month longitudinal study on the development of speaking ability among Japanese high school learners of English who participated in online speaking lessons. Students were divided into an experimental group of 32 students who took 20 monthly lessons and a control group of 22 students who took 3 lessons.

Both groups then took the Aptis speaking test three months after the last speaking lesson, the fall semester of the third year. Speaking data recorded from the test was transcribed using ELAN, and the durations of pauses and speaking time were measured. The transcribed data was then divided into AS-Units and analyzed from the perspectives of fluency and complexity.

For fluency, speech rate, ratio of pause time to speaking time, number of pauses per minute, ratios of self-corrections, repetitions, and fillers to AS-Unit were measured. The analysis showed that speech rate and the ratio of pause time to speaking time ratio showed the greatest difference. For complexity, the ratio of subordinate clauses to AS-Units and the mean number of words were measured. These measures differed less than the fluency measures between the two groups.

The analysis showed that the experimental group was significantly more fluent than the control group.

【主要参考文献】

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: OUP.

Foster, P., Tonkyn, A., and Wigglesworth, G. (2000). "Measuring Spoken Language: A Unit for All Reasons", *Applied Linguistics* 21 /3, 354-375, Oxford: Oxford University Press.

Housen, A., Kuiken, F., Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In Housen, A., Kuiken, F., Vedder, I. (Ed.). *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*. *Language Learning & Language Teaching* 32, p. 1-20.

Presentation 6

仁科恭徳(神戸学院大学)／赤瀬川史朗(Lago 言語研究所)

【題目】

日英・英日パラレルコーパスオンライン検索ツール『(仮称)パラレルリンク』(Ver.1.0)の開発に向けて【B】

【題目(英語)】

Toward the Development of "Parallel Link (tentative name)" (Ver. 1.0), an Online Search Tool for Japanese-English and English-Japanese Parallel Corpora (Interim Report)

【キーワード】

パラレルコーパス, ツール開発, オンライン, 日英・英日翻訳

【概要】

本発表では、我々が現在開発中の網羅型日英・英日パラレルコーパスオンライン検索ツール『(仮称)パラレルリンク』(Ver. 1.0)について中間報告を行う。まず、現在までに構築された日英・英日パラレルコーパスや検索ツール、それらを活用した研究を振り返り、今後の展望を述べる。次に、一般参照パラレルコーパスの構築に先駆けて、実験的に開発している『(仮称)パラレルリンク』(Ver. 1.0)の開発状況等を報告する。特に、現在までに無償公開された 9 種の日英・英日パラレルコーパスをオンライン上で網羅的に串刺し検索できる本検索システム(Ver. 1.0)は、擬似的な一般参照パラレルコーパスとして活用することができ、検索語に関する精緻な語彙プロファイルの獲得やオーセンティックな翻訳例の獲得において有益なツールとなる可能性があることも示唆したい。

【主要参考文献】

中條清美・西垣知佳子・赤瀬川史朗・内山将夫(2015)「レキシカル・プロファイリング型オンラインコーパス検索ツール LWP for ParaNews の英語授業における利用」『日本大学生産工学部研究報告 B』第 48 号, 45-57.

仁科恭徳(2020)「日英パラレルコーパス WikipediaKyoto-LWP を用いた和英辞典の記述改善案について—「X を固める」の場合—」『英語コーパス研究』第 27 号, 1-21.

染谷泰正・赤瀬川史朗・山岡洋一(2011)「大規模翻訳コーパスの構築とその研究および教育上の可能性」『日本メディア英語学会第 1 回年次大会発表資料』1-15.

Presentation 7

黒田絢香 (大阪大学大学院生)

【題目】

多様な指標を組み込んだトピックモデル可視化ツールの開発とテキスト分析への応用【B】

【題目 (英語)】

Visualization of Topic Models Using Multiple Measures: LDA for Text Analysis

【キーワード】

トピックモデル, 文学作品分析, Diagnostics, ビジュアライゼーション

【概要】

機械学習アルゴリズムの一つであるトピックモデルを用いたテキスト分析において, 各トピックの特徴や関係性, 構成単語や出現傾向など様々な要素を的確に可視化し, モデルの全体像を把握することが極めて重要である. 本研究では, 従来の分析で主に用いられていた「文書ごとのトピック出現確率」「各単語の weight」に加えて, coherence や exclusivity, effective number of words など様々な指標を組み込んだビジュアライゼーションツールを開発し, いかにか効果的にトピックを可視化できるか, それらがどのように文学作品分析に寄与するかを検討する.

【主要参考文献】

Jockers, M. and Mimno, D.: Significant themes in 19th-century literature. *Poetics* 41: 750-769 (2013)

田畑智司:FLOB コーハスの意味構造: 確率論的トピックモデルによる言語使用域の特徴付け
『統計数理研究所 共同研究レポート』386: 1-17 (2017)

Presentation 8

NEWBERY-PAYTON Laurence (東京外国語大学(非))

ニューベリー・ペイトン・ローレンス

【題目】

A Learner Corpus-Based Study of L1 Effects on L2 English Auxiliary Verb Use: The Case of “Will” 【A】

【キーワード】

learner corpus, auxiliary verb, modality, L1 influence

【概要】

This study analyzes use of the modal verb “will” by L1 Chinese and Japanese learners of English. Data is drawn from the Written Essays module of the International Corpus Network of Asian Learners of English (ICNALE). Both frequency and type of use of are predicted to be influenced by the presence (in Chinese) or absence (in Japanese) of functional equivalents to “will” in L1. Analysis reveals overuse of “will” by L1 Chinese learners across proficiency levels. This high frequency of use can partially be attributed to Chinese learners’ consistent use of “will” to express non-future (e.g. habitual or generic) meanings. Such uses are analogous to functions of the modal auxiliary “hui” in Chinese, suggesting potential L1 influence. In contrast, Japanese learners not only use “will” less frequently, but also consistently omit it in obligatory contexts. The two groups of learners also differ in their use of “will” in conditional sentences. With rising proficiency, “will” becomes increasingly restricted to conditional sentences in essays by Japanese learners, whereas the opposite trend is observed among Chinese learners. Finally, the study considers task-related effects, notably that the convergence on native speaker-like frequency of use is apparent in only one of the two essay topics.

【主要参考文献】

- Bardovi-Harlig, K. (2017). Beyond individual form-meaning associations in L2 tense-mood-aspect research. In M. Howard & P. Leclercq (Eds.), *Tense-aspect-modality in a second language Contemporary perspectives* (pp.27-52). Amsterdam: John Benjamins.
- Nakayama, S. (2021). Contrastive interlanguage analysis of Japanese EFL learners’ modal auxiliary verb use in conversation. *Journal of Educational Research and Review*, 4(1), 1-13.
- Tsai, W. (2015). On the topography of Chinese modals. In U. Shlonsky, (Ed.), *Beyond functional sequence* (pp.275-294). Oxford: Oxford University Press.

Presentation 9

大橋由紀子 (ヤマザキ動物看護大学) / 片桐徳昭 (北海道教育大学) / 押切孝雄 (戸板女子短期大学)

【題目】

授業コーパス構築のための自動タグ付けツール "Classroom Corpus Tagger" の開発【B】

【題目 (英語)】

Developing Classroom Corpus Tagger: A Spoken Language Tagger to Compile Classroom Corpora

【キーワード】

授業コーパス, tagger, reflection, 発話分析

【概要】

本研究では、文字起こしされた各発話に対してタグを自動生成する Classroom Corpus Tagger (CCT) を紹介する。授業コーパス構築には、話者・使用言語・活動等に関するタグ付与を要する (e.g., Ohashi & Katagiri, 2016)。手作業でのタグ付与は時間を要し、ミスが生じやすいことが課題であった。CCT は JavaScript を利用、ブラウザで作動し、日本語か英語を自動的に判別して言語タグを付与する。話者タグは任意に複数種類の設定が可能である。これにより授業コーパス構築が容易となる。手動と CCT でのタグ付けの妥当性の比較実験を行った結果、手動構築で見られるミスは、CCT を利用した場合は見られず、CCT を使用したタグ付けの正確性と、負担軽減が観察できた。発表では、CCT の実演と実験結果の詳細について報告する。

【主要参考文献】

Ohashi, Y., & Katagiri, N. (2016). The effects of explicit instructions observed in teacher transcripts and student impression remarks in elementary school. *HELES Journal* (16), 3-18.

Presentation 10

佐々木恭子(神戸大学大学院生)

【題目】

日本人学習者の英語原因表現使用:ICNALEに基づく量的概観【A】

【題目(英語)】

Use of English causal expressions by Japanese EFL learners: a quantitative overview based on ICNALE

【キーワード】

英語原因理由表現, because 過剰使用, 日本人学習者, 英作文

【概要】

日本人学習者は英作文において because を多用するとされるが(小林, 2009; 佐々木, 2021), 幅広い原因表現を対象にした計量的検証は十分になされていない。そこで本研究は, Altenberg (1984), Biber et al. (1999) などから取捨選択した 34 種の原因表現を対象に, ICNALE の母語話者作文と, CEFR A2 学習者, B1_2/B2+ 学習者作文を比較した。その結果, (1)母語話者は接続詞と動詞で理由を深化する使用傾向, (2)学習者は接続詞と reason など名詞使用により理由を羅列的に表す傾向などの知見が得られた。本研究の知見は, 高校段階での英語教育の改善に一定の示唆を有する。

【主要参考文献】

Altenberg, B. (1984). Causal linking in spoken and written English. *Studia Linguistica*, 38(1).20-69.

<http://dx.doi.org/10.1111/j.1467-9582.1984.tb00734.x>

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education Ltd.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Longman Group Limited.

Presentation 11

小島ますみ(岐阜市立女子短期大学)／金田拓(帝京科学大学)

【題目】

ライティング評価とテキストの特徴との相関関係:メタ分析による研究成果の統合【A】

【題目(英語)】

L2 Writing and its text features: A meta-analysis of correlation coefficients

【キーワード】

第二言語ライティング, ライティング評価, CAF, 結束性

【概要】

本研究は、小島・金田(2020)の研究対象を拡大し、2016年以降の研究や3種類の調整変数を加えたものである。第二言語(L2)学習者のライティング評価とテキストの特徴の相関関係について、103の研究(総参加者15,537人)のメタ分析を行った。結果より、ライティング評価と最も相関が高かったテキストの計量的特徴は流暢性であり、続いて正確性、語彙的複雑性、統語的複雑性、結束性の順であった。主観的な評価項目では、内容と言語使用の効果量が最も大きく、結束性、一貫性が最も小さい結果となった。また、書き手の年齢、母語、学習環境、ライティング評価方法、計量言語指標の種類が有意な調整変数であった。

【主要参考文献】

小島ますみ・金田拓(2020)「ライティング評価と言語的指標の関係—メタ分析による研究成果の統合」石井雄隆・近藤悠介(編著)『英語教育における自動採点—現状と課題』(pp. 33-72) ひつじ書房

Presentation 12

石川有香(名古屋工業大学)

【題目】

工学系大学院生のための教材開発—日英コーパスの分析【C】

【題目(英語)】

A Corpus-based Analysis of Japanese/English Journal Articles in Engineering Fields: Material Development for Students

【キーワード】

工学系論文コーパス、ライティング教材開発、ESP、JSP

【概要】

工学系大学では、大学院生に対しても、英語による論文執筆を求める声が強くなってきている。一方で、学部・大学院において、英語教育に割り当てられた時間は非常に少なく、英語論文執筆の体系的な指導はほとんどなされていない。工学系大学院生を対象とした教材開発が急務となっている。本プロジェクトでは、日英パラレルコーパスを作成し、日本語で研究活動を行っている工学系大学院生が、その研究成果を英語で発表するための教材開発を目指す。すでに、Academic Phrase BankやAWSuMなどすぐれたライティング支援ツールがあるが、英語使用に慣れていない工学系大学院生にはハードルが高くなっている。日本語の分析や日英の提示方法などの課題について考える。

【主要参考文献】

石川有香(2021)『ジャンルとしての工学英語』大学教育出版

Presentation 13

畔元里沙子(九州大学大学院生)

【題目】

CEFR 準拠教科書における英語コロケーションの難易度変化要因の特定【A】

【題目(英語)】

Identifying Difficulty Factors of English Collocations in CEFR-informed Textbooks by Deep L

【キーワード】

CEFR, 教科書コーパス, コロケーション, Deep Learning

【概要】

学習者にとって初級の単語から成る英語コロケーションが必ずしも難易度が低いとは限らないように、英語コロケーションの難易度の測定は難しい(内田, 2015)。そこで本研究は、英語コロケーションの難易度が変化する要因を客観的指標から特定することを目的とする。そのためにまず、自作の CEFR 準拠の教科書コーパスに含まれるレベル別サブコーパスを利用し、そこに一定の基準以上で出現する二語で構成されるコロケーションを抽出、コロケーションのレベル別リストを作成した。その後、統計的手法を用いてレベル間のコロケーションの特徴の違いを分析し、CEFR 準拠の教科書コーパス内でのコロケーションのレベル変化要因を明らかにした。

【主要参考文献】

内田諭. (2015). 基本動詞のコロケーション難易度測定—CEFR レベルに基づくテキストコーパスからのアプローチ—. 言語処理学会年次大会発表論文集, 21, 880-883.

Presentation 14

ISHII Tatsuya (神戸市高専) / KAWAMOTO Takeshi (広島大学)

石井達也 / 河本健

【題目】

N-grams at the Beginning of the Moves in the Results Section of Experimental Medical Research Articles【C】

【キーワード】

Move Corpus. Results Section. N-grams

【概要】

Using a corpus based on move analysis of experimental medical research articles (in total approximately 1.5 million words), Ishii & Kawamoto (2020) focused on the behavior of adverbs and successfully identified 26 lexical phrases for the three moves in the Results section: (RM1) Introducing experiments, (RM2) Announcing results, and (RM3) Commenting results. However, although there are cycles of the three moves, it was still unclear as to how the moves start and are connected. In this study, to identify the n-grams at the beginning of the three moves, we extracted and examined the first sentences of the three moves. After the first sentences of the three moves were extracted by CasualConc (2021) with the use of a wildcard, they were copied and pasted into Excel to divide them into independent words. The frequencies of the n-grams were counted with the help of CasualConc (2021). In conclusion, the observation of the n-grams led to the description of highly frequent phrases for starting and connecting moves; for example, the phrase to determine in (RM1), the phrase we observed in (RM2), and the phrase taken together these results in (RM3). This study will provide new insights for investigating a corpus based on move analysis.

【主要参考文献】

Ishii, T., & Kawamoto, T. (2020). The behavior of adverbs in the results sections of experimental medical research articles: A corpus-based move analysis. *English Corpus Studies*, 27, 23-52.

Presentation 15

清水眞(東京理科大学)／村田真樹(鳥取大学)

【題目】

生化学英語学術論文のための学術語彙リスト【A】

【題目(英語)】

Academic word lists for biochemistry

【キーワード】

学術語彙リスト、基礎単語、ESAP、EGAP

【概要】

Shimizu et al. (2018)は、Hyland & Tse (2007)にならい、有機化学論文誌、物理化学論文誌に掲載された論文からコーパスを編纂し、有機化学論文のための学術語彙リスト、物理化学論文のための学術語彙リストを作成した。この研究では、2016年と2020年に発行された生化学論文誌からコーパスを編纂し、生化学論文のための学術語彙リストをふたつ作成した。ふたつのリストの比較、ふたつのリストと有機化学、物理化学のリストとの比較を行った。どのリストにおいても、基礎単語は約800語しか用いられていないこと、基礎単語であっても、専門的な意味を持つものがあることなどがわかった。

【主要参考文献】

Hyland, K. & Tse, P. (2007). 'Is there an "Academic Vocabulary"?' *TESOL Quarterly* 41:2, 235-253.

石川 慎一郎(2012) 『ベーシックコーパス言語学』ひつじ書房

Shimizu, M., Murata, M, & Ramonda, K. (2018), 'Teaching English for Chemistry at a Japanese University', *The Online Journal of Science and Technology* - July 2018 Volume 8, Issue 3

Presentation 16

堀家利沙(神戸大学大学院生)

【題目】

高校英語指導における句動詞の扱い—教科書とセンター試験の分析から—【A】

【題目(英語)】

An investigation into the usage of phrasal verbs in high school English input sources—Analysis of English textbooks and the National Center Test

【キーワード】

句動詞, 日本人高校生, 検定教科書, 大学入試センター試験

【概要】

日本人高校生の英語アウトプットにおける句動詞運用を分析した研究は多いが, 高校生向けインプット資料における句動詞の実態を計量的に調査した研究は必ずしも多くない。そこで, 本研究は高校教科書 6 種, センター試験 10 年分をコーパス化し, BNC/ COCA と比較した。句動詞の全体使用量, 重要句動詞に対するカバー率, 特徴句動詞の 3 観点で調査した結果, (1) 初中級教科書における頻度は母語話者の基準以下, (2) 重要句動詞カバー率は, 上級教科書, センター試験併用型であっても7割弱, (3) インプット資料の上級化に伴い, 特徴句動詞の質も段階的に変化する, という3点が明らかになった。本研究の知見は, 句動詞指導の改善に一定の意義を持つ。

【主要参考文献】

Gardner, D. & Davies, M. (2007). Pointing Out Frequent Phrasal Verbs: A Corpus-Based Analysis, *TESOL Quarterly*, 41(2), 339-359.

石井康毅(2018)「話し言葉コーパスと検定教科書に基づく日本人英語学習者の句動詞使用実態の分析」*Learner Corpus Studies in Asia and the World*, 3, 101-119.

石川慎一郎(2019)「英語教育における連語:ターゲット・インプット・アウトプットの三元コーパス分析をふまえた English N-gram List for Japanese Learners of English (ENL-J) の開発と利用」『言語分析のフロンティア』金星堂, 32-47.

Presentation 17

中谷安男 (法政大学)

【題目】

オックスフォード・ユニオンにおけるリーダーシップ育成の示唆: 英語圏のリーダーの発話コーパス分析【A】

【題目 (英語)】

How to enhance leadership at Oxford Union: Corpus Data Analysis on Leaders' Speech

【キーワード】

リーダーコーパス, 発話, オックスフォードユニオン, コミュニケーション戦略

【概要】

世界有数のディベート組織であるオックスフォード・ユニオンや、TEDTalk で講演を行った政治・経済界の代表者の発話コーパスを分析することにより、リーダーに必要な聴衆を説得する Communication Strategy (CS) の検証を行った。120 名の発話データをテキスト化し約 50 万語のコーパスデータとして活用した。これを AntConc Windows (3.5.8) を使い FLOB 及び FROWN コーパスの合計 200 万語と比較し Keyword 分析により特徴語を抽出した。さらにこの特徴語のクラスター表現を抜きだし、リーダーたちが活用する CS を確認した。結果として、リーダーは聴衆の注意を喚起し、誘導し、積極的に行動を起こすように促す CS を効果的に活用していた。

【主要参考文献】

- Charteris-Black, J. (2006) *The Communication of Leadership: The Design of Leadership Style*. London: Routledge.
- Fetzer, A., and Bull, P. (2012) Doing leadership in political speech: Semantic processes and pragmatic inferences. *Discourse & Society*, 23- 2, 127-144.
- Kotter, J. P. (1999) *John P. Kotter on What Leaders Really Do*. MA: Harvard Business School Press.

Presentation 18

小谷尚子(東京外国語大学大学院生)／佐野洋(東京外国語大学)

【題目】

金融関連辞典と実務資料コーパスを用いた経済・金融分野の英語語彙リスト研究【C】

【題目(英語)】

A corpus based study of English vocabulary lists in Economy and Finance focusing on financial dictionaries and domain-specific text corpora

【キーワード】

ESP, 経済, 金融, 語彙

【概要】

本研究は、経済・金融分野を対象とした ESP 教育に関する教育的な観点からの語彙カテゴリ化と学習語彙のリスト化を検討している。発表では、辞書の見出し語における重要性と、実務文書における語の出現の有り様の二面から調査した結果を報告する。具体的には、英語及び日本語で出版されている経済・金融分野の用語辞書を対象に、それら辞書の見出し語を比較し、共通語彙を明らかにする。辞書の選択条件は、近年の出版であること、見出し語数が多いこと(5000 語以上)、知名度が高い(流通量が多い)ことである。共通語彙の通用性を確認あるいは検証することを目的として、主要米国企業30社の10K(米国の有価証券報告書に該当)からテキストを取り出し(数字からなる語を含む240万語程度)、使用されている語彙群をもとめ(170万語程度)、上記の共通語彙との重複や頻度傾向を調査した。経済・金融分野における実務上の観点からみた学習語彙のリスト化について、語彙の定性分析も含めて考察する。

【主要参考文献】

Dictionary of Finance and Banking (Oxford University Press, 2014)

『金融関連専門辞書経済・金融ビジネス英和大辞典』(日外アソシエーツ、2012)

Apple Inc. (2019). *Form 10-K 2018*. U.S. Securities and Exchange Commission.

Presentation 19

FUJITA Iku (大阪大学大学院生)

藤田郁

【題目】

LDA Topic Modelling of Tennyson's Poetry【A】

【キーワード】

Alfred Tennyson, LDA, Topic Modelling, Poem

【概要】

Topic modelling is considered a promising approach in text mining (Meeks and Weingart, 2012), and a number of studies have examined prose texts using topic modelling (Tabata, 2017; Kiyama, 2018; Huang, 2020). However, the application of topic modelling to poetry is still developing; this study thus will make a step further towards an in-depth investigation using LDA (Blei et al., 2003) on Alfred Tennyson's poetry works.

The data of this study is a Victorian poet Alfred Tennyson's 66 epic and lyrical poems over 1,000 words in length. In this study, some explicit features to characterize works, such as character names and honorific titles, are excluded from the analysis as stopwords.

Emerging results LDA have shown the latent topics hidden behind prominent elements of poems in the corpus, and the topics appeared in some works in common; of further interest is the latent connections between some works. In addition, this study discusses the possibility of detecting rhyming elements when LDA is run on poetry data as well as the issue of PoS tagging on verse texts, suggested by the results of LDA in hindsight, and conceivable future approaches for addressing the issue.

【主要参考文献】

Blei, M. D., Ng, Y. A., and Jordan, I. M. (2003). "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3, 2003, pp. 993-1022.

Meeks, E. and Weingart, B. S. (2012). "The Digital Humanities Contribution to Topic Modeling." *Journal of Digital Humanities*, vol.2, No. 1 Winter 2012, pp. 1-6.

Tabata, T. (2017). "Mapping Dickens's Novels in a Network of Words, Topics, and Texts: Topic Modelling a Corpus of Classic Fiction." *Japanese Association for Digital Humanities Conference 2017, September 2017, Doshisha University*.

Presentation 20

松田佑治 (立命館大学)

【題目】

have long V-ed 構文の典型例【A】

【題目(英語)】

The typical examples of the have long V-ed Construction

【キーワード】

have long V-ed 構文, have long V-ed 構文の典型例, have long V-ed 構文の V-ed, have long V-ed 構文の主語

【概要】

住吉 (2020) は、have long V-ed 構文の典型例を探るために、Web 版の COCA を対象とし、[have] long [VVN] という検索で、この構文の典型例を探り、複数の特徴を示している。しかし、住吉 (2020) の調査で用いた品詞タグ[VVN] は、確かに過去分詞を抽出するものの、been を抽出しないという致命的問題などが複数見られる。そこで本発表では、COCA-FullText を対象とし、正規表現を用いて、品詞タグに頼らずに過去分詞を抽出するという新たな分析を行った。その結果、新しい知見の中の一つとして、have long V-ed 構文の V-ed には、圧倒的に been が生起することが分かった。

【主要参考文献】

Hilpert, Martin. 2014. *Construction Grammar and its Application to English*, Edinburgh: Edinburgh University Press.

住吉誠. 2020. 「コーパスと英語語法研究」『コーパス研究の展望 (Aspects of English Corpus Studies)』(最新英語学・言語学シリーズ 第 11 巻) 石川慎一郎・長谷部陽一郎・住吉誠 (著). 東京: 開拓社.

滝沢直宏. 2016. 「コーパスからの情報抽出と抽出データの意味づけに関わる諸問題」『英語コーパス研究』23: 45-60.

Presentation 21

佐竹由帆(青山学院大学)

【題目】

英語の動詞-名詞コロケーション学習に対する DDL の効果【A】

【題目(英語)】

The effects of DDL on learning verb-noun collocations in English

【キーワード】

DDL, コロケーション, COCA

【概要】

コロケーションの知識は重要だが、日本の英語学習者の語彙学習は単語単位になりがちであり、コロケーション指導・学習は十分に行われていない。コーパスを参照して学習するデータ駆動型学習(DDL)の語彙学習に対する効果は様々に検証されているため、本研究はDDLの動詞-名詞コロケーション学習に対する効果を検証した。被験者は約20名の日本の大学2年生で英語中級学習者であり、筆者が選んだ2つの動詞-名詞コロケーションを現代アメリカ英語コーパス(COCA)で検索して用例を見る学習を毎週1度10週間行った。事前事後のテスト結果は、ウィルコクソンの符号順位検定で有意差有り、効果量大で($z=3.63$, $p=.000$, $r=.59$)、コロケーション学習に対するDDLの有効性が示唆された。

【主要参考文献】

Shin, D. & Nation, P. (2008) Beyond single words: the most frequent collocations in spoken English, *ELT journal* Vol. 62(4), pp. 339-348

Presentation 22

内田諭(九州大学)

【題目】

チャンクのコロケーション:spaCy を用いた共起分析の試み【C】

【題目(英語)】

Collocation of chunks: A preliminary attempt using spaCy

【キーワード】

コロケーション, チャンク, spaCy, 共起分析

【概要】

コロケーションの集計は多くの場合、単語単位で行われる。この集計方法の場合、higher education, the United States などのまとまりの情報(チャンク)が失われ、重要な共起情報を見逃してしまう可能性がある。本研究では、自然言語処理のアプリケーションを用いてチャンク的情報を保持した形でのコロケーションの集計を試みる。具体的には、Python のライブラリの一つである spaCy を用いて、名詞句のチャンクをタグ付けし、そのテキストデータに対して共起分析を行う。実施手順等を紹介した後、分析手法の妥当性、言語研究における有用性等について議論する。

【主要参考文献】

NA

Presentation 23

AMMA Kazuo (獨協大学)

安間一雄

【題目】

Distribution of repeated appearance of grammar items in junior high school textbooks through nonlinear regression【B】

【キーワード】

grammar item, junior high school textbook, nonlinear regression, factor analysis

【概要】

Opportunities for repeated learning is of vital significance especially in second language acquisition. However, in the Japanese junior high school (JHS) context the beginner-level grammar items are linearly arranged in the curriculum and occasional reviewing of past items seems to be neglected. This study is aimed at characterising the reappearance patterns of grammar items and differentiating them qualitatively, thereby suggesting the teacher's approach to individual items.

38 popular grammar items were selected for analysis in six MEXT-inspected JHS textbooks across three year grades (all published in 2016). The frequency was counted for each appearance printed in the student textbook as well as exercise answers and audio scripts in the teacher's manual; ie., for all exposures either visually or orally presented including repeated exercises.

A cumulative frequency data was collected to which a cubic regression was applied, resulting in high rates of squared residuals for high-frequency items ($R^2=0.95\sim 0.99$). The coefficients of the regression formulae were then used for factor analysis. The distribution of items revealed a new dimension of convex curve patterns and concave curve patterns, indicating how soon or late the items appear and reappear. It also showed textbook-specific patterns as well as universal ones.

【主要参考文献】

AMMA Kazuo. (2018). Extracting patterns from transition of occurrence frequency of grammar items in a junior high school textbook. *Proceedings of the 4th Asia Pacific Corpus Linguistics Conference (APCLC 2018)*, 219-226.

林正頼・石井康毅・高村大也・奥村学・投野由紀夫. (2016). CEFR-based Coursebook Corpus からの CEFR レベル別基準特性の特定. 投野由紀夫(代表)『学習者コーパスによる英語 CEFR レベル基準特性の特定と活用に関する総合的研究』(平成 24 年度~平成 27 年度科学研究費補助金(基板研究(A))研究課題番号 24242017 研究成果報告書).

石井康毅. (2016). CEFR-J Grammar Profile の構築のための英文法項目の選定・抽出・頻度集計・精度評価. 投野由紀夫(代表)『学習者コーパスによる英語 CEFR レベル基準特性の特定と活用に関する総合的研究』(平成 24 年度~平成 27 年度科学研究費補助金(基板研究(A))研究課題番号 24242017 研究成果報告書).

Presentation 24

西垣知佳子(千葉大学)／赤瀬川史朗(Lago 言語研究所)／川名隆行(千葉大学教育学部附属中学校)／中井康平(千葉大学教育学部附属中学校)／見目慎也(千葉大学教育学部附属中学校)／山崎達也(千葉大学教育学部附属中学校)

【題目】

中学生のための Web 版 DDL 支援ツールの開発と活用【A】

【題目(英語)】

Classroom Application of Web DDL Support Tool in a Secondary School

【キーワード】

DDL, teaching-oriented corpus, 気づき, 入門期 英文法

【概要】

DDL は大学生を対象とする活用事例は多いが, 中学生のような入門期では世界的に少ない。その理由には, 入門期レベルに適したコーパスと使いやすい検索ソフトの不足がある。

そこで発表者らは, レベルに配慮した teaching-oriented corpus を作成し, その検索と学習を助ける Web 版 DDL 支援ツールを公開した。本研究の目的は, 本ツールを使って中学生がどのように英語の学びを深めるか調査し, 有効性を確認することであった。国立大学附属中学校の生徒に対して, 誤り訂正タスクと Web 版 DDL ツールを組み合わせた授業を実施した。ワークシートを分析した結果, DDL が教師の期待する文法や語彙の気づきを生徒より引き出したことから, 学習に有効であると確認した。

【主要参考文献】

Crosthwaite, P. (2020). *Data-Driven Learning for the Next Generation: Corpora and DDL for Pre-tertiary Learners*. Routledge.

堀正広, 赤野 一郎 (2015)(監修), 投野 由紀夫 (編) 『英語コーパス研究シリーズ(第 2 巻): コーパスと英語教育』 ひつじ書房。

Timmis, I. (2015). *Corpus Linguistics for ELT: Research and Practice*. Routledge.

Presentation 25

木山直毅 (北九州市立大学) / 渋谷良方 (金沢大学)

【題目】

動詞の意味はトピックから推測できるのか: 英語の動詞 run を例に【A】

【題目(英語)】

Exploring the meaning of 'run' with topic models.

【キーワード】

多義語, Biterm topic model, トピック

【概要】

近年, コーパス言語学では多義語の意味を調査する上でコロケーションや語が現れる統語や形態素といった文法要素や共起語の意味カテゴリーを用いる手法が提案されている (e.g. Gries 2006, Heylen et al. 2012)。本発表では, 上述の要因に加え, 語が現れるトピックが多義語の意味を決める要因になることを提案する。英語の動詞 run は非常に多くの意味を持つが (Gries 2006, Langacker 1988, Taylor 1996), 例えばスポーツの話題で run が用いられれば「走る」の意味で, ビジネスの話題で run が現れれば「経営する」の意味だと考えるのが直感的には自然である。本発表では, News on the Web corpus (Davies 2016-)より抽出した run の事例に対し, トピックモデルの手法の一つである biterm topic model (Yan et al. 2013) を用いて以上の直感を実証する。

【主要参考文献】

- Gries, S. (2006). Corpus-based Methods and Cognitive Semantics: The Many Senses of to run. In A. S. Gries Stefan (Ed.), *Corpora in Cognitive Linguistics Corpus-Based Approaches to Syntax and Lexis* (pp. 57-99). Mouton de Gruyter.
- Davies, M. (2016). Corpus of News on the Web (NOW): 3+ billion words from 20 countries, updated every day.
- Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A Biterm Topic Model for Short Texts. *Proceedings of the 22nd International Conference on World Wide Web*, 1445-1456.

Presentation 26

小林純一郎(東京外国語大学学部生)／佐野洋(東京外国語大学)

【題目】

現代スペイン語における主語後置の数理モデル化【C】

【題目(英語)】

Model Building of Subject Postposition in Contemporary Spanish

【キーワード】

主語後置, 情報構造, 表層特徴, ロジスティック回帰

【概要】

スペイン語の主語後置の発生メカニズムを多変量カテゴリの分類問題に帰着させ, 大規模データ(約 20 億語規模のコーパス)を用いて分類器(ロジスティック回帰モデル)を構築したので, 本発表にて報告する。先行研究では, 情報の新旧などの情報構造に基づいた手法が多く, 情報の新旧などをカテゴリ変数とする線形重回帰分析もある。本研究では大規模データを高速分析すべく, 表層に出現するカテゴリ変数(冠詞の定性など)を用いてモデルを作成した。その結果, 先行研究で示された結果と同等かそれ以上の分類性能を持つことが分かった。スペイン語のみならず英語を含め, 通言語的な情報構造概念が表層特徴から近似されうることを示された。

【主要参考文献】

- Brunetti, L., & Bott, S. (2011). Subject inversion in Romance: a corpus-based study; Handout distributed at: Quantitative Investigations in Theoretical Linguistics QITL-4. <available in <https://edoc.hu-berlin.de/bitstream/handle/18452/2021/brunetti.pdf?sequence=1>>
- Hatcher, A. G. (1956). Theme and Underlying Question: Two Studies of Spanish Word Order. *Word* 12, supplement 3, 1-52.
- Müller, A. C. & Guido, S. (2017). *Introduction to Machine Learning with Python* O'Reilly Media, Inc., (アンドレアス・C・ミュラー & サラ・グイド 中田秀基 (訳) (2017). 『Python ではじめる機械学習: scikit-learn で学ぶ特徴量エンジニアリングと機械学習の基礎』, オライリー・ジャパン)

Presentation 27

ANTHONY Laurence (早稲田大学)

アントニ・ローレンス

【題目】

Introducing AntConc 4.00: A fast, powerful, and easy-to-use corpus analysis tool for small and large-scale corpus analysis and data-driven learning【A】

【キーワード】

AntConc, corpus tools, DDL, KWIC

【概要】

AntConc is a corpus analysis tool that has been repeatedly cited to be the most widely used desktop corpus tool in the world (e.g., Tribble, 2012). It has been downloaded over 2.5 million times by users in over 140 countries and its tutorial videos have been viewed over 500,000 times. While AntConc is a relatively powerful and easy-to-use tool, two of its most commonly cited weaknesses are its speed and handling of medium to large corpora of 10 million words or more. To address these issues, AntConc 4.00 has been completely rewritten on top of a modern database indexing system that scales to corpora of 100 million words and more and allows for results from large corpora to be returned in fractions of a second. Also, the interface has been redesigned to allow for pagination or thinning of large sets of results that are commonly generated with large corpora. In addition, AntConc 4.00 introduces a completely original Key-Word-In-Context (KWIC) concordance view that dramatically simplifies the use and interpretation of this tool. It is anticipated that this new view will become a standard in the field and greatly improve the utility of KWIC concordancing as part of Data-Driven Learning (DDL) approaches.

【主要参考文献】

- Anthony, L. (2021). AntConc (Version 4.0.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>.
- Tribble, C. 2012. Teaching and language corpora: Quo Vadis? 10th Teaching and Language Corpora Conference (TALC). Warsaw, 11th-14th July 2012.

英語コーパス学会第 47 回大会アブストラクトブック

刊行日 2021 年 10 月 2 日

刊行所 英語コーパス学会

事務局 〒819-0395 福岡市西区元岡 744 九州大学言語文化研究院 内田諭研究室気付
E-MAIL:jaecs.hq@gmail.com