

英語コーパス学会

第 **48** 回大会資料

日時：2022年10月1日（土）

会場：Zoom オンライン開催

英語コーパス学会 第48回大会 プログラム

日時 2022年10月1日(土) 9:30-17:40 Zoom開催

開会式 9:30-9:40

ワークショップ: 9:40-10:40

ワークショップ1【Running a Vocabulary Course With Lextutor】(場所: ルームA)

講師: Dr. Tom Cobb (Université du Québec à Montréal)

ワークショップ2【言語データを対象とした KH Coder の活用法】(場所: ルームB)

講師: 樋口 耕一 (立命館大学)

ワークショップ3【初めてのXML】(場所: ルームC)

講師: 永崎 研宣 (一般財団法人人文情報学研究所)

●研究発表第1セッション 言語資源開発 (場所: ブレイクアウトルーム1) 司会: 島津 美和子 (立教大学)

研究発表1: 10:50-11:10

The pre-processing of YouTube transcripts for corpus-based spoken language analysis Christopher Cooper (Rikkyo University)

研究発表2: 11:10-11:30

日英・英日パラレルコーパス検索ツール『パラレルリンク』(Ver.1.20) 仁科 恭徳 (神戸学院大学)
—インターフェース, 検索機能, 活用研究などについて— 赤瀬川 史朗 (Lago NLP)

研究発表3: 11:30-11:50

PEP コーパスプロジェクト: その設計と射程 神原 一帆 (京都外国語大学・立命館大学)
木村 修平 (立命館大学)
近藤 雪絵 (立命館大学)
山下 美朋 (立命館大学)
山中 司 (立命館大学)

●研究発表第2セッション 文法・統語 (場所: ブレイクアウトルーム2) 司会: 森下 裕三 (桃山学院大学)

研究発表4: 10:50-11:10

受動文における様態副詞の生起位置に関する一考察 西村 知修 (石川工業高等専門学校)

研究発表5: 11:10-11:30

起動表現の意味と補部にくる語との関係性に関する考察 藏菌 和也 (神戸学院大学)
—start NP を例に—

研究発表6: 11:30-11:50

学習英文法における as best as possible の位置付け 松田 佑治 (関西外国語大学大学院生)

●研究発表第3セッション 英語教育・中高生 (場所: ブレイクアウトルーム3) 司会: 田畑 圭介 (神戸親和女子大学)

研究発表7: 10:50-11:10

A longitudinal study of fluency based on learner corpus of English conversations Maxim Tikhonenko (Tokyo University of Foreign Studies, Graduate Student)
Keiko Mochizuki (Tokyo University of Foreign Studies)

研究発表8: 11:10-11:30

日本人中学生英語学習者の take の使用と語義指導への提言 澤口 遼 (関西大学第一中学校)

研究発表9: 11:30-11:50

学習者コーパスの品詞連鎖分析による母語の英語構文への影響 山口 一華 (東京外国語大学大学院生)
投野由紀夫 (東京外国語大学)

●研究発表第4セッション ESP (場所: ブレイクアウトルーム4) 司会: 中谷 安男 (法政大学)

研究発表 10: 10:50-11:10

研究論文における英語と日本語のアブストラクトは読み手を意識しているか
浅野 元子 (大阪医科薬科大学)
松田 紀子 (近畿大学)

研究発表 11: 11:10-11:30

高校生のライティングに対する添削者の添削エラー分析
- 時制の誤りに焦点を当てて -
竹下 綾音 (九州大学大学生)

研究発表 12: 11:30-11:50

Development of the Ballet English Corpus (Ver. 1.0) for art major students
Hiroko Usami (Tokai University)

会員総会 11:50-12:20 (場所: メインルーム)

<休憩 12:20-13:10>

●研究発表第5セッション 分析手法 (場所: ブレイクアウトルーム1) 司会: 西村 知修 (石川工業高等専門学校)

研究発表 13: 13:10-13:30

COCAにおける正誤表現と非標準表現について
田畑 圭介 (神戸親和女子大学)

研究発表 14: 13:30-13:50

CasualConc 3.0 - Universal Dependency タグを利用した文法検索の試み
今尾 康裕 (大阪大学)

研究発表 15: 13:50-14:10

量的概念分析再考: 動詞 explain を例に
菅原 裕輝 (大阪大学)
神原 一帆 (京都外国語大学・立命館大学)

●研究発表第6セッション 英語学・社会言語学 (場所: ブレイクアウトルーム2) 司会: 藏菌 和也 (神戸学院大学)

研究発表 16: 13:10-13:30

米国一般教書演説に出現する分離不定詞の効果に関する予備的研究
福本 広光 (大阪大学大学院生)

研究発表 17: 13:30-13:50

男性リーダーの発話データの特徴分析: オックスフォード・ユニオン
及び TED Talk 分析の示唆
中谷 安男 (法政大学)

研究発表 18: 13:50-14:10

憲法をめぐる日本の帝国議会・国会議事録における「翻訳」の用法および
米国議会議事録における translation の用法の分析
島津 美和子 (立教大学)

●研究発表第7セッション 英語教育・大学生 (場所: ブレイクアウトルーム3) 司会: 浅野 元子 (大阪医科薬科大学)

研究発表 19: 13:10-13:30

日本人英語学習者の「うなぎ文」の使用に関する分析
藤原 康弘 (名城大学)
岩男 考哲 (神戸市外国語大学)
伊藤 創 (関西国際大学)
仲 潔 (岐阜大学)

研究発表 20: 13:30-13:50

The use of nominalization features in the academic texts by Japanese learners
Kaede Hanawa (Tokyo University of Foreign Studies, Graduate Student)
Yukio Tono (Tokyo University of Foreign Studies)

- 研究発表 21 : 13:50–14:10
The misuses of English articles in compositions of L1 Chinese and Japanese learners: A corpus-based study Xiao Sun (Kyoto Sangyo University, Graduate Student)
- 研究発表第 8 セッション DDL (場所: ブレイクアウトルーム 4) 司会: 水本 篤 (関西大学)
- 研究発表 22 : 13:10–13:30
初等・中等教育向け DDL ツールが目指す「自律的なコーパスユーザーの育成」西垣 知佳子 (千葉大学)
赤瀬川 史朗 (Lago NLP)
- 研究発表 23 : 13:30–13:50
中学校通常英語授業における DDL の活用を目指して: 学習環境・タイミングの違いによる比較 中井 康平 (千葉大学教育学部附属中学校)
水本 篤 (関西大学)
西垣 知佳子 (千葉大学)
- 研究発表 24 : 13:50–14:10
英作文における動詞-名詞コロケーション産出に対する DDL の効果 佐竹 由帆 (青山学院大学)
- 研究発表第 9 セッション 分析手法・統計 (場所: ブレイクアウトルーム 1) 司会: 今尾 康裕 (大阪大学)
- 研究発表 25 : 14:20–14:40
A critical evaluation of the optimal association measures for creating L2 learners' collocations lists Kohei Fukuda (Tokyo University of Foreign Studies, Graduate Student)
Yukio Tono (Tokyo University of Foreign Studies)
- 研究発表 26 : 14:40–15:00
教科書コーパス分析における推定周辺平均値の有用性について 梶山 達也 (同志社大学大学院生)
ー英語法助動詞の頻度と意味の分析からー
- 研究発表第 10 セッション 英語学・英文学 (場所: ブレイクアウトルーム 2) 司会: 神原 一帆 (京都外国語大学・立命館大学)
- 研究発表 27 : 14:20–14:40
分布意味論の手法を応用したフレーム意味論の分析 森下 裕三 (桃山学院大学)
- 研究発表 28 : 14:40–15:00
Negated speech and thought presentation in contemporary present-tense fiction Reiko Ikeo (Senshu University)
- 研究発表第 11 セッション 英語教育・教材 (場所: ブレイクアウトルーム 3) 司会: 宇佐 美裕子 (東海大学)
- 研究発表 29 : 14:20–14:40
高校英語教科書におけるコロケーション: 基本動詞に着目して 畔元 里沙子 (九州大学大学院生)
- 研究発表 30 : 14:40–15:00
Corpus based research on vocabulary development: Focusing on phrasal verbs composed of A-level verbs and particles Kohei Takebayashi (Tokyo University of Foreign Studies, Undergraduate Student)
Yukio Tono (Tokyo University of Foreign Studies)
- 研究発表 31 : 15:00–15:20
A comparative study on collocations used in Japanese junior high school English textbooks and CEFR-based English coursebooks Noriaki Mikajiri (Tokyo University of Foreign Studies, Graduate Student)
Yukio Tono (Tokyo University of Foreign Studies)

基調講演 15:30–16:20 (場所：メインルーム)

What norms for language learners? A corpus-based research and teaching perspective

司会：Dr. Yasuhiro Fujiwara (Meijo University)

講師：Dr. Gaëtanelle Gilquin (The Université Catholique de Louvain, Belgium)

シンポジウム 16:30–17:30 (場所：メインルーム)

Introducing VOICE 3.0: ELF perspectives for Learner Corpus Research

司会：Dr. Shin'ichiro Ishikawa (Kobe University)

講師：Dr. Marie-Luise Pitzl-Hagin (Austrian Center for Digital Humanities and Cultural Heritage, Austrian Academy of Sciences, Austria)

閉会式 17:30–17:40 (場所：メインルーム)

閉会の辞

田畑 智司 (大阪大学)

学生優秀発表賞授賞式

プログラム (9/1現在)

09:30-09:40	開会式							
09:40-10:40	ワークショップ 1【Running a Vocabulary Course With Lextutor】2【言語データを対象とした KH Coder の活用法】3【初めての XML】							
	Session 1 言語資源開発(発表室1)【司会:島津】		Session 2 文法・統語(発表室2)【司会:森下】		Session 3 英語教育・中高生(発表室3)【司会:田畑】		Session 4 ESP(発表室4)【司会:中谷】	
10:50-11:10	Cooper	【1】The pre-processing of YouTube transcripts for corpus-based spoken language analysis	西村	【4】受動文における様態副詞の生起位置に関する一考察	チホネンコ・望月	【7】A Longitudinal Study of Fluency based on Learner Corpus of English Conversations	浅野・松田	【10】研究論文における英語と日本語のアブストラクトは読み手を意識しているか
11:10-11:30	仁科・赤瀬川	【2】日英・英日パラレルコーパス検索ツール『パラレルリンク』(Ver.1.20) - インターフェース, 検索機能, 活用研究などについて -	藏菌	【5】起動表現の意味と補部にくる語との関係性に関する考察 - start NPを例に -	澤口	【8】日本人中学生英語学習者のtakeの使用と語義指導への提言	竹下	【11】高校生のライティングに対する添削者の添削エラー分析 - 一時制の誤りに焦点を当てて -
11:30-11:50	神原・木村・近藤・山下・山中	【3】PEP コーパスプロジェクト: その設計と射程	松田	【6】学習英文法における as best as possible の位置付け	山口・投野	【9】学習者コーパスの品詞連鎖分析による母語の英語構文への影響	宇佐美	【12】Development of the Ballet English Corpus (Ver. 1.0) for Art Major Students
11:50-12:20	会員総会							
	Session 5 分析手法(発表室1)【司会:西村】		Session 6 英語学・社会言語学(発表室2)【司会:藏菌】		Session 7 英語教育・大学生(発表室3)【司会:浅野】		Session 8 DDL(発表室4)【司会:水本】	
13:10-13:30	田畑	【13】COCAにおける正誤表現と非標準表現について	福本	【16】米国一般教書演説に出現する分離不定詞の効果に関する予備的研究	藤原・岩男・伊藤・仲	【19】日本人英語学習者の「うなぎ文」の使用傾向に関する分析	西垣・赤瀬川	【22】初等・中等教育向けDDLツールが目指す「自律的なコーパスユーザーの育成」
13:30-13:50	今尾	【14】CasualConc 3.0 - Universal Dependency タグを利用した文法検索の試み	中谷	【17】男性リーダーの発話データの特徴分析: オックスフォード・ユニオン及びTED Talk分析の示唆	埴・投野	【20】The Use of Nominalization Features in the Academic Texts by Japanese Learners	中井・水本・西垣	【23】中学校通常英語授業におけるDDLの活用を目指して: 学習環境・タイミングの違いによる比較
13:50-14:10	菅原・神原	【15】量的概念分析再考: 動詞explainを例に	島津	【18】憲法をめぐる日本の帝国議会・国会議事録における「翻訳」の用法および米国議会議事録における translationの用法の分析	孫	【21】The misuses of English articles in compositions of L1 Chinese and Japanese learners: A corpus-based study	佐竹	【24】英作文における動詞-名詞コロケーション産出に対する DDL の効果
	Session 9 分析手法・統計(発表室1)【司会:今尾】		Session 10 英語学(発表室2)【司会:神原】		Session 11 英語教育・教材(発表室3)【司会:宇佐美】			
14:20-14:40	福田・投野	【25】A critical evaluation of the optimal association measures for creating L2 learners' collocations lists	森下	【27】分布意味論の手法を応用したフレーム意味論の分析	畔元	【29】高校英語教科書におけるコロケーション: 基本動詞に着目して		
14:40-15:00	梶山	【26】教科書コーパス分析における推定周辺平均値の有用性について - 英語法助動詞の頻度と意味の分析から -	池尾	【28】Negated speech and thought presentation in contemporary present-tense fiction	竹林・投野	【30】Corpus Based Research on Vocabulary Development; Focusing On Phrasal Verbs Composed of A-Level Verbs and Particles		
15:00-15:20					三河尻・投野	【31】A Comparative Study on Collocations Used in Japanese Junior High School English Textbooks and CEFR-based English Coursebooks		
15:30-16:20	基調講演 What norms for language learners? A corpus based research and teaching perspective 講師: Dr. Gaëtanelle Gilquin, (The Université Catholique de Louvain, Belgium)							
16:30-17:30	シンポジウム Introducing VOICE 3.0: ELF perspectives for Learner Corpus Research 講師: Marie-Luise Pitzl-Hagin (Austrian Center for Digital Humanities and Cultural Heritage, Austrian Academy of Sciences, Austria)							
17:30-17:40	閉会式・学生優秀発表賞受賞式							

【ワークショップ 1】

Running a Vocabulary Course With Lextutor

Dr. Tom Cobb (Université du Québec à Montréal)

The importance of vocabulary knowledge in any type of language course or curriculum is now acknowledged but still not easy to incorporate in a systematic manner. There are few dedicated vocabulary courses, and the ones there are do not match the increasing specialisation of many learner programs. Lextutor has been designed to basically take on the whole job of a dedicated vocabulary supplement from a practical corpus perspective. My workshop will show the main steps in this process, from building a corpus of learning materials, to placement testing, to text selection and adaptation, to assuring a manageable supply of new and appropriate items, to test writing that reflects what learners have actually been exposed to sufficiently to be tested on. The theme is 'corpora for courses' and I will share results from locales where this approach and technology is being deployed.

【ワークショップ 2】

言語データを対象とした KH Coder の活用法

樋口 耕一 (立命館大学)

報告者は、社会調査におけるテキスト分析を当初の目的として、「計量テキスト分析」の方法を提案し、実践のためのツール「KH Coder」を開発しています。本ワークショップでは、この方法とツールを用いて自由回答（日本語）を分析した研究事例を紹介したうえで、方法の特徴について述べます。また、ご紹介した研究事例で分析したデータを使用し、まったく同じ結果をえるための KH Coder 操作をデモ形式で示します。実際の研究データなので「OK」ボタン押せば完了というわけにはいかず、結果を見ながら微調整していく過程も示します。最後に、英語をはじめとする多言語対応状況についてお話しします。

【ワークショップ 3】

初めての XML

永崎 研宣 (一般財団法人人文情報学研究所)

コーパス研究には様々なアプローチがある。近年は自然言語処理技術に頼って超大規模コーパスを処理するという魅力的な手法が広まってきているが、一方で、単語や文章に事前にタグを付けた上でそのタグを利用することで分析を行う手法も未だ健在である。タグを付ける際にはいくつかの手法があり得るが、その一つの手法として XML (Extensible Markup Language) がある。このワークショップでは、XML の概要と応用事例を簡単に紹介した上で、XML タグを活用したコーパスを分析するための基本的な手法について扱う。ツールとしては、Google Colabo で提供される Python を用いることで、参加者の方々が再現・応用しやすいものとなることを目指したい。

【研究発表第1セッション】

【研究発表1】

The pre-processing of YouTube transcripts for corpus-based spoken language analysis

Christopher Cooper (Rikkyo University)

The pre-processing of texts is an important step in any corpus-based research, especially when dealing with internet-based texts, where the data tends to be 'noisy'. In this presentation, the pre-processing steps taken to prepare YouTube transcripts for a multi-dimensional analysis will be described. Examining the texts and keeping the end goal and purpose of the study in mind is essential when deciding what elements of texts should be edited. For this study, the transcripts had the following characteristics that needed to be cleaned: words that do not represent speech (e.g. '[laughter]'), censored words (represented as '[_]'), there were no sentence boundaries or punctuation, and all words were lower case (causing tagging problems for proper nouns and lower case 'i'). Potential solutions to these problems will be discussed including using regular expressions, using a Stanford NLP caseless model to capitalise proper nouns, using an open-source punctuation prediction model that is available as a Python library to add sentence boundaries, and replacing censored words with pseudo words. The accuracy and suitability of the solutions will be reported and any feedback from the audience will be very welcome.

【研究発表2】

日英・英日パラレルコーパス検索ツール『パラレルリンク』(Ver.1.20)ーインターフェース, 検索機能, 活用研究などについてー

仁科 恭徳 (神戸学院大学) ・ 赤瀬川 史朗 (Lago NLP)

仁科・赤瀬川 (2021, 2022) では, 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網羅的に振り返り, 日英・英日パラレルコーパスオンライン検索ツール『パラレルリンク』(Ver. 1.0) に搭載予定であった9種のパラレルコーパスの概要と, それらを再整備する上で施したテキスト処理やアノテーション, 全文検索インデックスの作成, ファイル整理などの一連の作業について詳説した。本発表では, 当該ツールのプロトタイプの具体的なインターフェースや検索機能, 実装されている統計処理, 想定される活用紹介などについて紹介する。

【研究発表3】

PEP コーパスプロジェクト: その設計と射程

神原 一帆 (京都外国語大学・立命館大学) ・ 木村 修平 (立命館大学) ・ 近藤 雪絵 (立命館大学) ・
山下 美朋 (立命館大学) ・ 山中 司 (立命館大学)

本発表の目的は立命館大学で実施されているプロジェクト型必修英語プログラムにおける学生の term paper を集積した学習者コーパスの設計案とその発展可能性を論じることにある。このプログラムでは自由度の高いプロジェクトが奨励される一方で複数教員の指導による表現や書式、評価の一貫性といった課題がある。本研究はこうした課題を解決するための前準備として位置付けることができる。本発表ではコーパスの設計案を論じ、その発展可能性として、このコーパスが(A)教員によるプロジェクト評価の傾向の調査、(B)ライティングの構造研究、(C)科学論文「らしさ」の測定のための基礎データとして働きうることを主張する。

【研究発表第2セッション】

【研究発表4】

受動文における様態副詞の生起位置に関する一考察

西村 知修 (石川工業高等専門学校)

様態副詞は受動文で動詞後方または **be** 動詞と過去分詞の間 (中間位置) に生起できるが、中間位置に生起するのが普通であるとされ (Swan 2016: § 201)、文成立に必須の副詞は動詞後方に生起できないという指摘もある。

(1) She was badly treated / *treated badly. (安藤 2005: 525)

BNC や COCA で観察すると副詞が受動文で中間位置に高い頻度で現れるのは事実のようだが、動詞と副詞の組み合わせによってその頻度は異なり、**be treated badly** も散見される。本発表では「副詞+形容詞 (+名詞)」の結びつきの強弱が受動文における副詞の位置に影響を与えている可能性を探る。

参考文献

安藤貞雄 (2005) 『現代英文法講義』 開拓社, 東京.

Swan, Michael (2016) *Practical English Usage*, 4th ed., Oxford University Press, Oxford.

【研究発表5】

起動表現の意味と補部にくる語との関係性に関する考察 —start NP を例に—

藏菌 和也 (神戸学院大学)

本発表の目的は、起動動詞 **start** の補部に名詞句 (Noun Phrase, NP) がくる表現 **start NP** にどのような語句が生起するかに関して一般化を行うとともに、なぜ特定の性質の語句が生起するかについて、意味的な観点から説明することにある。本調査では、**start NP** に生起する名詞句を *The British National Corpus* を利用して抽出し、手作業で質的に分類した。さらに、*The Corpus of Contemporary American English* を利用して **begin/start +NP** に生起する名詞句を調査した中村 (2018) や類義語の **begin NP** に生起する名詞句を調査した拙稿 (2021) の結果と比較し、さらにネイティブの直観や調査 (Freed 1979) を考慮することで、**start NP** には **start** の使役の意味から「起動後に自分で動いたり、進展していく性質のもの (a car/engine/a team/company/fire/trouble/game/a war)」が頻繁に生起することを主張の柱にすえて議論を進めていく。

【研究発表6】

学習英文法における **as best as possible** の位置付け

松田 佑治 (関西外国語大学大学院生)

as X as 構文の **X** には、通常は原形が要求される。しかし、**as best as possible** や、**as less wrong as possible** など、**X** に最上級や比較級が生起している事例が存在する。そこで、本発表では、大規模なコーパス調査に基づき、**as X as** 構文の **X** に最上級、比較級が生起する事例を正規表現で抽出し、それぞれの頻度を示す。そして、インフォーマントの意見も踏まえ、その分析結果として、現代英語では、**as X as possible** の形に限り、最上級 **best** のみが容認されていることを論証する。その一方、**best** の反対の **worst** や比較級にも、次第に拡張されている点をコーパス調査に基づいて説明する。

【研究発表第3セッション】

【研究発表7】

A longitudinal study of fluency based on learner corpus of English conversations

Maxim Tikhonenko (Tokyo University of Foreign Studies, Graduate Student)

Keiko Mochizuki (Tokyo University of Foreign Studies)

The purpose of this study is to analyze the longitudinal growth of fluency in English conversations by three high school learners. The learner corpus consists of video and audio recordings of 30-minute remote one-on-one speaking lesson with a native English-speaking instructor, recorded over a 20-month period from the first year to the third year of high school. In September 2020, three months after the 20th month lesson, all three students took the Aptis speaking test and their scores were 33 out of 50 points, which was judged as Cefr B level.

The method of fluency analysis was as follows. First, speaking data were manually transcribed and analyzed by ELAN software; second, "speech duration and silent pause length" were measured. Third, the dialogue texts were divided into AS-Units based on Foster, P., A. Tonkyn, and G. Wigglesworth (2000), and fluency in the 10th and 20th month lessons of the three learners was analyzed based on the following criteria.

- 1) Speech rate (words/total time)
- 2) Pause rate (total pause time/speech time)

The results of the speech rate analysis show that all three learners spoke more words per minute and the number of words per minute increased at the 20th month.

Reference

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

【研究発表8】

日本人中学生英語学習者の take の使用と語義指導への提言

澤口 遼 (関西大学第一中学校)

基本語 take の語義の多くは中学校で導入されるが、学習者の習熟度に応じてどの段階で各語義を提示・指導すべきかは十分に検討されていない。そこで本研究では、英語学習者コーパス JEFLL の中学生学習者が使用する take の語義を CEFR レベル別語義データベース English Vocabulary Profile によって分類し、各学年でどの CEFR レベルの語義が使用可能になるのか調査した。その結果、学習者の使用する take の語義はほとんどが A1~B1 レベルで、A2 レベルの語義の使用が少ないこと、学年が上がるにつれて B1 レベルの語義の使用が増加することが判明した。教科書等の教材編集への示唆として、A2 レベルの語義の記述を充実させる、学年を考慮して B1 レベルの語義を導入することが挙げられる。

主要参考文献

投野由紀夫 (編) (2007) 『日本人中高生一万人のコーパス: JEFLL Corpus』 小学館。

Capel, A. (2015). The English vocabulary profile. In J. Harrison. & F. Barker. (Eds.), English profile studies 5, pp. 9-27. Cambridge: Cambridge University Press.

【研究発表9】

学習者コーパスの品詞連鎖分析による母語の英語構文への影響

山口 一華 (東京外国語大学大学院生)

投野由紀夫 (東京外国語大学)

本研究は、第二言語習得における母語の英語構文への影響を調査するため、母語の異なる英語学習者の言語データを収録した学習者コーパス中の品詞連鎖を分析した。特に本研究では、ICCI コーパスのオーストリアサブコーパスと JEFLL コーパスを用いた。ドイツ語母語話者と日本語母語話者の初級~中級 EFL 英語学習者データに出現する単語を品詞タグに変換し、n-gram 形式で品詞連鎖を取り出し、その傾向を調べた。さらに、抽出された n-gram 統計をもとに、英語母語話者コーパスで得た n-gram 統計を基準に、2つの異なる母語の学習者データを比較し、ドイツ語母語話者と日本語母語話者それぞれに特有の品詞連鎖の過剰・過少使用を明らかにする。

【研究発表第4セッション】

【研究発表10】

研究論文における英語と日本語のアブストラクトは読み手を意識しているか

浅野 元子 (大阪医科薬科大学) ・松田 紀子 (近畿大学)

本研究は、論文における英語と日本語の抄録が各々の読み手を意識した言語的特徴を有するかどうかを探索した。2020年1年分の教育心理学研究誌 (EP) 30報と日本公衆衛生雑誌 (JPH) 30報の英日抄録テキストをコーパス化し、修辞の運びを分析した。EPとJPHの総語数は英文が約6,000語と14,000語 (CasualConc)、日本語は約9,000語と16,000語 (KH Coder) であった。双方の約半数は英文と日本語の論理の運びに差異があり、EPは英文で研究の対象、データ収集方法や手順、JPHは様々な部分の英文または日本語で詳述しており、各々の読み手を意識して調整されていることが示唆された。

【研究発表11】

高校生のライティングに対する添削者の添削エラー分析 – 時制の誤りに焦点を当てて–

竹下 綾音 (九州大学大学生)

本研究は、日本の高校生の英作文と添削者の英文を用いてパラレルコーパスを作成し、高校生の中間言語における時制に関する誤用と、それに対する添削者の添削エラーの傾向を分析することを目的とする。これまで学習者の中間言語に関する研究は多く行われてきたが、ライティングに対する添削やその誤りに関する分析はほとんど行われていない。本研究では、日本語を母語としない添削者 (フィリピン、バングラデシュ等在住者) による添削データを対象とし、時制に関する添削誤りを中心に考察する。分析の結果、日本語の表現 (~ている等) に引きずられた高校生の誤りについて特に添削エラーが多いことが明らかになった。

【研究発表12】

Development of the Ballet English Corpus (Ver. 1.0) for art major students

Hiroko Usami (Tokai University)

English for specific fields such as business, law, science technology, medicine, nursing, tourism, aviation has been examined, and specialised corpora for specific English fields have been constructed and applied to teaching. However, English used in the field of art, including music, performing arts, fine arts, and dance, especially classical ballet, has been insufficiently examined.

Various types of dance, including, in particular, classical ballet, have been enjoyed by people of all ages for different purposes in Japan. An increasing number of young Japanese dancers are learning classical ballet both abroad and in Japan and are dancing in overseas ballet companies.

Therefore, this study aims to introduce a specialised corpus, the Ballet English Corpus (BEC), which can be applied to the study of English for specific purposes. The BEC (Ver. 1.0) contains written texts in ten different categories related to classical ballet (ballet techniques, companies, studios, history, schools, theatres, people, narratives, repertoires, and miscellaneous) that are used by both adults and children. This study describes the design and analysis of the word lists in the BEC.

[reference]

Showa Academia Musicae. (2022). Nihon no ballet kyoiku ni kansuru zenkoku chosa - hokokusyo. [Nationwide survey on ballet education in Japan – Report]. Showa Academia Musicae.

【研究発表第5セッション】

【研究発表13】

COCAにおける正誤表現と非標準表現について

田畑 圭介（神戸親和女子大学）

誤用と区別される非標準用法や正用法の判定基準について、COCAで検出される頻度情報をもとに考察を行う。検証対象となる英語表現は *What do you got?/get me to doing/got A confused with someone else* 等である。帰結1：COCAにおいて正用法との比率が15%を超える用法でTV/Movieで一定の使用が確認できる用法は非標準用法としてその存在が認定できる。帰結2：COCAにおいて200以上の用例が確認され、正用法との対立がない表現、あるいは正用法と同程度の頻度と生産性が確認できる用法は、標準用法(=正用法)と認定できる。

【研究発表14】

CasualConc 3.0 - Universal Dependency タグを利用した文法検索の試み

今尾 康裕（大阪大学）

macOS用アプリケーション CasualConc は、基本的なコーパス分析ツールが備えている KWIC 検索、単語リスト作成、コロケーション検索などの機能に加えて、統計環境 R を利用したグラフ作成機能を備える統合的コーパス分析環境である。今春公開した新しいバージョンでは、XML タグを利用した簡易的な検索機能やファイルのフィルタリング、Stanford CoreNLP による Universal Dependency タグを利用した文法検索などの機能を追加した。本発表では、新機能を紹介するとともに開発の課題なども議論する。

参考文献 Imao, Y. (2002). CasualConc (Version 3.0.3) [Computer software].

【研究発表15】

量的概念分析再考: 動詞 explain を例に

菅原 裕輝（大阪大学）・神原 一帆（京都外国語大学・立命館大学）

近年、科学哲学分野において、内省に基づく分析に対する方法論批判を受け[1]、量的手法を組み込んだ概念分析の開発が行われてきている[2]。本研究では、量的な科学哲学の方法論の可能性と限界について、特に共起語情報のみに基づく量的な手法を批判的に検討し、量的な認知意味論の方法を組み入れた量的科学哲学の方法を提案・実行する。この目的の下、BNC から無作為に抽出した 657 件の動詞 explain を含む事例に対し、フレーム意味論的な解析に加え[3]、様々な構文解析[4]等を行った。この結果、動詞 explain が喚起する概念構造をより緻密に表すだけでなく、量的分析によって先行研究での結果の比較が可能となった。

[1] Overton, J. A. (2013). "Explain" in scientific discourse. *Synthese*, 190(8), 1383-1405.

[2] Pence, C. H., & Ramsey, G. (2018). How to do digital philosophy of science. *Philosophy of Science*, 85(5), 930-941.

[3] Fillmore, C. J., Johnson, C. R., & Petruck, M. R. L. (2003). Background to FrameNet. *International Journal of Lexicography*, 16(3), 235-250.

[4] Gries, S. T. (2010). Behavioral profiles: A fine-grained and quantitative approach in corpus-based lexical semantics. *The Mental Lexicon*, 5(3), 323-346.

【研究発表第6セッション】

【研究発表16】

米国一般教書演説に出現する分離不定詞の効果に関する予備的研究

福本 広光 (大阪大学大学院生)

本研究は、歴代米国一般教書演説 (State of the Union Address) のスクリプトをコーパスとして、そこに現れる分離不定詞(split infinitive, 下記例文参照)の効果および使用意図の一端を探ろうとするものである。

(例) We are working to finally end America's longest war and bring our troops back home. (Donald Trump, 2020)

The American Presidency Project 所収の演説原稿に品詞タグを付与し、AntConc でコンコーダンス検索を行って用例を抽出した。分離不定詞は時代や人物ごとの揺れはあるものの米国一般教書演説においても一定数用いられており、本発表では、"splitter" (to 不定詞を「割る」副詞相当語句)を軸として、収集した全用例を分類し、大統領が分離不定詞を使うことで強調していると考えられる事柄について、共起表現にも気を配りつつ、幾つかのコンテキストを分析する。

【研究発表17】

男性リーダーの発話データの特徴分析：オックスフォード・ユニオン及び TED Talk 分析の示唆

中谷 安男 (法政大学)

今や性別によるリーダーの役割の差をつけるべきではない。この点に配慮しつつ、本発表は、ビジネス・政治・社会活動に関わる男性リーダーの英語発話における語彙的特徴を調査した。世界のリーダーを多く輩出しているオックスフォード・ユニオンにおけるディベートやスピーチのコーパスデータを作成した。これに加えて、TED Talk のリーダーのコーパスを作成した。合計 164 人のデータを AntConc 4.0.10 を活用し、Keyword 分析により、男性リーダーの特徴語を抽出した。結果として you, get, go, this 等が特徴語彙として抽出された。本報告では、これらのクラスター分析も活用し特徴的な表現の具体例を示す。

【研究発表18】

憲法をめぐる日本の帝国議会・国会議事録における「翻訳」の用法および
米国議会議事録における translation の用法の分析

島津 美和子 (立教大学)

現在、日本政府が進める改憲の理由の一つに、現行憲法は GHQ から提示された英語原案を日本語に翻訳したものであり、日本が策定したものではないという主張がある。この主張の言語的妥当性を検証するため、本研究では、帝国議会・国会議事録をコーパスとみなし、憲法の文脈における「翻訳」の用例を抽出し、分析した。その結果、「翻訳」は「翻訳調」「翻訳臭 (い)」といった特定の句で用い、否定的なニュアンスを伴うことが分かった。一方、米国議会議事録から translate/translation と Constitution が共起する事例を抽出し、分析した結果、合衆国憲法の他言語訳は肯定的に受け止められていることが分かった。

【研究発表第7セッション】

【研究発表 19】

日本人英語学習者の「うなぎ文」の使用に関する分析

藤原 康弘 (名城大学) ・ 岩男 考哲 (神戸市外国語大学) ・ 伊藤 創 (関西国際大学) ・ 仲 潔 (岐阜大学)

本発表の主たる目的は、いわゆる「うなぎ文」が、日本人英語学習者による英語に確認されるか、また「英語力」の伸長と共にどのような変化がみられるかをコーパス言語学的手法で検証することである。日本語は名詞／形容詞述語文に「僕は珈琲だ」「今日は忙しい」のように解釈においてコンテキストに依存する部分が多い「うなぎ文」と呼ばれる表現がある。第二言語習得や英語教育の先行研究では、日本人英語学習者の英語に一定程度確認されることが指摘されてきたが、能力の伸長による変化の詳細はまだ明らかではない。本研究では、International Corpus Network of Asian Learners of English (Ishikawa, 2013) を用いて、うなぎ文の一部の使用傾向の分析結果を報告し、広く議論を行う。

【研究発表 20】

The use of nominalization features in the academic texts by Japanese learners

Kaede Hanawa (Tokyo University of Foreign Studies, Graduate Student)
Yukio Tono (Tokyo University of Foreign Studies)

Since the introduction of the term “grammatical metaphor” by Halliday (1985), nominalization features have caught attention in the research of academic written texts (Biber & Gray, 2013). The present study aims to provide a descriptive view of nominalization features found in the academic writing produced by Japanese L2 learners of English in comparison with L1 users of English.

Corpora used in this study are ICLE-JP for Japanese learners and LOCKNESS for L1 users. Sketch Engine is used for the collocation search. Using the CQL function, all the sentences which include “noun + preposition” are selected for the further annotation phase. Those items are then annotated and categorized under several types of nominalizations. The two sets of data are compared in terms of the type frequency and the verbs (both intransitive and transitive verbs) or adjectives which the nominalizations originate from.

References

Biber, D., & Gray, B. (2013). *The Verb Phrase in English: Nominalizing the verb phrase in academic science writing*.
Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.

【研究発表 21】

The misuses of English articles in compositions of L1 Chinese and Japanese learners: A corpus-based study

Xiao Sun (Kyoto Sangyo University, Graduate Student)

English articles belong to a category of high frequency words. But research on the article system has been mainly focused on the analysis of functions and description of usage, with few studies focusing on the distribution and features of English article errors (Zhou, Xia, Du & Yan-xia, 2015). At the same time, corpus-based research on English articles has also been limited. Hence, this study analyzes misuses of English articles made by L1 Chinese and Japanese learners of English using written corpora. Data is collected from the Nagoya Interlanguage Corpus of English Reborn (NICER) and the Chinese Learner English Corpus (CLEC). In these corpora a wide range of variability in misuse can be found. As 和泉 et. al (2004) have pointed out, the misuses of English articles can be categorized into three patterns: omission of articles, substitution of articles and over-employment of articles, but it still can be predicted that there are some differences existing between these groups of learners in their use of English articles. It is hoped that the findings of this comparative study can have a positive effect on the article acquisition education for L1 Chinese learners and L1 Japanese learners.

【参考文献】

Zhou, Xia, Du, & Yan-xia (2015). *An Investigation of the Misuse of English Articles of Chinese English Learners*.
和泉絵美・内元清貴・井佐原均 (2004). 日本人英語学習者の英語冠詞習得傾向の分析. 『日本人 1200 人の英語スピーキングコーパス』 東京:アルク

【研究発表第 8 セッション】

【研究発表 22】

初等・中等教育向け DDL ツールが目指す「自律的なコーパスユーザーの育成」

西垣 知佳子 (千葉大学) ・赤瀬川 史朗 (Lago NLP)

発表者らは小学生向けと中高生向けの「英語用例コーパス」と「DDL ツール」をそれぞれ開発し、初等・中等教育の英語授業における DDL (data-driven learning) の普及を目指し、学校現場で実践を行っている。発表者らの DDL 開発では、短期・中期・長期の 3 つの目標を設定している。最終的な到達点となる長期目標には「小学生の頃からコンコーダンスラインに慣れ親しんだ自律したコーパスユーザーの育成」を掲げている。本発表ではこの 10 年間にわたる DDL ツールの開発を振り返り、長期目標の「コンコーダンスラインに慣れ親しむ」ことを目指した中高生向けの DDL ツールの新たな機能について報告する。

【研究発表 23】

中学校通常英語授業における DDL の活用を目指して：学習環境・タイミングの違いによる比較

中井 康平 (千葉大学教育学部附属中学校) ・水本 篤 (関西大学) ・西垣 知佳子 (千葉大学)

本研究では、中学 1 年、2 年、3 年の約 260 名が、自主開発した中高生用 DDL 支援ツールを使って、自宅と学校の異なる学習環境で、また、予習、復習、あるいは教科書学習の補助という 3 種のタイミングで英文法を学習した後、質問紙に回答した。質問紙には 5 件法と自由筆記の調査があった。分析の結果、DDL の使用環境について自宅と学校では勉強のしやすさ等の点で差が見られなかった。また、発見活動にかかる時間の節約を目標にして行った教科書学習の補助の DDL が、生徒のわかりやすさにつながるという結果が得られた。これらの結果から、中学校の通常英語授業に恒常的に DDL を導入する場合の可能性、そして問題点を議論する。

【研究発表 24】

英作文における動詞-名詞コロケーション産出に対する DDL の効果

佐竹 由帆 (青山学院大学)

コロケーションの知識は重要だが、英語学習者にとってコロケーションを正確に使用することは難しい。筆者の先行研究 (佐竹 2021) はコーパスを参照して学習するデータ駆動型学習 (DDL) が動詞・名詞コロケーションを覚える上で有効であることを示唆したため、本研究は英作文における DDL の動詞-名詞コロケーション産出に対する効果を検証した。被験者は 20 名の日本の大学 1 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 Wordbanks Online で検索して用例を見る学習を毎週 1 度 9 週間行った。被験者が学習の前後に書いた英作文を比較した結果、事前英作文ではほとんど使用されなかった対象コロケーションが事後の英作文では使用されており、コロケーション産出に対する DDL の有効性が示唆された。

【研究発表第9セッション】

【研究発表 25】

A critical evaluation of the optimal association measures for creating L2 learners' collocations lists

Kohei Fukuda (Tokyo University of Foreign Studies, Graduate Student)

Yukio Tono (Tokyo University of Foreign Studies)

Association measures (AM) have been used to extract candidates for the collocations list for L2 learners. However, little attention has been paid to which AM is optimal for extracting collocations for pedagogical purposes. Furthermore, dispersion measures are often not considered for the evaluation of AMs. This study explores which AM is suitable for identifying pedagogically useful English collocations, considering both frequency and dispersion thresholds. Verb-Object and Modifier-Noun pairs were extracted from the syntactically parsed British National Corpus, and their AMs (MI, MI3, T-score, Z-score, logDice, Log-likelihood) were calculated. Dispersion was calculated using Gries' DP. AMs were evaluated against the gold standard, defined as the items found in CEFR-based English coursebooks and Oxford Collocations Dictionary (OCD). Precision and recall were evaluated by setting varying frequency and DP thresholds. The main results are as follows: (a) without DP threshold, T-score is the most closely associated with the selection of collocations in the gold standard; (b) moderate DP threshold can enhance the performance of AMs, (c) in most cases, DP threshold ≤ 0.4 has an adverse effect on the performance of AMs; and (d) frequency thresholds cannot enhance the performance of Log-likelihood, but DP threshold can make it surpass T-score in some cases.

【研究発表 26】

教科書コーパス分析における推定周辺平均値の有用性について —英語法助動詞の頻度と意味の分析から—

梶山 達也 (同志社大学大学院生)

ここでは、教科書コーパス分析に対する推定周辺平均値の有用性について発表する。現在、大規模コーパス分析にはカイ二乗検定がよく使われている。しかし、教科書コーパスのような規模の小さいコーパスに対して、大規模コーパス分析で使用されるカイ二乗検定を用いると、正しく検定されない可能性が指摘されている。そこで、2022年3月開催英語コーパス学会語彙研究会で発表した、英語助動詞の頻度と意味に関する中・高英語教科書コーパスの分析結果を批判的に考察しながら、新たに、推定周辺平均値を用いた分析を行い、これの有用性について議論する。

【研究発表第 10 セッション】

【研究発表 27】

分布意味論の手法を応用したフレーム意味論の分析

森下 裕三 (桃山学院大学)

本研究では、land と ground という名詞，および stingy, thrifty, generous, wasteful という形容詞を対象として認知意味論が想定する意味観に基づいた分析 (e.g. Fillmore 1982) をおこなう。本研究では、COCAにおいて、上記の語が生起する全用例を対象として分布意味論 (e.g. Montes and Heylen 2022) の手法に基づく分析をおこなった。結果として、認知言語学で主張されてきた land と ground という語の対称性が確認されたものの、stingy と thrifty という語の関係，ならびに generous と wasteful という語の関係については、先行研究において主張されていた以上に複雑な反義関係・類義関係にあることを示す結果を得た。

参考文献

Fillmore, Charles J. (1982) Frame semantics. In The Linguistic Society of Korea (ed.), *Linguistics in Morning Calm*. 111–137. Seoul: Hanshin Publishing Company.

Montes, Mariana and Kris Heylen (2022) Visualizing Distributional Semantics. In Dennis Tay & Molly Xie Pan (eds.), *Data Analytics in Cognitive Linguistics. Methods and Insights*. 103-136. Berlin: Mouton De Gruyter.

【研究発表 28】

Negated speech and thought presentation in contemporary present-tense fiction

Reiko Ikeo (Senshu University)

This paper shows how negated speech and thought presentation functions in narrative by using a corpus approach. A corpus consisting of texts from contemporary present-tense fiction was annotated with discourse presentation categories based on the Semino and Short model (2004). In this annotation process, a new subcategory ‘g’ has been introduced for negated discourse presentation. The data shows that negated thought presentation occurs more frequently than negated speech presentation. The subcategory “g” is attached to 29 out of 3,261 speech presentation cases, which accounts for 0.9% of all the speech presentation tags. In contrast, this subcategory is found in 150 cases out of 1,805 thought presentation cases, which accounts for 8.3% of thought presentation tags.

Negated cases of speech and thought presentation reveal characters’ inner worlds in a way that affirmatives do not. In contrast to events and states which are expressed by means of affirmative terms, “non-events” and “non-states” which are expressed by means of negatives are usually less salient and less informative. When non-events or non-states are expressed in narrative, they stand out and thus have special textual effects.

Reference

Semino, E. and Short, M. (2004) *Corpus Stylistics: Speech, Writing and Thought Presentation in a Corpus of English Writing*. London: Routledge.

【研究発表第 11 セッション】

【研究発表 29】

高校英語教科書におけるコロケーション：基本動詞に着目して

畔元 里沙子（九州大学大学院生）

本研究は、高校英語教科書で扱う基本動詞を含むコロケーションをリスト化し、一般コーパスにおけるコロケーションと比較することを目的とする。具体的には、高校英語教科書コーパスから基本動詞を含むコロケーションを抽出し、一般コーパスでの T スコア・頻度等を付与して重要度・出現率等を評価する。さらに、教科書コーパスを学年別に分類し、学年ごとのコロケーションの相違や、一般コーパスにおけるコロケーションカバー率の変化を分析する。

【研究発表 30】

Corpus based research on vocabulary development: Focusing on phrasal verbs composed of A-level verbs and particles

Kohei Takebayashi (Tokyo University of Foreign Studies, Undergraduate Student)

Yukio Tono (Tokyo University of Foreign Studies)

Despite the significance of phrasal verbs (PVs) in communication, learners have a noticeable tendency to avoid PVs in favour of one-word equivalents. The author argues that PVs would serve as significant building blocks to develop learners' vocabulary knowledge. To this end, this study explores the possibility of using PVs as a bridge between A-level verbs and B-level verbs as defined by the CEFR. A corpus of PV textbooks (size = 1.2 million) was compiled, and pairs of PVs and their single-word verb (SV) equivalents were retrieved. After producing a list of [PV – SV] pairs, the vocabulary levels of those verbs on the list were identified in accordance with the English Vocabulary Profile to investigate the extent to which PVs can be replaced with their SV counterparts. At the same time, their semantic opacity and the degree of semantic transformation between PVs and their SV equivalents were examined. The results show that PVs have the potential to serve as a bridge between A-level and B-level verbs, and a selected group of PVs will make a significant impact on the expansion of the range of meaning related to verb semantics.

【研究発表 31】

A comparative study on collocations used in Japanese junior high school English textbooks and CEFR-based English coursebooks

Noriaki Mikajiri (Tokyo University of Foreign Studies, Graduate Student)

Yukio Tono (Tokyo University of Foreign Studies)

Collocations are arguably one of the most important points for learning a language. The present study investigates the use of English collocations by comparing authorized English textbooks published in Japan and English native corpora in order to improve the contents of English textbooks. To this end, three corpora were used in this study: a corpus of English textbooks for junior high school students, a corpus of CEFR-based English coursebooks published in the UK, and the British National Corpus (BNC) as a reference corpus. From each textbook corpus, a list of verb-object collocations was extracted for each CEFR level via Sketch Engine. For each extracted collocation list, frequency and association measures such as t-score, MI score, logDice, Log-likelihood were obtained using the BNC. Based on those results, extracted collocations used in one corpus but not the other were compared. I also analyzed how the collocation lists would differ when evaluated by each association measure. The identification of collocations and association measures that are markedly different from the collocations used in the textbook and coursebook corpora will allow us to find other collocations to learn from the BNC and other native speaker corpora and make suggestions for textbook improvement.

【講演】

What norms for language learners? A corpus-based research and teaching perspective

Dr. Gaëtanelle Gilquin (The Université Catholique de Louvain, Belgium)

In both research and teaching, norms are often used as a reference against which to describe or evaluate learner language (L2). In learner corpus research, for example, the notions of ‘underuse’ and ‘overuse’ imply a comparison of frequencies in the L2 and in some variety corresponding to the expected target (see Granger 1996). In language teaching, textbooks mostly provide materials representing native varieties of the language and teachers often assess learners’ proficiency with reference to some native standard. Corpora, by giving access to data that are representative of a certain language or language variety, can serve as a basis to define a norm empirically (Klippel & Mukherjee 2007).

This presentation will provide an overview of the different types of corpus-derived norms that can be used in learner corpus research and foreign language teaching. We will consider whether a norm is always required and will examine the use of native vs non-native norms, novice vs expert norms, single vs multiple norms and research vs pedagogical norms, among others. It will also be shown that using different norms can lead to different results (see Chen 2013 for an example), which means that it is crucial to choose the most appropriate norm(s). In this respect, the importance of context will be underlined, including one’s research purposes (cf. Ädel 2006) and learners’ goals in studying the language.

It will be argued that norms are useful in most L2 research and teaching contexts but should be carefully chosen to reflect the most relevant target languages/varieties and should be applied with some flexibility to allow, for example, for linguistic creativity.

References

- Ädel, A. (2006). *The Use of Metadiscourse in Argumentative Texts by Advanced Learners and Native Speakers of English*. Amsterdam: John Benjamins.
- Chen, M. (2013). Overuse or underuse: A corpus study of English phrasal verb use by Chinese, British and American university students. *International Journal of Corpus Linguistics* 18(3): 418-442.
- Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg, & M. Johansson (eds) *Languages in Contrast. Text-based Cross-linguistic Studies* (pp. 37-51). Lund: Lund University Press.
- Klippel, F. & Mukherjee, J. (2007). Standards and norms in language description and language teaching: An introduction. In S. Volk-Birke & J. Lippert (eds) *Anglistentag 2006 Halle, Proceedings* (pp. 303-306). Trier: Wissenschaftlicher Verlag Trier.

Introducing VOICE 3.0: ELF perspectives for Learner Corpus Research

Dr. Marie-Luise Pitzl-Hagin (Austrian Academy of Sciences, Austria)

Since the first online release of the Vienna-Oxford International Corpus of English (VOICE) in 2009, thousands of users world-wide have used VOICE for linguistic research and university teaching to explore the nature spoken English as a lingua franca (ELF) interactions. This talk introduces the new version of VOICE – VOICE 3.0 – built in the VOICE CLARIAH project (2020-2021) and released in September 2021. It provides a tour of the new open-access interface VOICE 3.0 Online that was developed by an interdisciplinary team at the Austrian Centre for Digital Humanities and Cultural Heritage (Austrian Academy of Science) and the University of Vienna. During the talk, we will have a look at key features of VOICE 3.0 Online like its enhanced filter, style and search functions. In doing so, we will consider the expanded research potential of VOICE 3.0. New technical features include for instance being able to search for select spoken mark-up (such as pauses and overlaps) and being able to combine lexical and mark-up queries flexibly with queries for part-of-speech (POS) tags. The talk will also offer a brief glimpse at the technology used to create the new frontend and backend infrastructure for VOICE 3.0.

Following the rationale of VOICE as an ELF corpus, the second part of my talk will then propose and map out some ELF perspectives for Learner Corpus Research (LCR). Here, we will specifically concern ourselves with the issue of corpus annotation and highlight differences in approach and practice between ELF research and LCR. In order to make tangible these differences, the talk will, among other things, report on an ongoing research project (see Riegler in press) that develops a mark-up scheme for annotating pragmatic functions in VOICE. A major difference between an ELF approach and an LCR approach to corpus annotation lies in the way the two fields address the issue of norms and normativity. While LCR continues to orient quite strongly to the norms of the so-called target language (e.g. error tagging), ELF research promotes a very different orientation to linguistic norms. From an ELF perspective, norms are always flexible and seen as situationally negotiated and negotiable by participants in a particular interaction. This orientation in ELF research has clear implications for developing pragmatic corpus annotation and, as a next step, direct consequences for teaching implications and applications that can be drawn developed on the basis of such annotation.

References:

- Riegler, Stefanie. Accepted/forthcoming. Annotating VOICE for pedagogic purposes: the case for a mark-up scheme of pragmatic functions in ELF interactions. In Harrington, Kieran & Patricia Ronan. *Demystifying corpus linguistics for English language teaching*. London: Palgrave Macmillan.
- VOICE. 2021. The Vienna-Oxford International Corpus of English (version VOICE 3.0 Online). Founding director: Barbara Seidlhofer; Principal investigators VOICE 3.0: Marie-Luise Pitzl, Daniel Schopper; Researchers: Angelika Breiteneder, Hans-Christian Breuer, Nora Dorn, Theresa Klimpfinger, Stefan Majewski, Ruth Osimk-Teasdale, Hannes Pirker, Marie-Luise Pitzl, Michael Radeka, Stefanie Riegler, Barbara Seidlhofer, Omar Siam, Daniel Stoxreiter. <https://voice3.acdh.oeaw.ac.at> (last accessed 24 August 2022).

JAECs48 発表についての注意事項

今回の大会では Zoom を使用します。Zoom の操作については、様々な解説サイトがありますので適宜ご利用ください。

<https://support.zoom.us/hc/ja/articles/206175806-Top-Questions>

■事前準備

- ・ Zoom でパワーポイント（とくにアニメーション機能）を使うと、タイムラグが発する場合があります。可能な場合はパワーポイントに代えて PDF などの使用もご検討ください。
- ・ スライドで使用する画像などについてはご自身で肖像権や著作権への対応をお願いします。
- ・ 当日の発表のために、安定したネット環境と、騒音のない静かな環境を事前に確保くださるようお願いいたします。
- ・ Zoom 上での画面共有をご経験でない方におかれては、当日までに、ご自身で、基本的な操作手順について確認して下さるようお願いいたします。

■発表室への入室と発表の準備

- ・ ご自身の発表時間の「5分前までに」（厳守）発表室に入室ください。
- ・ 司会が発表者を認識できるように、ご自身の名前の前に、発表番号（01～31の通し番号。プログラム、大会資料に記載）をつけてください。（例：21 岡田花子（ABC大））
- ・ 発表開始時間の約1分前になると、司会が発表準備を指示します。発表者は、マイクをオンにして、「画面共有」を開始し、スタンバイしてください。
- ・ 発表時に、音声ファイルや、音声付き動画ファイル（発表の前撮り動画も含む）を再生される場合は、画面共有時に「音声を共有する」のチェックをお忘れなくお願いします。
- ・ Zoom をスピーカービューにしたり、発表者をピン止めしたりする作業は、必要に応じて、発表者が自分で行って下さい。

■発表時間・質疑時間

- ・ 質疑込みで20分です。終了時間はくれぐれも厳守をお願いします。
- ・ 質疑時間の長さは、発表者自身が決められます。標準は「発表15分+質疑5分」です。
- ・ ご自身の発表タイプや聴衆の数などを考慮し、発表時間と質疑時間の割合をお決めいただけますが、短くても質疑の時間をとっていただくようお願いいたします。
- ・ （実行委員会側の責任によるものでない限り）開始が遅れたり、発表中にネットトラブルが起こったりした場合も終了時間は延長できません。
- ・ 開始後「21分」を超えた場合は、司会側で強制終了を行います。次の発表者の迷惑にならないよう、時間厳守を重ねてお願いいたします。

■発表時の資料配布

- ・ 帯域への負荷を軽減するため、発表時に資料（PDF、パワポなど）をチャット機能で送信することは禁止とさせていただきます。
- ・ 資料を配布されたい場合は、たとえばチャットで自身のメールアドレスを送り、スライド希望者にメールで連絡させるなどのご対応を検討ください。

■司会

- ・ 司会は原則としてタイムキーピングとトラブル対応のみを担当します。
- ・ 質疑開始のアナウンス、聴衆への質問の依頼、質問方法の指示（チャット、挙手機能、直接の声出し etc.）、質問への回答などは発表者自身をご対応下さい。
- ・ チャットにせよ、声出しにせよ、1人が多くの質問を一度にすると時間内に処理できなくなりますので、発表者のほうから、事前に「短く、1問のみで」のような指定をなさることをお勧めします。
- ・ 実行委員会では、別途、専用の質問受付フォームを用意しています。時間内にできなかった質問やコメントはフォームに書くよう参加者に案内します。フォームに書かれた質問は、大会終了後、発表者各自にまとめてお送りしますので、可能な範囲で、該当者にメールで回答ください。

■万一のトラブルが生じた場合の対応

- ・ 実行委員会側の責任によるトラブルで、万が一、研究発表をしていただけなくなった場合は、大会終了後、発表ビデオをご提出いただき、参加者に期間限定で公開して閲覧していただく対応をとります。この措置により、学会として、当日の発表が成立したものと扱います。

■録音録画

- ・ 上記の場合を除き、発表の録音・録画・後日の公開などは予定しておりません。

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

会長 田畑智司 事務局 〒501-0192 岐阜市一日市場北町7番1号岐阜市立女子短期大学英語英文学科 小島ますみ研究室気付

e-mail: jaecs.hq@gmail.com

URL: <http://jaecs.com/>
