

英語コーパス学会 Newsletter No. 84

July 25, 2018

■会長 投野 由紀夫
■事務局 〒157-8511 東京都世田谷区成城6-1-20 成城大学社会イノベーション学部 石井康毅研究室気付
■郵便振替口座 00930-3-195373 (英語コーパス学会)
■URL: <http://jaecs.com/> ■e-mail: jaecs.hq@gmail.com

JAECS
Japan Association for English Corpus Studies

■ 春季研究会報告

2018年4月21日に東京外国語大学(府中キャンパス)にて、英語コーパス学会2018年度春季研究会が開催されました。2017年度に誕生した英語コーパス学会の5つの研究会(SIG; Special Interest Group), ESP研究会, ツールと統計手法研究会, DDL研究会, コーパスとCEFR研究会, 語彙研究会のお披露目となる5つのシンポジウムが行われ、活発な議論が交わされました。

最初に行われた全体会では、投野由紀夫会長からSIGの設立趣旨の説明があった後、各SIGの代表者がSIGに関する簡単な紹介をしました。その後、参加者は各SIGが主催するシンポジウムに分かれて参加しました。SIGが主催する各シンポジウムでは、1時間の持ち時間を工夫して、SIGの紹介や研究発表がなされました。

以下に、各SIGが実施したシンポジウムの概要を記します。

◇ ESP研究会シンポジウム

テーマ: 「ESP研究におけるコーパス利用」

石川有香先生(名古屋工業大学)の司会のもと、(1)「ESP研究とコーパスの接点」野口ジュディー先生(神戸学院大学)、(2)「医療系ESP教育におけるコーパス利用」藤枝美穂先生(大阪医科大学)、(3)「工学系ESP研究におけるコーパス利用」川口恵子先生(芝浦工業大学)、(4)「工学系ESP語彙教育におけるコーパス利用」石川有香先生(名古屋工業大学)の4つの発表が行われました。

◇ ツールと統計手法研究会シンポジウム

テーマ: 「価値あるコーパスツールと統計手法とは何か」

SIGの代表者であるアントニローレンス先生(早稲田大学)より、(1)「SIG設立の趣旨と方針」、(2)「価値あるコーパスツールとは何か」、(3)「価値ある統計手法とは何か」に

関する説明がなされました。また、説明後にはオープン・ディスカッションが行われました。

◇ DDL研究会シンポジウム

中條清美先生(日本大学)司会のもと、(1)「The concepts of Contemporary DDL and Classic DDL」Gregory Hadley先生(新潟大学)、(2)「2つの開発-SCoREと小中学生向けDDL英語学習支援サイト」赤瀬川史朗先生(Lago言語研究所)、(3)「教育用例文コーパスの公開とDDL実践」中條清美先生(日本大学)・濱田彰先生(日本大学)、(4)「SCoREを利用したハンドアウトの作成とそれを利用したDDL実践」若松弘子先生(筑波大学)、(5)「DDLで育む小・中学生の英語力」西垣知佳子先生(千葉大学)・石井雄隆先生(早稲田大学)、(6)「英作文のエラー・コレクションに対するコーパス使用の効果」佐竹由帆先生(駿河台大学)の6つの発表が行われました。

◇ コーパスとCEFR研究会シンポジウム

テーマ: 「CEFRとコーパス分析の接点」

宇佐美裕子先生(東海大学)の司会のもと、(1)「SIG設立の趣旨とCEFRとコーパス分析の最新動向」投野由紀夫先生(東京外国語大学)、(2)「CEFR(-J)準拠コーパス構築とCEFR-J Grammar Profileにおける文法項目使用例抽出手法」石井康毅先生(成城大学)の発表が行われました。また、発表後にはオープン・ディスカッションが行われました。

◇ 語彙研究会シンポジウム

杉森直樹先生(立命館大学)の司会のもと、第1部「コーパスと語彙」、第2部「コーパスと連語」の発表が行われました。第1部では、(1)「英単語のCEFRレベルとコーパ

ス頻度」内田諭先生（九州大学）、(2)「縦断的コーパスを用いて英語スピーキング力の発達を探る：語彙の観点から」阿部真理子先生（中央大学）の発表が、第2部では、(3)「TIO (Target-Input-Output) 連動分析で見る L2 英語 n-gram」石川慎一郎先生（神戸大学）、(4)「語の意味的粒度とコロケーションに関する試論」長谷部陽一郎先生（同志社大学）の発表が行われました。

英語コーパス学会第44回大会のお知らせ

日時：2018年10月6日（土）～10月7日（日）

場所：東京理科大学（神楽坂キャンパス）

土曜日はワークショップ・総会・研究発表・シンポジウム・懇親会、日曜日は研究発表・講演の予定です。詳細については8月前半までに当学会ウェブサイトに掲載予定の大会資料をご覧ください。大会資料の冊子は、登録されている住所宛にお送りします。

■ 新入会員紹介

迫田久美子（広島大学）
Daniel Teuber（大阪産業大学）
浅井静代（立命館大学）
神田みなみ（千葉県立保健医療大学）
今滝暢子（日本大学）
川口恵子（芝浦工業大学）
甘利実乃（東京外国語大学, S）
坂場寛子（東北大学, S）
松本優美（聖心女子大学）
守家 輝（京都大学, S）
福本広光（大阪大学, S）
三隅 光（東京外国語大学, S）
福永真理子（京都大学, S）
工藤洋路（玉川大学）
松田節郎（松江工業高等専門学校）

Jonathan Ferries

(University of Sunderland, S)

(Sは学生会員)

(2017年12月2日から2018年7月6日の入会者)

■ 理事会の決定事項について

2018年4月21日（土）に東京外国語大学（府中キャンパス）において理事会が開催されました。承認された人事についてご報告いたします。（理事会後にメール審議で承認された内容も含まれます。）役員・委員の新任・再任の先生方には学会運営で重要な役割を果たしていただくことになりま

すが、よろしく願い申し上げます。

(1) 理事

・理事（退任）

地村彰之先生（岡山理科大学）

園田勝英先生（北海道大学）

・理事（新任）

アントニ・ローレンス先生（早稲田大学）

金澤俊吾先生（高知県立大学）

小島ますみ先生（岐阜市立女子短期大学）

水野和穂先生（広島修道大学）

2018年3月末をもって地村先生と園田先生が理事をご退任されました。各種委員会等で当学会の発展に多大なるご尽力をいただきました。ありがとうございました。今後ともご指導をいただけますようお願いいたします。

アントニ先生、金澤先生、小島先生、水野先生、これからの学会運営にお力を貸していただけますようお願いいたします。

(2) 事務局

・会計（退任）

小島ますみ先生（岐阜市立女子短期大学）

事務局の大黒柱であった小島先生、長きにわたり事務局を支えてくださり、ありがとうございました。

(3) 編集委員会

・委員長（退任）

中尾佳行先生（福山大学）

・委員長（新任）

田畑智司先生（大阪大学）

・委員（退任）

家入葉子先生（京都大学）

瀬良晴子先生（兵庫県立大学）

中尾佳行先生（福山大学）

中尾先生、編集委員会委員長の重責を担ってくださり、ありがとうございました。また、2011年度～2016年度に編集委員長を務めて下さった瀬良先生、編集委員を長らく務めて下さった家入先生、長きにわたり学会誌の編集にご尽力いただき、誠にありがとうございました。

(4) 学会賞選考委員会

・委員（退任）

地村彰之先生（岡山理科大学）

田畑智司先生（大阪大学）

・委員（新任）

五百藏高浩先生（高知県立大学）

石川保茂先生（京都外国語大学）

地村先生，田畑先生，学会の活性化のためにご尽力をいただきましてありがとうございました。

(5) 大会企画委員会

・委員長（再任）

金澤俊吾先生（高知県立大学）

・委員（退任）

滝沢直宏先生（立命館大学）

西村秀夫先生（三重大学）

・委員（新任）

西部真由美先生（愛知大学）

渡辺拓人先生（熊本学園大学）

滝沢先生，西村先生，長きにわたり，大会企画に携わってくださり，ありがとうございました。

<会誌『英語コーパス研究』第26号論文投稿募集について>

『英語コーパス研究』編集委員会委員長
田畑智司（大阪大学）

『英語コーパス研究』第26号の原稿を次の要領で募集いたします。会員各位の積極的な投稿をお待ちしております。

【原稿の種類】

1. 英語コーパス利用・コンピュータ利用を中心に据えた「研究論文」，「研究ノート」，「総説論文」，「書評論文」，「実践報告」
2. 「書評」，「コーパス紹介」，「ソフトウェア紹介」，「海外レポート」，「論文紹介」などの各種情報あるいは紹介原稿

【原稿提出期限】2018年11月30日（金）

電子メール添付にて提出してください。提出方法等についての詳細は学会 Web ページの投稿規定 http://jaecs.com/jnl/jnl_kitei.pdf を参照してください。

【問い合わせ先・原稿提出先】

『英語コーパス研究』編集委員会

E-mail : jaecs.ed@gmail.com

【採用通知】2019年1月

【発行日】2019年3月31日

（発送は2019年5月下旬の予定）

■ 今後の大会日程と開催校

第45回大会は2019年10月5日（土）・6日（日）に高知県立大学にて開催する方向で現在調整を行っています。

■ 事務局から

事務局からは情報発信のツールとして，ホームページ，ニューズレター，JAECSメーリングリストでイベントの案内などを随時行っております。

◇会費納入のお願い

2018年度会費（一般6,000円，学生3,000円）未納の方は，6月にお送りした払込取扱票を使ってお納めいただきますよう，ご協力をお願いいたします〔振替口座：00930-3-195373〕。払込取扱票を紛失された方は，郵便局に備え付けのものに加入者名「英語コーパス学会」とご記入の上お納めください。

過年度会費未納の方は，2018年度分と併せてお納めください。過年度会費未納の場合，機関誌などの送付を一時中止させていただいております。

住所，所属などに変更や異動のある方は，学会ウェブサイトの「会員情報の変更」からのお手続きをお願い申し上げます。

※会員の皆様には，日頃より会費の当該年度内納入にご協力をいただきまして，お礼申し上げます。会費を滞納されますと，退会時に滞納分をまとめてお支払いいただくといった事態にもなりかねません。会員の皆様におかれましては，円滑な学会運営のためにご協力いただけますようお願い申し上げます。なお，退会を希望される場合は，当該年度内に学会ウェブサイトの「退会案内」からのお手続きをお願い申し上げます。

◇Forumの原稿募集中！

英語コーパス学会 Newsletter では会員の皆様からの Forum への投稿を募集しています。国際学会報告，研究会の紹介，新刊紹介など，会員の皆様の情報交換の場として Forum が活用されることを願っております。以下，詳細を記します。

- **Forum のテーマ**：国際学会報告，研究会の紹介，新刊紹介など英語コーパス学会にとって有益と思われる情報
- **締め切り**：5月末あるいは10月末
- **分量**：800-1,600字程度（画像も可です）
- **送付先**：jaecs.hq@gmail.com

FORUM

■新刊紹介 1 : Biber, D. & Reppen, R. (eds.) (2015) *The Cambridge Handbook of English Corpus Linguistics*. Cambridge University Press.

報告 : 高橋有加 (広島大学)

本書はコーパス言語学及びコーパス言語学の研究者や大学院生レベルを対象とした概論書で、多岐にわたる言語学の領域でコーパス言語学の手法を駆使した、出版時点で最新の研究が集積されている。岩崎研究会コーパス部会では2016年から半年に1度、本書を用いた輪読が行われている。本書の目的は、Biber が特に得意とするレジスター変異と言語使用に関する研究設問を広範囲な応用分野においてコーパス言語学の視点からアプローチすることである。具体的には *lexical variation*, *grammatical variation*, *historical change*, *linguistic description of dialects and registers*, *language teaching and translation* などの領域が含まれる。序章の後に続く本論は第1部 *Methodological considerations*, 第2部 *Corpus analysis of linguistic characteristics*, 第3部 *Corpus analysis of varieties*, 第4部 *Other applications of corpus analysis* の4つのパートで構成されており、全28本の論考が収められている。各論考の構成としては厳選された先行研究、事例研究、研究についての批評が含まれている。

第1部 *Methodological considerations* は現在使用可能なコーパス及びインターフェイスの紹介、そしてリサーチデザインや統計手法について概観された3本の論考からなっている。

- 1. *Corpora: an introduction* (Mark Davis), 2. *Computational tools and methods for corpus compilation and analysis* (Paul Rayson), 3. *Quantitative designs and statistical techniques* (Stefan Th. Gries)

上記の Paul Rayson による論考では、アンテーションや書き起こしを含むデータ処理の際に使われる多様な分析ツールの種類と手法について紹介されている。調査として、2012年に出版された1) *International Journal of Corpus Linguistics*, 2) *Corpora*, 3) *Corpus Linguistics and Linguistic Theory*, 4) *ICAME Journal* の中から、32本の研究で用いられているツールを3つのカテゴリ (*Retrieval*, *Compilation*, *Annotation*) に分類し

たところ、コーパスからの情報抽出: 84%, コーパス構築: 38%, 言語情報付与: 25%であったとし、大部分が既存のツールを使用した情報抽出で、コーパス構築に分類されるような独自開発したツールや R を用いた研究は少なかったと報告している。コーパス構築・研究と既存の様々なツールの関係を把握するのに非常に役に立つ章である。

第2部 *Corpus analysis of linguistic characteristics* はイントネーション, コロケーション, 文法, 談話分析, 語用論といった様々な言語学の領域・研究分野に焦点を当てた11本の論考からなっており、いずれもコーパスを用いた分析結果と知見が示されている。

- 4. *Discourse intonation: a corpus-driven study of prominence on pronouns* (Winnie Cheng), 5. *Keywords* (Jonathan Culpeper and Jane Demmen), 6. *Collocation* (Richard Xiao), 7. *Phraseology* (Bethany Gray and Douglas Biber), 8. *Descriptive grammar* (Geoffrey Leech), 9. *Grammatical variation* (Daniela Kolbe-Hanna and Benedikt Szmrecsanyi), 10. *Grammatical change* (Martin Hilpert and Christian Mair), 11. *Lexical grammar* (Susan Hunston), 12. *Using corpora in discourse analysis* (Alan Partington and Anna Marchi), 13. *Pragmatics* (Brian Clancy and Anne O'Keefee), 14. *Historical pragmatics* (Irma Taavitsainen)

7. *Phraseology* (Douglas Biber) では、Longman *Spoken and Written Corpus* を使い、高頻度の4語から成る語彙的フレーム (e.g. *in the *of*) を取り出す際に、1) 高頻度連鎖に焦点を絞ってから語彙的フレームを取り出した *hybrid-corpus based* な方法 (Biber, 2009) と、2) Perl を用いた *corpus-driven* な方法で抽出した場合の語彙的フレームの数を比較した。その結果、前者で抽出できたのは後者の約2分の1の数であったことを明らかにした。前者は、高頻度な語彙的フレームはある程度頻度の高い連鎖を含んでいるであろうことを前提にした方法であるため、フレームのスロット (e.g. *in the *of*) を埋める単語の *type-token ratio* (*variability*) が低い場合は取り出せたが、高い場合 (スロットに様々な単語が入る高頻度フレーム) は抽出できていなかったことがわかった。どの方法なら何が取り出せるのかを理解した上で手法を使い分ける必要があるとしている。

12. *Using corpora in discourse analysis* (Partington and Marchi) では、ブッシュ大統領とオバマ大統領の *Whitehouse briefings* のコーパス内の高頻度語彙を時間を追って比較し、同一語彙

に関してはどのような使用の違いがあるかを詳細に分析している。例えば、それぞれのコーパス内で *job* という単語がどのような単語と共起し、どのような意味で使われ、どのような政治的・社会的な出来事と関連しているかという点について分析している。コーパスのような大量のデータを分析して得られる結果は予測するのが難しいことも多く、いくつか立てた仮説にあてはまる場合と、仮説とは全く違う予想外の結果が出る場合があるため、両方に備える必要があることにも言及している。

第3部 *Corpus analysis of varieties* は英語の *variety* について書かれた以下の9本の論考からなる。

➤ 15. *Spoken discourse* (Shelly Staples), 16. *Corpora and written academic English* (Ken Hyland), 17. *Register variation* (Susan Conrad), 18. *Diachronic registers* (Merja Kyotö and Erik Smitterberg), 19. *Literary style and literary texts* (Michaela Mahlberg), 20. *Dialect variation* (Jack Grieve), 21. *World Englishes* (Marianne Hundt), 22. *New answers to familiar questions: English as a lingua franca* (Anna Mauranen, Ray Carey, and Elina Ranta), 23. *Learner language* (Gaëtanelle Gilquin and Sylviane Granger)

15. *Spoken discourse* (Shelly Staples) では、看護師と患者の会話における4つの *phase* (opening, exam, council, closing) とその中の *stance* を表す法助動詞、副詞、*that*-clause, *to* clause の頻度とその具体的な会話例が示されている。

23. *Learner language* (Gaëtanelle Gilquin and Sylviane Granger) ではLOCNECとLINDSEIを使用し、11の異なる言語を母語とする英語学習者グループと英語母語話者を対象に、2語の連鎖からなるディスコースマーカの使用傾向を比較している。その結果、フランス語話者は *in fact* を、英語母語話者とスウェーデン語話者は *sort of* を多用する傾向があるとし、L1の影響だけでなく学習者の習熟度の高さなども考慮に入れた考察をしている。また、最も多く頻度が観察された *you know* がL1グループによって文の異なる位置に使用されていることについても考察している。

第4部 *Other applications of corpus analysis* では、教材としてのコーパス、語彙研究、辞書学、通訳といったテーマの論考が5本集められている。

➤ 24. *Vocabulary* (Ron Martinez and Norbert Schmitt), 25. *Lexicography and phraseology*

(Magali Paquot), 26. *Classroom applications of corpus analytics* (Thomas Cobb and Alex Boulton), 27. *Corpus versus non-corpus-informed pedagogical materials: grammar as the focus* (Fanny Meunier and Randi Reppen), 28. *Translation* (Silvia Bernardini)

26. *Classroom applications of corpus analytics* (Thomas Cobb and Alex Boulton) では、1989年から2012年までの研究を対象に、コーパスを用いた教室での指導効果についての21本の先行研究(語彙・文法・ライティング・リーディング指導など)のメタ分析を行い、結果とその効果量を比較した上で、様々な指導においてコーパスを活用した指導が有効であることを示している。

各専門分野の著名な著者による最新の研究を紹介するとともに、これまでの先行研究のなかでも厳選されたものが多く紹介されている。ハンドブック全体からは言語学の領域を中心とした様々な分野の研究例におけるコーパス言語学の役割とその重要性を知ることができ、コーパス言語学が方法論として切り離せない存在になっていることが示されている。どの研究にも共通する点として、コーパスやツールによって格段に研究が進んだこと、そしてツールを使った量的分析と人の手による最終確認や個人データを対象とした質的分析の両方を組み合わせる必要性が述べられている。コーパス言語学、言語学、方法論の3つに関わる多数の研究例に1度に触れることができ、この領域における研究を広く深く見渡すことができる1冊である。

■新刊紹介2 : Szudarski, Pawel (2018) *Corpus Linguistics for Vocabulary*. Routledge.

報告 : 英語コーパス学会語彙研究会

■はじめに (石川慎一郎 : 語彙研究会代表)

先般、英語コーパス学会ニューズレター編集委員会より、P. Szudarski (2018) *Corpus Linguistics for Vocabulary* の紹介記事の執筆の依頼を受けた。

ちょうど、語彙研究会が本格的に活動を開始したところだったので、分担執筆を研究会メンバーに呼び掛けたところ、鎌倉義士氏(4章)、小屋多恵子氏(6章)、佐藤恭子氏(9章)、中西淳氏(2章、7章)、能登原祥之氏(5章)より申し出があった(50音順)。そこで、以上の5氏に、それぞれに希望される章の紹介を依頼した。また、残りの1, 3, 7, 8, 10章については筆者が担当した。

およそ1章につき1~2ページ程度で原著の概要をまとめるという目安を決めたほかは、報告の様式は担当者の裁量にゆだねた。このため、担当者により、文体や報告内容の精粗の度合いに違いがあるが、あえて編集は行わず、各人の報告をそのまままとめてレポートとして提出する。なお、草稿については、杉森直樹氏はじめ、研究会メンバーからご意見をいただいて修正を行った。

本書は全10章構成(224ページ)で、はじめに、コーパス言語学の概論(1章)、コーパス分析法の概説(2章)、語彙論の概説(3章)を行った後、語彙頻度研究(4章)、連語研究(5章)、語彙指導(6章)、学習者語彙使用(7章)、専門語彙(8章)、語用論(9章)のトピックについて議論が行われ、最後にまとめと研究事例の紹介(10章)がなされる。

本書には関連する先行研究が手際よく紹介されており、事例も豊富で読みやすい。コーパス語彙研究を生業としている研究者にとってはいささか物足りないと感じられるかもしれないが、大学院生や学部生が初めてコーパスや語彙研究を学ぼうとする際、本書は推薦できる入門書の1つと言える。広く一読を勧めたい。

本書の著者である Szudarski 氏は、現在、ノッティンガム大学英語学科の Teaching Associate の職にあり、オンライン修士コースの指導のまとめ役をしている。最近の論文・著述として、“Learning and teaching L2 collocations: insights from research” (2017)、“The role of input flood and input enhancement in EFL learners’ acquisition of collocations” (2016)、“Formal instruction in collocations: Mixed methods approach” (2015) 等がある(共著含)。

■Ch. 1 What is corpus linguistics? (石川)

コーパス言語学とは「コーパスの構築と分析」に関わる研究と定義できる。相対的に新しい分野だが、言語研究の革新に大きく寄与してきた(1.1節)。

特定の言語集団を代表するコーパスを作るには設計が重要である。その際、検討すべき点が複数ある。サイズについては、研究課題によって適切なサイズが決まる。代表性・均衡性に関してはコーパスに包含されるセクションごとの分量がほぼ同等語数になるべきで、COCA が良い例である。アノテーションについてもその適切な方法は研究課題による(1.2節)。

コーパス研究は、実証研究である点、慣用連語(phraseology)を重視し、語彙文法(lexicogrammar)研究につながる点、コミュニケーションの文脈を重視する点にその価値があるが(1.3

節)、含まれているものしか調査できない点、母集団から見ると小さすぎる点、文脈が切断されている点、データ解釈までは示してくれない点に限界がある(1.4節)。

コーパスには様々な種類がある。特殊コーパスとしては、MICASE(大学講義)、BAWE(学生作文)、BASE(大学講義)、HK Engineering Corpus(工学)、ICLE(学習者)、VOICE(ELF発話)、ELFA(大学ELF発話)、WrELFA(ELF論文等)、ACE(アジアELF)、EuroCoAT(エラスムス学生対話等)等がある。また、話し言葉コーパスとしては、CANCODE、Santa Barbara Corpus、HK Corpus of Spoken English、BNC2014等がある。時系列コーパスとしては、COHA(1800年代~)やARCHER(1600年代~)がある。パラレルコーパス・比較コーパスとしては、Oslo Multilingual Corpus(独・仏・フィンランド語)、Digital Corpus of the European Parliament(EUの各国語)、Lancaster Corpus of Mandarin Chinese(Brownと対応する中国語版)等がある。ウェブ資料を検索するコーパスとしては、WebCorp(ジャンル指定でウェブ検索)、Google Books Advanced(米国版Google Books検索)、NOW(オンライン新聞等370億語検索)等がある。

◎報告者メモ：バランスの取れた紹介文。ただし、「コーパスに包含されるセクションごとの分量がほぼ同等語数になるべき」という記述(p.6)については、セクションごとの実際の資料の比率を再現すべきという考え方もあるのでやや誤解を招くかもしれない。

■Ch. 2 Corpus Analysis: Tools and Statistics (中西)

本章では、コーパスを用いて語彙研究を進めるにあたって必要となるツールと統計手法が紹介されている。

最近では、AntConcのように幅広いコーパス処理に対応したソフトウェアも存在するが、本章では、主な分析技術ごとに異なるソフトウェアを紹介している。例えば、頻度検索にはBYU-BNCのウェブ検索インタフェースが、ワードリストには、同じく、BYUの検索インタフェースが、共起語検索にはAntConcが、特徴語検索には、Lextutorが紹介されている。

また、統計手法については、特徴度指標であるLog-likelihoodや共起強度指標であるTスコアとMIスコア、多様性指標であるTTRといった、コーパス研究を行うにあたって不可欠な統計量が紹介されている。

AntConcやBYU-BNCなどはコーパス研究者の

間でも広く知られているが、本章で紹介されている Lextutor や Log-likelihood and effect size calculator についてはまだ十分に知られておらず、価値があると考えられる。特に、Log-likelihood and effect size calculator は同時に effect size (効果量) が出力される。近年、効果量の表示が求められるようになってきていることをふまえると、このツールは幅広い研究者に有益と言えるだろう。

■Ch. 3 What is vocabulary? (石川)

語彙は言語の重要な構成要素である。語彙サイズは学習者の言語能力に関係しており、語彙サイズの IELTS スコアに対する説明力 (r^2) は 40-60% である。L2 習得には上位 2,000 語の獲得が重要である。内容理解には、書き言葉だと 9,000 語 (単語家族単位)、話し言葉だと 3,000 語程度 (7,000 語という見解もあり) の知識が必要である。NS の語彙サイズは、固有名詞を除き、1.6 万語～3 万語 (単語家族単位)、学習者の語彙サイズ (最大値) はその半分と考えられている (3.1 節)。

語彙研究では、表記形 (word form)、語彙素 (lexeme)、レマ (lemma)、単語家族 (word family)、語彙項目 (lexical item) 等の用語および概念の理解が必要である (3.2 節)。

語彙知識とは複合的な構成概念である。一般には、形態 (発音・綴り・接辞等)・意味 (語義・指示概念・連想)・用法 (文法・連語・使用制限) の 3 種を区別し、さらに、それぞれについて受容知識と発信知識を区別する。また、語彙知識の広さ (語彙サイズ) と深さ (連語・連想・制約) を区別したり、語彙知識を心的ネットワーク (メンタルレキシコン) ととらえて心理言語学的に分析したりすることもある。学習者は L2 処理時にも L1 語彙知識を媒介的に使用している場合が多いが、習熟度の上達につれて L2 の意味に直接アクセスがなされる (3.3 節)。

コーパスは、多義語の分析 (例: 名詞・動詞を区別して bear を検索)、類義語の分析 (例: problem と issue の共起語を比較)、メタファー/成句性分析 (例: 頻度や形態固定性を確認)、言語使用域分析 (例: 学術/話し言葉コーパスの高頻度語を比較) などに有益である。Mark Davies 氏の開発した Word and Phrase. Info というサイトでは、COCA に基づき、任意の語の類義語を探したり、共起語を確認したりすることができる (3.4 節)。

◎報告者メモ: 英語教育の視点からの語彙研究について基本事項がうまくまとまっている。Word and Phrase. Info は有用だがややインタフェースがわかりにくい。

■Ch. 4 Frequency and vocabulary (鎌倉)

本章では頻度と語彙についての概要を説明している。コーパス言語学にとって頻度は最も重要な尺度である。この章は頻度の重要性から始まり、頻度分析、語彙の一覧、教育への応用と続き、最後に Nation and Anthony (2003) の語彙と読解に関する研究を紹介している。

4.1 「コーパス言語学における語彙の重要性」は 5 つの項から構成される。Tognini-Bonelli (2010) や O'Keefe et al. (2007) を引用し、頻度によって経験的データに基づく分析が可能と述べ、続く項で話し言葉と書き言葉そして機能語と内容語の違いを BYU-BNC や COCA からの頻度を用いて説明している。話し言葉のコーパス分析例の代表として Carter and McCarthy (2017) を代表として挙げるが、Longman Grammar of Spoken and Written English (Biber et al. 1999) も含むべきではないだろうか。そして Zip の法則、談話内に占める既習語の割合を示す lexical coverage、基本と上級の語彙に関する説明が続く。この節はコーパスを用いた語彙研究の入門書となる基本的な情報を概括している。

4.2 「頻度による語彙分析」では VocabProfile (Cobb 2015)、RANGE (Cobb & Horst 2015)、AntWordProfiler (Anthony 2014) などのプログラムによるテキスト内での高・中・低頻度語彙分布表の作成方法が紹介されている。

4.3 「頻度に基づく有用な語彙リスト」では Basic English (Ogden 1930, Richards 1943) と General Service List of English Words (West 1953) がどちらも主観的な判断による語彙選択と指摘し、LOB, BNC, EnTenTen12 で頻度分析した New General Service List (Brezina & Galasova 2015) と CCL (Gardner 2013) を語彙学習の頻度分析の最新例として挙げている。

4.4 「頻度と言語教授法」では、言語教育において教員や教材編集者が優先する語彙と指導すべき高頻度語句の選出に頻度が指針となることを説明。しかし頻度は語彙指導の基準の一つであり、言語教育の他の要素も留意すべきと記している (Kennedy 1998, Granger 2011, Leech 2001)。

4.5 「中頻度語句による読本 (Nation & Anthony 2013)」は頻度に基づく分析の応用例を紹介している。BNC と COCA から 25 群の中頻度語を提示し、学習到達度別の読本作成に頻度情報が有用であることを示している。

この章は Paul Nation の研究や論文を中心に引用し、L2 の語彙学習の分析と応用について記述されている。

■Ch. 5 Corpora, phraseology, and formulaic language (能登原)

本章前半では、コーパス言語学、句表現研究(phraseology)、そして、定型表現(formulaic sequences)研究の各分野の背景と基礎概念が簡潔に紹介される。

次に、(1)コロケーション(collocation)、(2)品詞連鎖(colligation)、(3)意味的嗜好(semantic preference)、(4)意味的色合い(semantic prosody)の4つの言語特徴がBYU-BNCとBNCwebを通した検索式の入力例や検索結果とともに丁寧に紹介される。

これら4つの言語特徴を重視する新ファース学派(neo-Firthian school)は、近年(3)や(4)など意味に焦点を当てた研究を進め、認知言語学の構文文法(construction grammar)研究を背景に、句表現研究や定型表現研究の幅をさらに広げている(e.g., McEnery & Hardie, 2012)。本章は、そのような最近の研究動向も漏れなく紹介する。

研究手法や指標については、残念ながら、代表的なt-scoreとMI-scoreの2つの指標しか解説されていない。最近公開されたSpoken BNC 2014などでは、さらに5つMI3、Z-score、Log-likelihood、Dice coefficient、Log ratioを加え、計7種の指標を比較できるようになっている。実際に研究を行う際には、他の指標の特徴も吟味し(e.g., 石川, 2012)、研究目的に合わせ適切なものを選ぶ必要がある。

本章後半では、句表現の振る舞いを使用域(register)別に比較する研究や、コーパスを参照した句表現リスト研究(phrasal expressions list)が紹介される。ここでは、アカデミック定型表現(academic formulas)の研究動向や、句表現リスト研究(e.g., Martinez & Schmitt, 2012)の概要を確認できるだけでなく、英語教育へつなげる教育的配慮も学ぶことができる。

最後に、もし余裕があれば、本章をきっかけにLindquist (2009)の4章も合わせて参考にしたい。Firth以前に日本でなされたH. E. Palmer (1933)のコロケーション研究の意義や位置付けを確認でき、この分野の興味をより一層広げ深めることができるだろう。

■Ch. 6 Corpora and teaching vocabulary (小屋)

本章では、コーパスが語彙指導に与える効果と語彙指導におけるコーパスの活用について論じている。コーパスは、これまでの母語話者の直感や経験といった主観的な判断ではなく、実際に使用されている客観的な言語運用傾向を提示してくれることから、語彙指導に必要なテキスト、辞書、ワードリストやオンライン教材がコーパスに基づ

いて編纂されるようになっていく点を個々に説明している。コーパス情報やコーパス研究の成果をいかに反映させることができるかがカギとなると指摘している。次に、コーパス情報を直接学習者に提示し、そこから帰納的に語法や文法情報を見つけ出すData-driven learning (DDL) や指導の目的に応じて作成された自作コーパス(例: 教員養成のためのコーパスや授業で使用するためのコーパス)の活用が紹介されている。目的、学習者のレベルやニーズに合わせて使用を思案・準備することが重要であると主張している。

著者は、語彙指導におけるコーパスの有用性をこれまでの研究を紹介しながら示している一方で、課題も言及している。例えば、コーパスに基づく教材の編纂においては、コーパスを最大限に活かそうとする意識や効率より効果に重点をおいた編纂が少ないことをあげている。また、学習者の指導においては、インストラクターのコーパスリタラシーの重要性を訴えている。インストラクターが、コーパスを正しく理解し、コーパスを分析するソフトを使いこなし、自信を持って分析結果の解釈ができるようになってこそ効果的な活用が期待できるということである。

コーパスの有用性は広く支持されるようになってきたが、コーパスを効果的に語彙学習に取り入れるのには、教材作成においても、指導においても、試行を重ねていくことが必要であろう。とくに指導においては、活用できる人材の育成と並行して、自信がつくまでコーパスは活用しないのではなく、まずはコーパスベースの教材を多く使用する、紙ベースでよいからコーパスデータの一部を授業で使用するなどといった試みから初めてみるべきではないだろうか。

■Ch. 7 Corpora and Learner Vocabulary (中西)

従来、学習者の語彙研究は学習者コーパスを用いた対照中間言分析(Contrastive Interlanguage Analysis: CIA)を基盤とする、第1言語(L1)話者と第2言語(L2)学習者間の比較や異なるL1を持つL2学習者間の比較などが広く行われてきた。しかし、近年では、リングフランカ研究や習熟度に伴う語彙変化の研究など、学習者コーパスを用いた研究の多様性は広がってきており、本章では、それらの研究に対応した学習者コーパスが広く紹介されている。

また、これらの学習者コーパスを教育的に応用した研究についても触れられている。たとえば、Cambridge Learner Corpusを用いたEnglish Vocabulary Profile (EVP)では、語彙やセットフレーズにCommon European Framework of Reference for Languages (CEFR)のレベルが割り当てられて

おり、この指標を基に作られた Text Inspector というツールを用いることで、テキストの語彙難易度を測定することができる。これらはいずれも登録すれば無料で使用することが可能であり、英語学習者の語彙学習や英語教師の語彙指導の際に重宝されるものであると考えられる。

■Ch. 8 Specialized corpora and vocabulary (石川)

大規模コーパスは文脈の切断されたデータであると批判されることがあるが、この点をふまえると、特定の文脈と密接に連結した小規模な特殊コーパスを使う研究にも利点が多い。もっとも、特殊コーパスを構築する際にも代表性への配慮が不可欠である。有名な Coxhead の学術語彙研究についても、元となったコーパスが偏っている（法学と商学だけで 50%になる）という批判もある（Durrant, 2014）（8.1 節）。

小規模な特殊コーパスは、言語使用域やジャンルの分析に有益である。Biber et al. の LGSWE は会話・小説・新聞・学術文といった言語使用域ごとに英語の違いを浮き彫りにした。また、ジャンルについては、ビジネス会合・コールセンターでのやりとり・書簡・顧客と会社のやりとり等について言語分析がなされている。言語使用域とジャンルは似た概念だが多くのコーパス研究は伝統的に前者を重視している（8.2 節）。

特定の談話集団による LSP（専門分野のための言語）研究にもコーパスは有益である。英語では EAP（学術英語）や EOP（医学・法学・ビジネスなどの職業別英語）の研究が盛んであるが、特定の領域で典型的に使用される専門用語（technical vocabulary）をコーパスから特定することができる。これらの語を習得することは特定の談話集団に参画する一助となる。ただし ESP 語彙のサイズははっきりしておらず、およそ 1,000 から 5,000 語程度であろうという見解が示されている（Nation, 2008）。Chung（2003）は、解剖学コーパスと応用言語学コーパスを構築し、一般コーパスと語彙比較を行った結果、解剖学テキストの 31% が専門用語であったのに対し、応用言語学テキストの専門用語は 20% だけで、しかも他分野でも多用されていることを明らかにし、分野により専門用語への依存度は異なると述べている。ESP コーパスは、教育現場において、直接利用されるほか（例：データ駆動型の L2 教育）、間接利用されることもある（例：Tim Johns 氏の Kibbitzers [英語語法観察ガイド] では、「as 節と that 節」「By と Using」「data の単複性」など、EAP コーパス分析で分かった、間違いやすい類義語等の違いを解説している）（8.3 節）。

EAP 研究については Coxhead の学術語彙リスト（AWL）が有名である。これは人文学・商業・法律・科学の各分野の資料を集めた 350 万語の学術英語コーパスから選んだ重要語 570 語（単語家族）のリストで、これだけで学術文の 10% をカバーする。AWL は Lextutor などのサイトにも組み込まれており容易に使用できるが、一般語が入っていないことや、分野差への意識が低いことなどへの批判もある。最近では、Gardner & Davies（2014）が、COCA の 9 種のジャンルデータを分析し、頻度（学術文で他より 50% 以上多く出る）・レンジ（9 ジャンル中 7 ジャンル以上で出る）・分布（コーパス全体で均等に出る）・一般性（いずれかのジャンルで平均の 3 倍以上出てはいけない）といった基準を満たす 3,014 語（レマ）を選び、新たな Academic Vocabulary List（AVL）として発表した。また、Simpson-Vlach & Ellis（2010）は MICASE, BNC, 学術論文コーパスを分析し、学術連語リスト（Academic Formulas List）を公開した。これは書き言葉・話し言葉共通の重要連語 207 種、書き言葉で重要な 200 種、話し言葉で重要な 200 種を含む。この研究では、頻度に加え、専門家の考える重要度を計量化した formula teaching worth（FTW）指標を語彙選定資料としている（FTW に基づく上位 3 連語は in terms of, at the same time, from the point of view）。Ackermann & Chen（2013）も大規模な EAP コーパス頻度と専門家評価を組み合わせ、2,400 種の EAP 連語リストを発表した。Liu（2012）は BNC と COCA の学術英文データを用い、228 種の学術チャンクリストを発表した。これは、言語機能によって、参照・観念構成表現、スタンス標識表現、談話構成表現の三大区分、および、細かな下位区分に分類して提示されている（8.4 節）。

他の特殊コーパスを使った語彙研究もある。Biel（2012）は英語とポーランド語の法律文を集めたパラレルコーパスを比較し、高頻度語は 2 言語ともに類似した概念メタファーを持つが、コロケーションは異なることを明らかにした。また、Mahlberg et al.（2013）は Charles Dickens の作品コーパスを分析し、後置引用節（“...,” X said）がディケンズを特徴づけていることを明らかにした。Barbieri（2008）は米語会話コーパスを使って若年者と中高年の発話を比較し、若者発話が俗語・フィラー・感情関連語・強調語・談話標識・1/2 人称の多用によって特徴付けられることを明らかにした（8.5 節）。

◎報告者メモ：内容豊富な章。AWL だけでなく最新の AVL についての紹介もある点が有益。ただし既存学術語彙リストについてもう少し批判的観点

からの比較があってもよかった。なお、原著 p.142 の kibitzers は Kibbitzers の誤植か？

■Ch. 9. Discourse, pragmatics and vocabulary (佐藤)

本章では、談話と語用論の観点からの語彙の機能について、コーパスがどのように寄与できるかを論じている。談話分析の質的方法とコーパス言語学の量的手法は、その研究方法に違いがあるが、自然な言語データを扱う点では一致しており、それぞれが統合しあい、補いながら「コーパスに基づいた談話研究」がなされるとしている。

9.3 では談話における語彙の重要な役割として、まず語彙的結束性を挙げ、語彙の繰り返しや同じ意味を表す関連語（同義語、近似同義語、上位語）の例や、CANCODE のデータを用いた Buttery and McCarthy (2012)を紹介している。次にテキスト間相互関連性における機能を挙げ、Flowerdew and Forest (2009, 2015)の Ph.D.論文の文献レビューのコーパスを用いた研究を取り上げている。さらに談話標識の研究にも触れ、学習者コーパスを用いて、母語話者との使用の違いを比べた研究を紹介している。

9.4 ではコーパスと語用論について見ている。語用論的分析にコーパスを用いる際には、その観点から情報が付与されたコーパスを用いる必要があり、その点ではまだ十分ではないと課題を指摘している。9.4.1 では O’Keeffe et al. (2007)の CANCODE の分析結果を紹介し、多くの語用論的機能を果たす語彙が会話を構築していく上で、重要な機能を果たしているとしている。さらに曖昧語の研究について、Hong Kong Corpus of Spoken English の2つの subsection と Limerick Corpus of Irish English を比較した Cheng & O’Keeffe (2015)を挙げている。9.4.2 では Adolph (2008)の発話行為の「提案」の分析を紹介しているが、語用論的観点から情報付与されたコーパスは殆どなく、the Corpus of Verbal Response Modes, the Speech Act Annotated Corpus for Dialogue Systems と MICASE の一つのサブセクションがある程度であるとし、まだ十分ではないとしている。Adolph (2008)では、発話行為の分析の手順を「謝罪」を例に分かりやすく示しており、コーパス研究を今から始める人には有効であろう。9.4.3 では、意味的色合い(semantic prosody)の語用論的機能を扱い、Sinclair (2004)の Bank of English を用いた true feelings の意味的韻律の研究を紹介している。

最後に 9.5 でコーパスと談話分析を扱った適例として 12 種類の休日を過ごすパンフレットからなる小規模コーパスを用いた Baker (2013)を紹介

し、語彙リストや n-gram, 語彙の分散の程度、さらには年齢、性別といった社会言語的要因から、休日を扱う談話の構造を語彙的観点から明らかにしている。

■Ch. 10 Summary and research project (石川)

今後のコーパス言語学は、超大規模コーパス開発、発話コーパスの設計法の改良、新型モニタコーパス（毎日データ更新）開発、バーチャルコーパス（カスタマイズ可）開発、コーパスデータを解釈する新たな技法の開発（グラフ表示など）、クラウドソーシング型コーパス開発などの方向に進んでいくと思われる。とくに方法論の多元化 (triangulation) が必要である。語彙については、書き言葉・話し言葉別の語彙パターン解明、日常語研究、研究成果の教育応用、コーパス文体論、豊富な情報を提供する新型電子辞書開発、マルチモーダル環境での語彙使用の解明などが課題になる (10.1 節)。語彙研究は活発な研究分野であり、多くの問題にコーパスが活用可能である (10.3 節)。[※10.2 節は研究事例 (4 種) であるため、紹介を省略する。]

2018年7月25日発行

編集・発行 英語コーパス学会
会長 投野 由紀夫
事務局 〒157-8511 東京都世田谷区成城 6-1-20
成城大学社会イノベーション学部
石井康毅研究室気付
e-mail: jaecs.hq@gmail.com
twitter: @JAECs2012
URL: <http://jaecs.com/>
