

巻頭言
『英語コーパス研究』第30号
(英語コーパス学会創立30周年記念号) 刊行に寄せて

田畑 智司

本学会創立の翌年、平成6年(1994年)に創刊された『英語コーパス研究』がこの度、第30号を迎えることは、まことに喜びに堪えません。本号を含め、これまでに掲載された原著論文は98篇、研究ノート36篇、誌上シンポジウム論文67篇、特別寄稿・講演13篇、コーパス・ソフトウェア紹介12篇、書評・図書紹介10篇、海外リポート5篇、教育実践報告1篇、合わせて242篇にのぼります。また、本誌は巻末に当該年度の大会プログラムと資料を記載しており、当学会の歩みを記す貴重な年譜ともなっています。

この30年間に掲載された論文や記事を振り返ると、最初の10年は語法・文法や通時的変異、史的言語研究、多変量解析によるテキスト類型論などを扱った論文に加えて、コーパスおよびソフトウェア紹介や、書評・図書紹介、海外リポートが多く掲載されているなど、黎明期にあった当学会の姿が如実に反映されているように思えます。次の約10年は、誌上シンポジウム論文が積極的に公表されるようになる一方、掲載論文で扱われるトピックの範囲が大きく広がりました。特に、パラレルコーパスや学習者コーパス、母語話者・非母語話者による言語産出を収録した対照コーパスなどの構築や活用に基づき、学習者やアカデミックライティングの言語特徴を明らかにする英語教育学関連の論文が増加したほか、コーパスとしてのWeb、認知言語学におけるコーパス利活用、特定の言語使用域や集団の特徴抽出の方法論などが盛んに議論されるようになり、本学会の発展の様子がうかがい知れます。

さらに、最近の10年余りでは、データ駆動型学習、ESPやEAP、アカデミックライティングのムーブ分析、言語処理を応用した分析や作文を支援する新たなツールおよびシステムの開発、そして理論言語学・認知言語学とコーパス研究との融合などから見て取れるように、コーパス研究はさらに発展展開を続けてきました。なかでも、機械学習に基づく分類器や特徴(量)抽出、そして、トピックモデリングやword2vecなど、急速に進歩する自然言語処理の知見がコーパス研究に導入されるようになってきたのも新たなトレンドだと言えるで

しょう。

この間、コーパス研究の外側では、自然言語処理が長足の技術的進歩を遂げています。とりわけ、Bidirectional Encoder Representations from Transformers (BERT) や Generative Pre-trained Transformer (GPT)-3/GPT-4 に代表される、transformer を基盤とした深層学習による大規模言語モデルは、正確な数値は公開されていないものの、極めて巨大な文章データをコーパスとして学習させています。その結果、爆発的な勢いで汎化性能を高め、生成 AI としてまさに時代の寵児となる一方、研究倫理や規範、著作権、個人情報に対する脅威とも目されるほどに高度な文章生成能力を備えていることは、もはや人口に膾炙するところとなりました。

生成 AI は文章校正や、参考文献リストの整理・確認、プログラミングコード生成など、すでに学界でも利用が始まっています。他方、研究倫理の面からその利用ガイドラインが議論され始め、本誌の投稿規定にも生成 AI の利用についての条項が追加されることが決まっています。私たち英語コーパス研究にたずさわるものにとって、これからの学術研究は好むと好まざるとに関わらず、AI と共生・共存していくことになるでしょう。これまで、私たちはコンピューター、プログラミング、統計学、言語処理などをスキルとして味方にする術を模索してきました。そして、コーパスを通して初めて可視化される言語使用の実際を観察し、記録することにより、理論や仮説の発見・検証、さらには新たな問いの創成など、知の蓄積と継承に取り組んできました。そのことは 30 年にわたって本誌に掲載されてきた論考に克明に刻まれています。生成 AI をうまく利用することにより、コンピューターにさせる仕事と、人間がすべき仕事一着想、分析、議論、そして洞察一とを適切に切り分け、私たちが後者に専念し、その質をより一層高めるためにはどうすればよいのか。本誌『英語コーパス研究』が、まさにその答えに至る道標を照らす論考の宝庫^{アーカイブ}となることを大いに願う次第です。

(田畑 智司 大阪大学)