

学習者コーパスによる事例研究：英語教育への応用¹⁾

野澤和典

はじめに

応用言語学の一つとしてのコーパス言語学 (Corpus Linguistics) の研究は欧米を中心に盛んにされてきている一方、日本でも英語研究を主たる目的とする Brown Corpus, LOB Corpus, The British National Corpus²⁾, CobuildDirect Corpus³⁾ などの大規模なコーパスを利用して盛んにされてきている。さらに、最近では第二言語 (外国語) としての英語学習者のライティングやスピーチの大規模データである学習者コーパスの構築と利用が世界各地で試みられ、新しい波として脚光を浴びてきている。

本稿では、世界各地に散在する大規模・小規模の学習者コーパスが外国語 (英語) 教育・学習にどのように利用されてきたか、幾つかの事例から、それらの理論的背景と実践的手法を紹介すると共に、今後の学習者コーパスを利用した英語教育・学習にどんな新展開が待ち受けているのかを考察する。特に、1990年に出版され、1997年に改訂された Chris Tribble と Glyn Jones の言語教育における教育資源としてのコンコダンス (concordance) 利用の考え方やコーポラ (corpora) を言語教育に利用する考え方など、さらには日本で先端的な研究をしている 齊藤・中村・赤野 (1998) や 鷹家・須賀 (1998) などのコーパス言語学の基本的研究、杉浦 (1993)、朝尾 (1998)、投野 (1998)、投野 (1999)、井上 (1999) などの有益なコーパス利用の実践研究についても触れ、今後の研究発展およびその利用について論じる。

1. 学習者コーパスの定義と目的

齊藤ほか (1998) によれば、学習者コーパス (learner corpus または learner's corpus) は、特定の

言語を外国語として学んでいる学習者が書いたり、話したりしたものを、コンピュータ可読な形式で多量に記録したものと定義されている。母語話者のコーパスと異なり、学習者コーパスで記録されるものは、発達段階にある中間言語 (interlanguage) を反映した言語運用であり、そのデータ収集方法の違いから、一般コーパスのジャンル分けは利用できないなどの特徴がある。

一般コーパスの主たる目的が言語そのものの分析であるのに対し、学習者コーパスの主たる目的は、学習者が言語習得の過程で形成する中間言語を研究する、第二言語習得研究である。

2. コーパスの外国語 (英語) 教育・学習への応用

2.1. 海外での研究

当時までの研究成果をまとめた Tribble & Jones (1990) では、第4章がコンコダンサーによってプリントアウトされたものを教材として利用する方法を幾つか示している。必ずしも的確なものとは限らないが、文脈から得られるヒントによって未学習語句 (キーワード) の意味を考えさせたり、前置詞、reported speech、比較級、同音異義語・同義語、感覚動詞などの文法的特質を Q&A、グループ・ワーク、穴埋め問題、左右の内容にあった並べ替えのスタイルで学習させる具体的なワークシートを提供している。学習者コーパス的な観点からの例もあり、学習者のライティングをコンコダンサーで分析し、頻繁に起こるエラーを自己診断的な発想で利用する再学習のための方策を提言している。

第5章も、かなり特定化された例とはいえ、コンコダンサーを授業の中で使用できる19の実践的なシラバスを提供していて大変役に立つ。第6章では、外国語としての英語教育で上中級レベル学習者

や Cambridge Proficiency Examination を受ける学習者を対象として文学を教える方策を提示している。取り上げている題材は、簡易な語彙で構成されている Katherine Mansfield (1923) の “The Fly” (2,100 語) とやや難解な Samuel Beckett (1967) の “Text for Nothing, I” である。前者では、スキヤニング読解、内容に関する簡単な Q&A、話の流れを考えさせるグループ作業、発見した情報と意見をまとめるライティングなどの方法を示し、後者では、語彙リストとその出現頻度による理解、スキヤニング読解、異なった内容のプリントアウトを利用したグループ理解とストーリーの再現、グループ代表者による内容報告とまとめライティングといった方法を紹介している。

Kettleman (1994) は、アメリカ文学作品の一つ Mary Freeman Wilkins (1890) “The Revolt of Mother” を題材に、ロングマン・ミニ・コンコーダンサー (LMC: Longman Mini-Concordancer) とオックスフォード・マイクロコンコード (OMC: Oxford MicroConcord) を使って実験した結果を報告している。テキストをスキヤンした後、LMC での分析ではトークン・タイプ比が 4.5 でやや語彙のバリエーションが低かったが、Grammatik V で文体チェックをした結果では、Flesch Reading Ease スコアが 83 で読みやすく、平均的な文の長さは 12.8 で比較的短く、容易な統語構造で、一語あたりのシラブルも 1.3 で出現頻度が高い文章と言う。その結果を踏まえ、機能語と内容語の違いや定冠詞と非定冠詞の使い方を説明できたり、出現頻度の高い内容語から作品の内容を考えさせたりする方法を提示している。OMC での分析では、使用されている動詞や副詞のスタイルをチェックし、主人公の行動分析をし、伝統的な男女の役割を理解させたり、登場人物の詳細な人物像と感情を説明している。こういったコンコーダンサーと様々な文学作品テキストを頻繁に使うことによって、英語力の向上と文学を楽しむ方策を提示した。

Kettleman (1995) は、OMC とそのコーパス A&B (210 万語) を使い、英語教育における文法、語彙習得に役立つための方策を示している。具体的には、Kettleman (1994) で取り上げた同じ文学作

品の分析結果を加えながら、ドイツ人英語学習者が対象で文法学習上の問題点がある if 構文、reported speech、時制、for and since を取り上げ、語彙習得では、文脈から複数の意味を帰納的に学習するという観点からの例を幾つか挙げている。

Stevens (1995) は、多くのコーポラとコンコーダンサーを使った研究事例を引用しながら、言語学習者は、帰納性、オーセンティシティ、学習に対する自らの責任を求める教材から多くを学べると主張している。特に、こういった学習法を導入するに当たって、教師及び学習者の役割と責任の所在の在り方を訓練する必要があると主張し、CALL の研究発展上で生じている問題を次第に解決していくだろうと予想する。

Tribble (1997a) は、大規模コーパスの利用と共に、CD-ROM 百科事典のテキスト (マイクロ・コーパス) とコンコーダンサー利用による文法教育教材の例を挙げて、その有効性を主張している一方、個々の教師が構築した小規模なコーポラであっても発見学習スタイルの語彙・文法学習やプロジェクトにおける有益な資源となると主張している。執筆時点で利用できるコーポラを挙げ、その利用法を示しながらも、英語教育における純粋な言語サンプルが必ずしもオーセンティックな言語使用そのものや効果的な言語学習活動を助長するとは限らないと言う。学習者コーパスについては、種々の問題点を抱えてはいるが、英語学科や英語学校に眠っている資源である英語学習者が書いた標準的でないコーパスの利用を奨めている。

Tribble&Jones (1997) の 1990 年からの改訂版では、第 4 章の中で、具体例は同じ一つだけであるが、重要度が増してきた学習者コーパスの利用方法を独立させて論じている。また、第 7 章では、利用するコンコーダンサーが、Word Perfect 5, Basic, WORD 4 から、DOS のパラグラフ・コンコーダンサーと WordBasic で書かれたコンコーダンサーに置き替わられている。

Dagneaux, et al. (1998) は、学習者コーパスを利用した誤用分析に関する論文で、1970 年代に盛んになった誤用分析の欠点をふまえた上で、コンピュータを利用した学習者コーパスの誤用分析を提案

している。フランス語を母語とする英語学習者のコーパスを、International Corpus of Learner English（上級者）からと、著者ら独自のもの（中級者）とから15万語集め、エラー・タグをつけた。エラー・タグは言語学的な記述とし、エラーの原因などの分類は主観的になるのでつけなかった。作業は、まず、英語母語話者が、間違い部分に訂正した語句をつけ足し、次に、その部分に英語教師がタグをつけていった。タグづけには専用のツールを使った。出来上がったコーパスを使った成果としては、例えば、中級者と上級者のエラーの割合を比較することにより、学習の発達の様子を把握することができる。エラー・タグなしの学習者コーパスでは、分析者が思いもよらない使い方をしている例や、使うべきところに使っていない例などを検索することはできない。従って、エラー・タグがあってこそ、そういう例を漏らさず分析することができるようになる。

Turnbull & Burston (1998) は、非英語母語話者であるモナッシュ大学大学院生2名の書いたものをコンコーダンス・ストラテジーを使って英語を学習する際に検証するケース・スタディを報告している。特に、特異な教育環境で学習者自身がどの程度の自学自習ができるのか、適当な研究方法を開発するために、どれほどの手助けが与えられるかあるいは与えるべきか、こういった新しいアプローチがこれまでの教授法に取り込めるかの3点を検証したものである。このケース・スタディの結果からは、コンコーダンス・ストラテジーでの学習で効果を認めつつも、コンコーダンス・データだけでは必ずしも演繹的な学習ストラテジーを助長しないという結果を得ている。この結果から、言語学習に関する以前の経験からの影響や個々の学習スタイルがあるので、コンコーダンス・ストラテジーを使う前の十分なガイダンスや実践訓練の必要性を主張している。

2.2. 日本での研究

本稿の執筆時点では、斉藤ほか(1998)、鷹家・須賀廣(1998)、井上(1999)などの関連する英語母語話者の中・大規模コーパス利用研究は数多くあるものの、日本での学習者コーパスを利用した英語教育への試みについての研究報告はほとんどないと

言っても良かったが、小規模学習者コーパスの英語教育への応用などの先駆的な研究は幾つかされている。

杉浦(1993)は、コーパスを利用した言語学的手法を英語研究に応用し、帰納的アプローチによる実例を示し、限られた小規模学習者コーパスでも利用可能なことを示した。

Ozeki & Sugiura (1997) は、マッキントッシュ・コンピュータ上でハイパーカード(HyperCard)とKWIC(Key Word In Context)コンコーダンス・プログラムを使って、英語母語話者が話している映像クリップに出てくる会話語句を帰納的に学習する方法を提言している。これは非英語母語話者である日本人英語学習者自身が録画・録音したものを題材としていないので、学習者コーポラの直接的な利用ではないが、英会話パターンなどの例文を発見学習のスタイルで分析させ、文法知識の運用能力を高めさせる方法である。即ち、まだ主流をなしている自動化された「チュートリアル」スタイルやドリル練習の演繹的な学習教材のCAI教材とは大きく異なるものであり、コンピュータと「生きた」データベースがあってこそ可能な学習スタイルという点では、学習者コーパス利用の英語教育・学習への基本の一例を示している。

斉藤ほか(1998)の第12章は、英語教育へのコーパス利用を扱っている。日本では、文系でのコーパス利用以前に理工系でのコーパス利用が「英語を書くためのツール」として利用されてきているが、科学技術論文、表現辞典、検索ツールについて内外の例を幾つか挙げながら論じている。学習者コーパスと第二言語習得研究との関連からの研究(運用語彙の発達、文法項目修得の過程の分析など)については、執筆担当者である朝尾氏(東海大学)が小規模学習者コーパス(ライティング・データ)の平均語数やタイプ/トークン比から分析し、「英語を書く力(writing proficiency)」がある程度成長していることを示している。しかし、ライティング・テーマによっては、平均語数に伸びがみられたものの、タイプ/トークン比に変化が見られないとの結果も同時に示している。

『現代英語教育』に連載した投野(1998/99)は、

過去5～6年で学習者コーパスの研究が盛んになってきた経緯から、世界的には規模や内容的に充実したものは少ないと報告しながらも、本格的な整備がなされれば、英語教育においても大きな恩恵に浴することができると主張し、学習者コーパスの利用の分野を大きく3つに分けて、具体例を挙げながら論を展開している：(1)文法・語彙指導、(2)第二言語習得研究、(3)辞書・教科書・教材開発。(1)では、指導現場で学習者コーパスが威力を発揮するケースを挙げ、学年ごとにあるいは到達度レベルごとに学習者の単語使用をより客観的に見極めたり、文法的・語彙的エラーに関する情報から分析し、その頻度、エラーの種類、学習段階との係りなどを把握できるので、教師側にとって有益な情報資源になるばかりでなく、学習者自身が活用することによって、和英辞典などに勝る「表現辞典」として利用でき、自身が書いたものと比較検討しながらの発見学習ができると主張する。(2)では、学習者コーパスは、学習段階ごとにコーパス・データを整備することに

より、中間言語の記述に役立つ基礎データを提供できるとし、中高の文法的形態素の定着度の推移を見た例や、エラー分析の手法で母語が異なる学習者コーパス間で共通に起こる「普遍的エラー (universal errors)」と母語の干渉による「母語関連エラー (L1 related errors)」の仕分けをする研究を紹介している¹⁾。

朝尾氏を研究代表とし、筆者を含めたプロジェクト・グループによる「日本人英語学習者コーパス」プロジェクトが1997-99年度文部省科学研究費助成金に採択され、本格的に始められている²⁾。コーパス・データの増殖、入力フォーマットの標準化、標準エラー・タグの検討を行っている段階である。書き言葉コーパス (written corpus) については相当量の構築ができたが、話し言葉コーパス (spoken corpus) についてはまだ不十分な状態であるのは否めない。詳しくは、Corpus of English by Japanese Learners (<http://www.lb.u-tokai.ac.jp/lcorpus/>) にアクセスされたい。

Location: <http://www.lb.u-tokai.ac.jp/lcorpus/>

Corpus of English by Japanese Learners

にほんご
Japanese

CEJL Project	CEJL Corpus Data	How to Use a Corpus	Publication
Learner Corpus Projects	Corpus Linguistics Pages	Corpus Sources	Corpus Linguistics Scholars

The number of visitors to this page since December 19, 1996 **002682**

Contact Address < elc@lang.nagoya-u.ac.jp >

3. 学習辞典編纂への応用

斉藤ほか（1998）の第12章では、学習者コーパスを作成し利用する目的の一つに、学習辞典編纂のための基礎資料収集も挙げている。基本的な英語辞典編纂に利用されるのは、英語母語話者のコーパスからであるが、Longman Dictionary of Contemporary English 第3版（LDOCE3）のように学習者コーパスを取り入れたものも出てきている。そのLongman Learner's Corpusから得られた知見はUsage Noteにおける解説に利用されている。

外国語学習辞典は、母語話者用の辞典（monolingual dictionary）とは大きく異なる。外国語学習辞典は、英語のような目標言語の分析だけでなく、対象学習者によって編纂されなければ、真に役立つ辞典とはならないからである。その意味でも、外国語学習辞典編纂に役立つ大規模な日本人学習者コーパスの構築が望まれるし、筆者達のグループによるものも含めて、あちらこちらでその試みが始まっている。

4. おわりに

英語学習・教育の利用できるツールとしてのコンコーダンスや大容量のデータベースとしての学習者コーパス利用の研究や実践はまだ始まったばかりで、それらを利用した研究や実践報告はまだ少ない。しかし、今後の応用言語学研究の大きな柱と成りうる分野であると言えよう。投野（1999）が課題を指摘し、今後の展望をしているように、教育目的に特化したコーパスの必要性、テクノロジーと教育との融和、言語学習システム全体の改革、データ、ツール類の共有・啓蒙などを積極的にしていかなければならない。演繹的な学習から機能的な学習へ、あるいは経験主義的な英語教育から実証主義的な英語教育への転換を促進させるために、学習者コーパス自体の更なる構築とそれらの研究と成果の公開、英語教育現場での積極的利用は重要な役割を演じることになるであろう。

注記

- 1) 本稿は1998年10月に岡山女学院大学（愛知県名古屋）で開催された英語コーパス学会第12回大会におけるシンポジウムで発表した資料に加筆・修正したものである。
- 2) 詳しくは<http://info.ox.ac.uk/bnc/>にアクセスのこと。
- 3) 詳しくは<http://titania.cobuild.collins.co.uk/>にアクセスのこと。
- 4) 投野氏は本稿執筆時点で連合王国のランカスター大学大学院博士課程在学中であり、<http://www.lancs.ac.uk/postgrad/tono/>で世界中の多くの関連情報を集積し提供している。
- 5) 正式には科学研究費補助金－基盤研究(B)「第二言語習得研究のための英語学習者コーパスの構築とその利用」課題番号09558018である。

参考文献

- 井上永幸（1999）. 「コンピュータを使った英語教師のための英語研究」『Cornucopia』 Winter, 3-6.
- 斉藤俊雄ほか編（1998）. 「英語コーパス言語学-基礎と実践」研究社出版
- 杉浦正利（1993）. 「コーパスを利用した言語学の英語教育への応用」『Language Laboratory』 語学ラボラトリー学会, 30, 95-113.
- 鷹家秀史・須賀廣（1998）. 「実践コーパス言語学 - 英語教師のインターネット活用」 桐原ユニ
- 投野山紀夫（1998/99）. 「学習者コーパスと英語指導」『現代英語教育』 研究社出版, 1998年4月号～1999年3月号
- 投野山紀夫（1999）. 「英語教育におけるコーパス利用-最新動向レポート」『Cornucopia』 Winter, 7-13.
- Dagneaux, E., S. Denness, & S. Granger（1998）. "Computer-Aided Error Analysis." *System*, 26, 163-174.
- Flowerdew, John. (1996). "Concordancing in the classroom" in Martha Pennington (ed.), *Power of CALL*, Houston, TX: Athelstan, 97-113.
- Kettleman, Bernard. (1995). On the use of concordancing in ELT, *TELL&CALL*, 4, 4-15.
- Kettleman, Bernard. (1994). *Concordancing in stylistics teaching*, Festschrift zum 30-jährigen Bestehen des Instituts für Anglistik der Universität Salzburg, Salzburg.
- Ozeki, Shuji & Sugiura, Masatoshi. (1997). Quick-KWIC-Dialog: A hypermedia-based corpus for EFL learners. (<http://langue.hyper.chubu.ac.jp/ozeki/CALICOQKD.html>)
- Stevens, Vance. (1995). Concordancing with language learners: Why? When? What? *CAELL journal*, 6(2), 2-10.
- Sugiura, Masatoshi. (1998). Web based dynamic reference

- system for English as a foreign language learners. (<http://oscar.lang.nagoya-u.ac.jp/paper/presentation/calico98/>)
- Tribble, Chris. (1991). "Some uses of electronic text in English for academic purposes" in J. Milton & K. Tong (eds.) *Text analysis in computer assisted language learning*. Hong Kong: Hong Kong Univesity of Science and Technology, 4-14.
- Tribble, Chris. (1997a). "Improvising corpora for ELT: Quick-and-dirty ways of developing corpora for language teaching." A paper presented at the first international conference: Practical applications in language corpora (1997), University of Lodz, Poland.
- Tribble, Chris. (1997b). "Corpora, concordances and ELT" in Boswood T. (ed.) *New ways of using computers in language teaching*, Alexandria, VA: TESOL.
- Tribble, Chris & Jones, Glyn. (1997). *Concordances in the classroom: A resource book for teachers*. Houston, TX: Athelstan.
- Tribble, Chris & Jones, Glyn. (1990). *Concordances in the classroom: A resource book for teachers*. London: Longman.
- Turnbull, Jill & Burston, Jack. (1998). Towards independent concordance work for students: Lessons from a case study. *ON-CALL*, 12(2), 10-21.
- 世界の主要な学習者コーパス関連リンク
- International Corpus of Learner English (ICLE)
<http://www.fltr.ucl.ac.be/FLTR/GERM/ETAN/CECL/abs.html>
 - HKUST Learner Corpus (John Milton)
No URL: It hasn't made the data publicly available yet, as it's still being used for J. Milton's Ph.D research.
 - Longman Learner's Corpus
<http://www.awl-elt.com/dictionaries/lclearn.html>
 - Corpus of Japanese Learners of English (Kojiro Asao's Group)
<http://www.lb.u-tokai.ac.jp/lcorpus>
 - Resources on Learner Corpus (Yukio Tono)
<http://www.lancs.ac.uk/postgrad/tono/lclresource.html>
 - Virtual DDL Library (Tim John)
http://web.bham.ac.uk/johnstf/ddl_lib.htm