

〈特別寄稿〉

## 学校でのコーパスの活用

東京外国語大学大学院 教授 投野 由紀夫

### 1. はじめに

「コーパス (corpus)」という言葉は、10年ほど前に比べると英語教育に携わる者にはかなり身近な用語となってきた。私がNHKで『100語でスタート！英会話』を担当したのが2003年。その当時は一般の英語教師には「コーパス」という言葉はまだ耳慣れなかったはずだ。現在はコーパスを使った語学教材や辞書もかなり豊富に出揃ってきたし、一般の人が触れることのできる便利なコーパスも登場し、日本の言語教育におけるコーパス利用を推進してきた筆者としては隔世の感がある。

しかし、用語は知っていても実際にコーパスに直接触れたり、その恩恵を実感したりしている英語教師はまだ少数なのではないだろうか。そこでこの寄稿では、学校におけるコーパス活用の視点とその導入法に関して考えてみたい。

### 2. 英語力の「エッセンス」を知る

コーパス利用でまず英語教師に知って欲しいのは、「英語を使える」という能力の中核は何か、である。これを知るためには英語母語話者の言語使用の最もコアになる部分の力をコーパスで調べてみるとよい。そうすると、だいたいどんな会話でもその9割は2000語程度の基礎語彙でまかなえること、その2000語のうち半分は名詞であること、かつ2000語のうち最もよく使う100語で会話の7割近くを占有し、その100語のほとんどが基本動詞と機能語（代名詞・前置詞・冠詞・助動詞・接続詞・副詞など）であること、などがわかってくる（投野2006）。

英語の文の骨組みは「主語＋動詞＋X」という型で始まり、動詞が決まることでXが決まる。動詞の使いこなしが文の骨組みを決めるわけだが、実は have = 「持っている」というだけでは不十分だ。「have＋X」と言った場合の X の要素のヴァリエーションをどれほど知っていて活用できるかが have の使いこなしの力量になるわけである。ちなみに have の後ろの名詞には中学入門期に練習する I have a pen. のような具

体名詞よりも、time, idea, problem, のような抽象名詞の方がずっと高頻度に現れる。それらを習熟させるように教材化したのが、NHKの『100語』であり、投野（2008）などの一連のコーパス準拠教材であった。

こういった教材をもとに、英語力のエッセンスに関する正しい認識が持てると、教科書の各レッスンを機械的に教えるというような単調な授業ではなく、英語力の「幹」をしっかりと太くする指導とはどのようなものかを考えながら授業計画を練るようになる。その「視点の変化」が重要なのである。

### 3. 生徒の使う英語教材の「目利き」をする

英語力の中核が見えてくると、指導にもメリハリが出てくる。必ず教えるべき項目とそうでないもの、絶対に活用できるようにしたい単語と覚えなくてもいい単語など、生徒の触れる教材の有用度を英語教師が大胆に仕切ってやるのである。その基礎資料を提供するのもコーパスである。

私は『エスクラウン英和辞典』（2008、三省堂）を編纂した際に、この「仕切り」を大胆にやった。トップ100語の基本語彙は学習辞典の常識を打ち破る2ページ見開きのフォーカスページという特集記事を随所に入れて、コーパス・データを駆使した単語のプロフィールを作り、「最低限これだけはやる」という仕切りを示した。そしてトップ2000語までは活用語彙、5000語までは認識語彙として、高校卒業までに身につけて欲しい単語として位置づけ、それ以外のランク外の単語には大幅に紙数を減らして「意味だけわかればよい」「覚えなくて辞書を引けばよい」という単語であることを明示した。こういった「仕切り」が学習上重要なのである。

日本の高校英語教科書をコーパスで分析した際にも、教科書にはたまたまある教科書を選んだから出てきた、というような単語が全体の40%もあることがわかった。つまり高校レベルの教科書では、そのレッスンで出てきた新出単語のうち10個に4個は覚えなくてもよい単語の可能性があるのである。新出単語をすべて

\*山形県英語教育\*

単語テストで無理矢理覚えさせている教師は罪深い。  
「この単語は忘れてもいい」という勇気と確信を教師は持つべきである。そういうことを正しく「目利き」  
できるような先生に生徒は絶大な信頼を置く。

#### 4. 学校でそろえておきたい英語コーパス

学校で実際にコーパスを使ってみたいという場合、  
基本的には以下のようなセットを勧めたい：

- (a) 自然な英語をチェックできるコーパス
- (b) 学習者のエラーをチェックできるコーパス
- (c) 教科書・入試問題などのインプットのコーパス

この3つをそろえておき、初めは教員の教材研究や  
タスク開発のための資料として参照し、だんだんと環  
境が整ってくれば実際に生徒に調べさせてみるという  
のも面白い。それぞれのコーパスに関して代表的なも  
のを紹介しよう。

(a)に関しては、英語母語話者のコーパスになる。現  
在最も信頼できる大規模コーパスとしては、イギリス  
英語の1億語のコーパス British National Corpus  
(BNC)、または4億語のアメリカ英語コーパス Corpus  
of Contemporary American English (COCA) がある。  
それぞれWebページで検索可能なサイトが整備され  
ている(参考URL参照)。

また有料にはなるが、親しみやすい日本語インタ  
フェースとして小学館コーパス・ネットワーク(SCN)  
が提供するBNCおよびWord Banks Online (COBUILD  
英語辞典の元データとなった英米語コーパス)のサー  
ビスも使いやすい。

(b)に関しては、筆者が無料で公開しているJEFLL  
Corpus(1万人の中高生英作文コーパス)が参考にな  
るだろう。これで該当する単語や文法のフレーズを検  
索すれば、さまざまな誤用を含む例文を閲覧して、よ  
くやる間違いなどの観察をすることができる。

(c)に関しては現状では既製品はない。そこで英語教  
師であれば自分でテキストを集めてコーパスにすること  
をお薦めする。教科書であれば、教科書会社からの  
データが使える。個人閲覧用であれば検索しても著作  
権上問題ない。入試問題もCDで入試問題のテキスト  
を対策用に販売している会社がある。中学・高校の英  
語入試問題を収集するのはそれほど難しくない。

自分のパソコンに保存したテキストを検索する時  
にはフリーの検索ソフト(コンコーダンサーという)を

使おう。AntConcというツールが無料で最も使いや  
すい(参考URL参照)。

#### 5. 毎日の活動とコーパスをリンクする

コーパス検索に少しずつ慣れてきたら、1時間の授  
業の中で何か1つはコーパス・データからの有用な情  
報を盛り込む工夫をしよう。文法事項ならば、(b)の  
コーパスを調べてつまずきやすいポイントを解説して  
もいいし、会話フレーズならばCan I have ~?の  
~に入る名詞を(a)のコーパスで調べてドリルに盛り込む  
のもいい。

高校レベルでCALL教室が使えるならば、新しい文  
法事項の導入をいきなりルール提示から入るのではな  
く、(c)の教科書や入試問題コーパスから10個くらい例  
文をまとめて取り出し、そのコンコーダンスラインを  
見せながら、文法規則を類推させるような発見学習的  
なアプローチも試してみたい。

最後にライティングのような発信活動で、生徒自ら  
がコーパスに触れて例文を考えたり出来る環境があれば  
素晴らしい。その作文データを教師が保存しておい  
て自分のクラスの作文コーパスを作りながら、指導に  
活かせれば、まさしくコーパス主導で行うアクショ  
ン・リサーチに他ならない。

「英語指導をデータ駆動型(data driven)に転換し  
ていく」これがコーパス準拠の英語教育の醍醐味だと  
言えよう。ご健闘をお祈りする。

#### 参考文献

- 投野由紀夫(2006)『コーパス超入門』小学館
- 投野由紀夫(2008)『コーパス練習帳パーフェクト』  
NHK出版
- 参考URL
- AntConc  
[http://www.antlab.sci.waseda.ac.jp/antconc\\_index.html](http://www.antlab.sci.waseda.ac.jp/antconc_index.html)
- British National Corpus  
<http://corpus.byu.edu/bnc/> (BYU版)  
<http://phrasesinenglish.org/searchBNC.html>
- Corpus of Contemporary American English  
<http://www.americancorpus.org/>
- 小学館コーパス・ネットワーク  
<http://www.corpora.jp>
- JEFLL Corpus  
<http://scn02.corpora.jp/~jefll04dev/>