

## はじめに

**学** 習者コーパスは、第1回でも記したように、その応用範囲は専門的な第2言語習得研究から明日の授業にすぐ応用できる表現語彙リスト作成まで幅広い。毎回、いろいろな角度から学習者コーパスを用いた分析結果を見て、そこから英語教育への示唆や提案をしてみたい。今回は、そんな事例の1つとして、学習者コーパスをもとにした topic vocabulary の選別やリスト化、その活用を考えてみよう。

## どんな語彙を使うか？

Topic vocabulary は、ある話題に関して自分のアイデアを表現する時に必要になる語彙である。読者の先生方は、毎回の授業で生徒に自己表現活動をさせたい時に「このトピックならどんな単語をあらかじめ与えておけばいいかな？」と迷ったことがおありではないだろうか？教科書の題材などに合わせて、自己表現のトピックもうまく考えついたとして、円滑に生徒に表現活動をさせるには単語の問題がどうしても出てくる。

こんな時に、整備された学習者コーパスがあれば、あるトピックに必要な語彙リストを比較的簡単に抽出して、授業に応用できる。「そんなことならば、10年も昔からとっくにやっている」とおっしゃる先生もおられるだろう。しかし、大切なことは学習者コーパスを作ることによって、もっと体系的・組織的に表現語彙リストに関してアプローチすることが可能だという点だ。一般性の高い話題別語彙リストなどは今後、教材作成や英語表現辞典のような分野の基礎資料として非常に重宝がられるだろうと思われる。

さて、では次に実際の topic vocabulary の抽出方法を紹介してみよう。

## キーワード分析による抽出

あるジャンルのテキストからキーワードを抽出

する方法はいろいろ考えられるが、現在最も頻繁に行なわれている方法は、基準となる標準コーパスがあってその語彙の頻度リストが与えられており、あるジャンルのテキスト中にその基準コーパスに照らして著しく頻度の高い単語があった時、それをキーワードとして抽出するという方法である(Scott 1997 参照)。

学習者コーパスが理想的に整備されていけば、各学年ごとのデータ以外に、学習者の表現したいアイデアやトピックごとのファイル情報で検索が可能になるだろう。例えば、文化祭とか学園祭についての英文ばかりを集めたデータがあれば、上記のような topic vocabulary のリストは比較的容易に作成できる。他にも、クラブ活動、好きな食べ物、将来の夢などトピックごとに検索できるようになっていけば、これから紹介するツールを使うことでかなり自作の語彙リストを容易に作れるようになる。

## WordSmith を用いたキーワード分析例

それでは TGU Learner Corpus (東京学芸大学で筆者が整備していたコーパス)のデータの一部分を用いて、WordSmith というソフトでキーワード分析を行なってみよう。WordSmith は Oxford University Press から発売されている Windows 用コーパスツールである。なかなか高機能で、専門的な知識があまりなくても使いこなせるので是非入手してみるとよいだろう\*。

例として、食事に関するキーワードリストをコーパスから抽出してみよう(実際には「朝食にはパンがいいかご飯がいいか」というタイトルで書かせたエッセイのコーパス)。このためには、まず基準コーパスの頻度データが必要になる。今回は実験的に、個人的に百科事典などのテキストを独自にコーパス化したデータ(500万語くらい)を使ってみた。この単語頻度データを WordSmith で作っておいて、wordlist ファイルとして保存しておき、これを基準に食事に関する作文コーパスのキーワードをピックアップすればいい。

図1が WordSmith の分析結果の画面である。

# 学習者コーパスと

WORD	FREQ	BASE FREQ	LOG LIKELIHOOD	P	KEYNESS
I	2,617	10.14	3,025	0.05	20649.7
BREAKFAST	688	2.67	59		7004.8
MORNING	482	1.87	169		4448.3
BREAD	491	1.90	269		4304.1
RICE	551	2.14	704	0.01	4222.4
EAT	540	2.09	813	0.01	4005.6
DONT	355	1.38	133		3252.0
HAVE	986	3.82	17,163	0.31	3133.2
USUALLY	376	1.46	1,738	0.03	2085.5
MISO	180	0.70	1		1924.6
MY	253	0.98	396		1858.9
SO	409	1.59	3,391	0.06	1840.1
LIKE	340	1.32	2,208	0.04	1678.9
SCUP	178	0.69	56		1658.4
MILK	272	1.05	926	0.02	1652.9
VERY	234	0.91	613	0.01	1525.7
HUNGRY	131	0.51	51		1194.0
THINK	158	0.61	303		1110.2
BUT	460	1.78	12,849	0.23	1072.4
LUNCH	111	0.43	28		1054.7
GET	152	0.59	769	0.01	817.5
IM	96	0.37	69		809.1
CANT	85	0.33	36		767.4
EVERY	179	0.69	1,868	0.03	729.1
BECAUSE	229	0.89	4,142	0.07	706.3
ME	102	0.40	269		663.4
BETTER	118	0.46	775	0.01	579.3
WANT	100	0.39	430		566.4
AM	79	0.31	164		544.8
ALWAYS	101	0.39	580	0.01	520.2

図1 WordSmithの分析結果

Keyness というコラムは、基準コーパスと比較して、ある語がそのテキストの中でどれだけ特徴的であるかを示しており、この値が高ければ高いほどキーワード性が強い。第1位にはI(私)が来ている。これは作文コーパスが一般のコーパスに比べて、自分自身のことを書いたものが多いことを意味する。ここでは「食事」に関するキーワードをリストするので、こういった一般語はどんどん編集してカットする。WordSmith上ではこのような編集作業も簡単に出来る。

編集したものを元に作った単語リストが表1である。書かせたエッセイのテーマによって多少単語に偏りがあるが、それは食事に関するコーパス・データをもっとバランスよく収集できれば解決する問題である。ここでは30語のリストしか示していないが、一瞬にして200語くらいのリストが出来てしまうのだから、その便利さは大変なものである。

表1 キーワード分析による上位30語

N	WORD		
1	BREAKFAST	16	HAYABEN
2	MORNING	17	COFFEE
3	BREAD	18	CUP
4	RICE	19	DINNER
5	EAT	20	MISOSHIRU
6	MISO	21	EGG
7	LIKE	22	DELICIOUS
8	SOUP	23	NATTO
9	MILK	24	HEALTH
10	HUNGRY	25	COOK
11	LUNCH	26	MISOSOUP
12	TOAST	27	MEAL
13	DRINK	28	BUTTER
14	EATING	29	SLICE
15	TEA	30	FOOD

# 英語指導

.....(3) Topic Vocabulary Listを作る

words 投野由紀夫

N	Concordance
46	I usually have rice and miso soup as breakfast. Because I need to bring a be
47	So I make it a rule not to eat bread as breakfast.
48	I usually eat bread and milk as breakfast. But I don't have any remarks
49	have rice. I have never eaten bread as breakfast for many years. I think bread
50	school. My father only takes drink as breakfast. I think it is one way to hav
51	me, it't not the problem what I have as breakfast. I'm satisfied enough if I ca
52	morning, but I like sleeping as as breakfast. it is difficult for me. I'm
53	So I want to eat rice and つけもの as breakfast. We should eat rice and つけ・
54	I get up late, I have bread and milk as breakfast. Why I use "as"? Because I d

表2 have A as breakfast のコンコーダンス

それだけではない。例えば、breakfast, lunch, dinner, meal などの単語は高頻度なので、当然これらを用いた「食事をする」という表現を適切に使えるようになって欲しい。そんな時、実際に学習者がどのような作文を書くのか、これらの単語を用いた例が瞬時に検索できる。WordSmith では単語リストからすぐに該当する用例コンコーダンスに移動できるので非常に便利だ。

breakfast の使い方で気になるのは、「A を朝食としてとる」という表現で、have A for breakfast をかなりみなきちんと使っているが(688 用例中 53 例)、have [eat] A as breakfast (8 例)という誤用も目立った。表 2 がその使用例である。

このような誤用の分析はもっと学年ごとに精密に行なってみると興味深い発達のな特徴がわかったり、この表現がこの時点でたくさん出てくる深刻な誤りか否かなどに関してもかなり客観的なデータを提示することが出来る。Topic vocabulary を与えるだけでなく、このような活用例文や誤りやすいポイントについても学習者データをもとに指導・助言できるのだから、この恩恵は計り知れない。

以上のように、キーワード分析による topic vocabulary list の作成が容易に出来るうえ、それらを活用語彙として生徒にコロケーションごとに練習させようと思ったりした時に、学習者コーパスが完備されてアクセスが可能であれば、こういった情報は即座に自分で検索して教材作成や指導に生かすことが出来るのである。

また学習者コーパスの整備が進めば、このよう

なキーワード分析を先生方がソフトを使って四苦八苦してやる必要はなくなるだろう。ほとんどの基礎データは大学の研究者サイドでどんどん整理した形でホームページなどを通して利用可能になるから、今後 5~10 年後にはかなり信頼できる語彙表や topic vocabulary の一覧が Web 上で公開されたり、自分の教材作成用にダウンロードできたりするようになるだろう。

### 終わりに

今回は、topic vocabulary list の作成を通して、学習者コーパスの整備によってどのような学習者支援の教材作成が可能か、それが単なる教師の「カン」ではなく、実際のデータに基づくことによる新しい発展の可能性を示してみた。学習者コーパス構築が是非とも順調に進むように、全国レベルでの展開と協力を引き続きお願いしたい。今回は応用例のうち、もう少し語彙習得の基礎研究に焦点を当てて、基本動詞の習得の特徴を中心に試してみることにしよう。

<註>

・コンタクトの URL は Oxford University Press の WordSmith のページ

(<http://www1.oup.co.uk/elt/catalogu/multimed/4589846/4589846.html>) 参照

<参考文献>

Scott, M. (1997) "PC Analysis of Key Words ... and Key Key Words." *System*, Vol. 25, No. 1, pp. 1-13.

(とうの・ゆきお / 元東京学芸大学講師

ランカスター大学博士課程在籍)

学習者  
コーパスと  
英語指導