



英語で言いたかった表現を集めた日英パラレル・ コーパス : 高校生が発信のために必要とする語彙 : EasyConcの開発と活用

日臺, 滋之

(Citation)

Journal of Corpus-based Lexicology Studies, 3:31-46

(Issue Date)

2021-02-28

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/81012539>

(URL)

<https://hdl.handle.net/20.500.14094/81012539>



英語で言いたかった表現を集めた日英パラレル・コーパス:

高校生が発信のために必要とする語彙

—EasyConc の開発と活用—

日臺 滋之(玉川大学)

hidai_ki@lit.tamagawa.ac.jp

A Japanese-English parallel corpus consisting of expressions Japanese students
want to use in English: Productive English lexis for senior high school students

—the compilation and utilization of ‘EasyConc’—

HIDAI, Shigeyuki (Tamagawa University)

概要

The purpose of this study is to seek out the vocabulary Japanese senior high school students find difficult to use to express themselves and to compile a Japanese-English parallel corpus. In order to compile the corpora, whose software is EasyConc.xlsm for Windows and Easy Conc.fmp12 for iOS, second grade senior high school students were asked to write the expressions they wanted to use in speaking tests but could not in Japanese. Later, students' questions are translated into simple English by the Japanese teacher of English and the native speaker of English. The result was that the majority of words that students wanted to use were included within 2000 high-frequency words. If teachers make use of EasyConc.xlsm to make worksheets for their English classes, then I believe that teachers can help students improve their speaking skills.

キーワード

授業の振り返り, 英語で言えなかった表現, 発信語彙, 日英パラレル・コーパス

1. はじめに

冬休み明けの中学校3年生の授業で, 冬休みの生活について二人一組になりチャットの活動を行い, その後の振り返りの活動として, 「英語で言いたかったけれど言えなかった表現」を生徒に日本語で書いてもらおうと, 「おせちを食べた」「おみくじを引いた」「初日の出を見た」といった表現が言えないことが分かった。全く同様に高校3年生に冬休み明けの授業で同じ活動をする中学生と同じ質問に加えて, 「三泊四日」「他に何をしましたかの『他には』をどう言ったらよいのか」「お節料理をおなか一杯食べた」「時給はいくらか」といった質問が寄せられたのである。日臺(2013)が指摘

するように、学年が上がると自然に英語の力が付き言えるようになるわけではないことがわかる。生徒が英語で言えないことを教師が知り、その表現を授業に取り入れ活かす指導を行うことで生徒の表現力を伸ばせるのではないかと考えている。文部科学省(2019:33)では、受容語彙と発信語彙とを区別し、「特に高等学校においては、中学校までに学習した 2,500 語程度をできるだけ発信語彙として用いる機会を豊富に提供することが重要である」としている。生徒の表現力を伸ばすためにどのような発信語彙が必要となるのかその選択についての研究が必要なのである。そのためにも、この英語で言いたかったけれど言えなかった表現からなるコーパスが必要であると考えている。

2. 先行研究

本研究は、高校生が英語で言えなかった表現を日本語で収集し、英語に直し、日本語と英語の対応する日英パラレル・コーパスを構築することを目的としており、英語で言いたかった表現語句としてどのような発信語彙を必要としているのか調査することである。

2.1 日英パラレル・コーパスの先駆的研究

先ず既存の日英パラレル・コーパスを概観する。国内で、公開されている日本語と英語の一対一対応のコーパスは決して多くはない。日本語と英語の一対一対応の先駆的研究として、「日英新聞記事対応付けコーパス」があげられる。現在、内山・高橋(2003)が構築した「読売新聞と The Daily Yomiuri」から自動作成された日英新聞記事対応付けコーパスのデータが入手可能である。この日本語と英語の対応データを Anthony(2021)の AntPConc, あるいは Barlow(2021)の ParaConc に取り込むことで日本語あるいは英語で検索することができる。この日英パラレル・コーパスは社会人や一部の大学生にとって時事問題について表現を調べるうえで実用価値は大変高いが、高校生が日常表現をどう表現したらよいのか調べるといったニーズには残念ながら合致しない。

2.2 学習者が英語で表現したかった語彙を調査するために日本語使用を許容するコーパス

投野(2007)が開発した JEFLL Corpus がある。JEFLL Corpus は、日本人中高生英作文1万件以上を収集した学習者コーパスが構築され、web 検索(http://scn.jkn21.com/~jefll03/jefll_top.html)でアクセスすることができる。特筆すべきことは学習者が英語で言えなかった表現は日本語で記述することを許容しており、どのような日本語を学習者が英語で表現できないのかわかる。本研究で構築しようとする日英パラレル・コーパスはこの学習者が英語で表現できずに日本語で表現してしまったことに着目し、後に英語に直しフィードバックすることに教育的価値を置き、その英語で言えなかった表現語句を探ろうとする研究なのである。投野(2007)によると JEFLL Corpus は、6種類の英作文テーマ urashima, earthquake, breakfast, festival, dream, otoshidama に限定されている。学習者コーパスとして大変有益なものであるが、6 種類以外のトピックについて中高生が自己表現としてどのような表現を難しく感じているのかについて話題が限定的であるために、話題のバラエティが必要である。

2.3 語彙の使用頻度別の分類と発信語彙について

Nation(2013:96)は、語彙を、high-frequency words, mid-frequency vocabulary, low-

frequency vocabulary, academic vocabulary, technical vocabulary の五大別に分類している。それぞれ、ワード・ファミリーで、high-frequency words は 2,000 語, mid-frequency vocabulary は 7,000 語, low-frequency vocabulary は 11,000 語, academic vocabulary は 570 語, technical vocabulary は 2,000 語に分類している。通常のテキストで、テキストカバー率は high-frequency words (88%), mid-frequency vocabulary (10%), low-frequency vocabulary (2%) で 100%を占めるとしている。日本の中学や高校の学習指導要領の発信語彙の選定を考慮するとき、high-frequency words の 2000 語の習得が極めて重要であることを示唆している。

3. リサーチデザインと手法

3.1 研究目的と RQ

本研究では、学習者が表現できたことを集めることを目的としたコーパスではない。高校生が英語で言いたかったけれど言えなかった表現をどのように調査し、調査結果からコーパスを構築し、そしてコーパスからの発見を学習者にフィードバックすることが目的である。コーパス構築過程は以下のとおりである。

高校 2 年生のスピーキングテストの直後に、活動の振り返りとして英語で言いたかったけれど言えなかった表現を日本語で書いてもらい提出する(図1A)。その後、高校 1 年生程度までの英語で生徒からの日本語を日本人英語教師と英語母語話者とで英訳する(図1B)。日本語の質問と英語の質問を Excel 上で一対一対応にする。続いて VBE(Visual Basic Editor)で検索機能をプログラミングした EasyConc for IzumiHS.xlsm (Windows 仕様)に取り込む(図1C)。

図 1 日英パラレル・コーパス EasyConc for IzumiHS の 3 部構成システム

- | |
|--|
| A. 英語で言えなかった表現をまとめた日本語表現データベース |
| B. 日本語表現を英語化したデータベース |
| C. 日本語表現およびその英語化したデータベースの同時検索システム(EasyConc for IzumiHS.xlsm) |

研究設問(RQ)として以下の項目を設定した。

RQ1: 英語で言いたかった表現を集めたコーパス EasyConc for IzumiHS にどのような特徴が見られるか。構築したコーパス EasyConc for IzumiHS に通常の大規模コーパス COCA には見られない特徴はあるか。あるとすればそれはどのような特徴と言えるか。

RQ 2: 構築したコーパスに出現する語彙と高頻度語彙(2,000 語)とを比較する。語彙を比較することを通して高校生の語彙の発信語彙を知る。発信語彙の視点から、学習者の表現能力を伸ばす教材作成につながると考える。高校生は同じ質問を繰り返すのか。もし繰り返すというのであれば質問の頻度が証明することになる。それはどのような表現かを明ら

かにする。

RQ 3: 構築したコーパスに文法標識のタグを付与することによって、どのような教材作成が可能となるか。文法タグを付与することにより文法用語からの検索の幅が広がる。文法の視点として、不定詞、接続詞、関係代名詞、仮定法等に着目しワークシートを作成することで検定教科書を補足する教材として文法指導に役立つのではないかと考えている。

3.2 データ

宮城県泉高校 2 年生を被験者として、渡邊崇教諭がスピーキングテストを実施した。概略は以下のとおりである。

【2018 年度スピーキングテスト】(2018 年度 2 月実施)

目的: Can-Do リストに基づき、普段の授業で培ったスピーキング力を評価

使用教科書と範囲: コミュニケーション英語 II Revised LANDMARK II L.8, L.9

後期スピーキングテストの形式: Group Discussion

被検者: 高校 2 年生の 3 クラスで実施。

方法: 生徒 3~4 名からなるグループに対して教師 1 名で、自分の意見を伝える力と相手とのやり取りする力を評価。

【2019 年度スピーキングテスト】(2019 年度 6 月実施)

目的: Can-Do リストに基づき、普段の授業で培ったスピーキング力を評価。

使用教科書と範囲: コミュニケーション英語 II Revised LANDMARK II L.1, L.2

前期スピーキングテストの形式: Summary Retelling + My opinion

被験者: 高校 2 年生 2 クラスで実施。

方法: 生徒 1 名に対し教師 1 名で、生徒はキーワードを用い、既習のストーリー全体をリテリングし、新情報と意見を加えて話す。その後、教師からの質問に答える。個人プレゼンテーション+インタビュースタイルの形式で評価。

上記二つのスピーキングテスト実施後に、調査票を配布し、英語で表現したかったけれどもできなかったことがあれば日本語で記入する調査を実施した。文脈が分かるように語句だけでなく文で記載するように依頼した。英語で言えなかった表現として、2018 年度では、112 件の質問が、2019 年度では、65 件の質問が日本語で寄せられた。図 2 を参照されたい。

図 2 スピーキングテスト直後の調査票と記入例

調 査 票

英語で言いたかったけれど言えなかった表現について

宮城県立泉高等学校 (2) 年生

実 施 日 平成(元)年(6)月(28)日

令和0

【英語で表現したかったけれどできなかったこと】

このアンケートは、コミュニケーション活動で、高校生のみなさんが英語で表現できなかったことについて尋ねるものです。

今日の授業であなたが、英語で表現したかったけれどできなかったことがあれば日本語で書いて下さい。

【記入の注意】 良くない例：人の気持ち 良い例：人の気持ちはお金では買えない。

※ 良くない例のように、「人の気持ち」だけだと何を英語で言いたかったのが、わかりません。良い例のように、日本語の文で書いてください。

【記入欄】

「私もしんごのように強く諦めることがいい
スポーツプレイヤーになりたい」

【研究へのご協力をお願い】

本アンケートへの回答は、本研究の目的以外に使用されることはありません。研究結果の公表にあたっては、個人が特定されることのないようにデータは全て匿名化されます。以上のデータ使用目的及び個人情報の取り扱いについて同意していただける場合は、以下の口に入れた上で、下記の署名欄にサインをしてください。

☒ 上記のデータ使用目的と個人情報取り扱い方法を理解し、その範囲でのデータ使用を承諾します。

署名 _____

3.3 手法

3.3.1 データの事前処理

図2の高校2年生の英語で言いたかったけれど言えなかったことについての日本語による質問件数は延べ数で177件であった。このデータの収集から、本研究のコーパス構築に至るまでの過程は以下のとおりである。

Stage 1：泉高校2年生スピーキングテストの実施。

Stage 2：スピーキングテスト実施後、生徒は英語で言いたかった表現を日本語で書く。

Stage 3：生徒からの質問をExcelのセルに入力。

Stage 4：生徒からの各質問を英語母語話者と日本人英語教師とで英訳作業。Excel上で日本語と英語との一対一対応データを作成。さらに、各データについてはトピックからも検索できるように15の話題を設定し、データの内容に応じて該当する話題を付加した。図3のD列を参照されたい。文部科学省(2019)の学習指導要領に「日常的な話題」、「社会的な話題」、「身近な出来事や家庭生活などの日常的な話題」という表現が散見するが、コミュニケーション活

動においてはどのような話題が学習者の関心事であるかを知るうえで話題の分類は大切と考える。本研究で絞り込んだ 15 の話題は以下のとおりである。

【15 の話題の分類】

1.日常生活 / 2.学校（勉強・勉強以外含む） / 3.仕事 / 4.旅行・買い物 / 5.ファッション（服飾・美容） / 6.趣味・娯楽 / 7.家族・友人関係（プレゼントほか含む） / 8.天気・気候・地震 / 9.健康 / 10.食事・食べ物・料理 / 11.地理・交通 / 12.政治・社会問題 / 13.日本紹介 / 14.日本語独特の表現（擬態語・擬音語なども） / 15.感情表現（喜怒哀楽、やる気、怠けるなども）

図 3 対応データからの抜粋例

A	B	C	D
1	AIが増えるということは私の仕事場が減るということだからとても恐ろしい。	The increase of AI is reducing the number of workers, so I am really worried about that.	12.政治・社会問題
2	初めは怖いものだと思っていたが、今では私たちの生活をよりよいものに変えるものだろう。	At first I was afraid of AI, but now I think it will change our life for the better.	12.政治・社会問題
3	家庭内で働いてくれるAI良いが、人の仕事をうばうAIはよくない。	AI helping us with our housework is good, but I don't like AI taking (or stealing) our jobs from us.	12.政治・社会問題

Stage 5: 日本語または英語で検索するための工夫

Stage 4 のデータを VBE(Visual Basic Editor)でプログラミングした EasyConc for IzumiHS.xlsm (Windows 仕様) に取り込む。

Stage 6: 文法用語で検索するための工夫

Stage 5 の EasyConc for IzumiHS.xlsm の英語部分を, Free CLAWS web tagger (<http://ucrel-api.lancaster.ac.uk/claws/free.html>) に送信し, CLAWS7 の文法標識でタグ付けし, VBE でプログラミングした EasyConc_tagged for IzumiHS.xlsm (Windows 仕様) に取り込む。

Stage 7: 表現・語句あるいは文法用語で検索し, 英文と日本語を抽出し, 教材作成を行う。

3.3.2 分析の手順

RQ1 英語で言いたかった表現を集めたコーパス EasyConc for IzumiHS にどのような特徴が見られるかについては, 大規模コーパスの COCA と比較して論じたい。

RQ2 コーパスに出現する語彙と高頻度語彙(2,000 語)との比較については, AntWordProfiler を使って, 高頻度語彙との重なりを調べる。また, 高校生が同じ質問を繰り返すことを調べるために, N-grams 分析を行い使用頻度から明らかにする。

RQ3 構築したコーパスに文法標識のタグを付与することによって, どのような教材作成が可能と

なるかについては, EasyConc for IzumiHS に文法タグを付与した EasyConc_tagged for IzumiHS.xlsm を活用し, ワークシートの教材作成例を示す。

4. 結果と考察

4.1 RQ1 英語で言いたかった表現を集めたコーパス EasyConc for IzumiHS にどのような特徴が見られるか。

EasyConc for IzumiHS のデータ収集後の英訳作業(stage 4)で気づいていたことがある。それは高校生ならではの興味関心がうかがえる質問が多く見られるということである。例えば, 調査票の中に「ドラえもん」に関する質問が多くみられたことである。以下に例を示す。EasyConc for IzumiHS で実際に'Doraemon'で検索した結果は以下のとおりである。

図 4 EasyConc for Izumi HS 0.7.xlsm (検索語:Doraemon 6 例ヒット)

38	ドラえもんは道具を出したり、非常にユニークだが、AIは道具を出さないのでドラえもんとAIは違う。	The cartoon character, Doraemon , takes out useful tools (or gadgets) from its pocket and this is unique, but AI doesn't do that and so it is different from Doraemon.
54	グルメテーブルかけ(ドラえもんの秘密道具)は食べたいものを思えば何でも出してくれる。	The Gourmet Table Cloth, one of Doraemon's gadget serves any food people want to eat.
55	どこでもドア(ドラえもんのひみつ道具)を使ってどこへ行きたいか。	By using Anywhere Door, one of Doraemon's gadgets, where would you want to go?
101	AIといえば、ドラえもんだとも思います。	When I think of AI, I think of Doraemon .
122	もしもボックスは私たちの願いを叶えてくれる。	Our wish would come true, if we had Doraemon's gadget "What-If Phone Box". / Doraemon's gadget "What-If Phone Box" could make our wish come true.
130	どこでもドアはAIを利用した道具なのかな。	I wonder if Doraemon's gadget "Anywhere Door" is a tool that uses AI.

高校生のみならず一般の人たちにとってもドラえもんはよく知られた国民的なアニメ・キャラクターであるが, COCA で'Doraemon'を検索すると以下の結果が得られる。

図 5 Corpus of Contemporary American English

検索語:Doraemon 10 例ヒット (2021 年 2 月 8 日アクセス)

Stranger Things theme # 3. In The End by Linkin Park # 4. Doraemon theme # 5. Noddy theme # 6. Old Town Road by Lil Nas
celebrate Astro Boy -- a wildly popular superhero robot -- and the robotic manga cat Doraemon. Many of those benevolent characters
Europeans are very fat fat, Europe is the fattest place on Earth. Doraemon: 4th Aug 2010 - 16:14 GMT # Amazing photos. # The one ma
reawakened psychic powers to save the world, sacrificing himself in the process. # Doraemon is fond of this. Many time the character
by the ascension - a cutesy gesture, perhaps - of a cartoon character, Doraemon, as the country's roving cultural ambassador. # The ar
understanding of Japan so they will become our friends. " For his part, Doraemon - who is as familiar to young people in Asia as Snoop
the walls be decorated with pictures of cartoon characters, such as, Pikachu and Doraemon, and paintings with relaxing, subdued colo
show videos and completed garments. Turning Japanese: A still from The Return of Doraemon (1998) and Chinatsu Ban's Elephant Unc
of the Lee Teng-hui presidency. A Japanese newspaper notes: # Hello Kitty and Doraemon Nihon junkies here are known by the specia
paintings. - What's that? - It's a seascape. What's Doraemon doing in a seascape? That's not a cat. It's a cloud

大規模コーパス COCA で'Doraemon'を検索しても、高校生が表現したいと思う「グルメテーブル」,「どこでもドア」,「もしもボックス」にたどり着くことはできないのである。

図 4 からわかるように EasyConc.xlsm には、高校生の興味関心のある事項が反映されているのに対して、図 5 の COCA では、高校生の興味関心の中身である「グルメテーブル」,「どこでもドア」,「もしもボックス」といった表現にまで触れられていない点である。Bowker & Pearson (2002: 48)が「コーパスサイズが、およそ 1 万語程度のものから数十万語のものであってもよく整備されたコーパスは LSP(language for special purposes)の研究に大変役立つのである」と述べているように、EasyConc for IzumiHS のような小さいコーパスであっても学習者の表現活動を支援するうえでその可能性が大きいと言えよう。

4.2 RQ2 コーパスに出現する語彙と高頻度語彙(2,000 語)と比較する。

EasyConc forIzumiHS は高校生が英語で表現できなかったことを日本語で書いてもらい、それを英語に翻訳する過程を踏んで構築されたコーパスである。英訳作業では筆者と筆者の同僚の英語母語話者とで高 1 までのシンプルな英語で翻訳することを心がけた。そこで使用される英語と Nation(2013)が分類した語彙分類との重なりについて調査する。重なりを調査するにあたり EasyConc for IzumiHS の英文を AntWordProfiler に取り込むことによって、表 1～表 3 ままで出力され、高頻度語彙の最初の 1000 語、次の 1000 語、アカデミック語彙の 570 語との比較が可能となる。比較するにあたり、EasyConc for IzumiHS の英語部分のみを抽出し、EasyConc for IzumiHS1.1_e.txt を作成し使用した。以下に EasyConc for IzumiHS1.1_e.txt で使用される英語と高頻度語彙(2,000 語)との比較を示す。

表 1 EasyConc for IzumiHS1.1_e.txt の異なり語数と総語数

Number of types	Number of tokens
705	2678

表 2 Level List Statistics

File name	Number of types	Number of groups (families)
1	4114	998
2	3708	988
3	3082	569

表 3 Lexical Profile Statistics

LEVEL	FILE	TOKEN	TOKEN%	CUMTOKEN%	TYPE	TYPE%	CUMTYPE%
1	1_gsl_1st_1000.txt	2192	81.85	81.85	451	63.97	63.97
2	2_gsl_2nd_1000.txt	136	5.08	86.93	94	13.33	77.3
3	3awl_570.txt	96	3.58	90.51	58	8.23	85.53
0	-	254	9.48	99.99	102	14.47	100
TOTAL		2678			705		

表 1 の EasyConc for IzumiHS1.1_e.txt というのは, EasyConc for IzumiHS の英語の異なり語数(types)が 705 語, 総語数(tokens)が 2,678 語であることを示す。

表 2 の Level List Statistics は, もっともよく使用される語が File name 1 (1_gsl_1st_1000.txt) のファイルに, ワード・ファミリーで 998 語あり, 次にもっともよく使用される語が, File name 2 (2_gsl_2nd_1000.txt)では, ワード・ファミリーで 988 語, そして, アカデミック語彙として File name 3 (3awl_570.txt)には, 569 語のワード・ファミリーが含まれていることを示す。

表 3 の Lexical Profile Statistics には, 今回使用した EasyConc for IzumiHS のサブ・コーパスである 泉 高 校 生 の EasyConc for IzumiHS1.1_e.txt と 3 種類 の コーパス (1_gsl_1st_1000.txt, 2_gsl_2nd_1000.txt, 3awl_570.txt)の語彙との重なりを示している。

表 3 の異なり語数(TYPE%)の重なりに着目すると, EasyConc for IzumiHS で使用される語彙は, もっともよく使用される語 (1_gsl_1st_1000)とは 63.97%の重なりが見られる。次にもっともよく使用される語 (2_gsl_2nd_1000)とは, 13.33%, アカデミック語彙 (3awl_570)とは 8.23%の重なりがあることが分かった。

表 3 の累積異なり語数(CUMTYPE%)の重なりに着目すると, EasyConc for IzumiHS で使用される語彙は, もっともよく使用される 1,000 語との重なりは 63.97%であり, 次にもっともよく使用される語の 1,000 語との重なりも含めると 77.3%の重なりを示す。アカデミック語彙まで含めると重なりは 85.53%にまで上がることが分かった。

学習者コーパス中にはドラえもんといった固有名詞も含まれており, 表 2 の File name 1,2,3 の語以外の語も含まれている。それらの語を除いても 85.53%の重なりはかなり高い数値を示してい

ると考えられる。この構築したコーパス中の 85.53%の語彙に習熟するだけでも学習者の表現能力を向上させる可能性を秘めているということである。

RQ 2 でのもう一つの課題として、高校生は同じ質問を繰り返すとすればどのような質問を繰り返すのかについて英語部分を対象にして N-grams 分析で出現頻度を調べる。

表 4 EasyConc for Izumi1.1_e.txt の N-grams 分析による比較

順位	頻度	bigram	頻度	trigram	頻度	four-gram	頻度	five-gram
1	32	in the	20	the edo period	14	in the edo period	11	people in the edo period
2	23	want to	14	i want to	11	people in the edo	3	i want to be a
3	20	edo period	14	in the edo	9	i would like to	3	i would like to be
4	20	the edo	11	people in the	5	the edo period i	3	in the edo period i
5	18	i want	10	i would like	4	do you want to	3	the life of people in
6	14	of ai	9	in the future	4	i want to be	3	to be a person who
7	14	to be	9	would like to	4	to be a person	2	ai is good but i
8	13	i think	7	you want to	3	be a person who	2	are you doing anything to
9	13	people in	6	i don t	3	i don t think	2	be a person who can
10	12	i am	6	we need to	3	i think ai is	2	come and see his game

表 4 から, the Edo Period, in the Edo Period, people in the Edo Period, want to, would like to, in the future, a person who, need to といった語句が繰り返し出現するのが分かる。表 4 の N-grams 分析の結果から、高校生は同じ質問を繰り返すとすればどのような質問を繰り返すのについて確かに学習者はチャンクとして同じ表現を繰り返し使用していることが読み取れるが、EasyConc for Izumi1.1.xlsm のサイズがまだ小さいこともあり、今後の継続研究が必要である。

4.3 RQ3 構築したコーパスに文法標識のタグを付与することによって、どのような教材作成が可能となるか。

EasyConc for IzumiHS.xlsm は、日本語の語句、あるいは英単語から検索することができるが、文法用語で検索することはできない。この EasyConc for IzumiHS.xlsm に、不定詞、受け身形といった文法用語から検索することができ、尚且つ検索結果には文法標識のタグがない形で表示できるソフトウェアの開発が必要である。そこで EasyConc for IzumiHS.xlsm の英語部分に CLAWS7 で文法標識を付与した。文法標識を付与するにあたっては、Free CLAWS web tagger を活用した。さらに、文法用語から検索する機能をつけ EasyConc_tagged_for IzumiHS1.2.1.xlsm を開発した。

使い方であるが、始めに EasyConc_tagged_for IzumiHS1.2.1.xlsm を起動する。続いて、

「Ctrl+K」キーで検索画面を起動する(図6)。受身形の検索では、「A21」をダブルクリック, 続いて「A22」をダブルクリックする。検索条件の Or, 表示形式のタグなしを選択する。最後に, OK をクリックする。文法標識のタグがない結果が表示される(図7)。

図6 EasyConc_tagged_for IzumiHS1.2.1.xlsm の文法用語での検索方法—「受身形」の検索

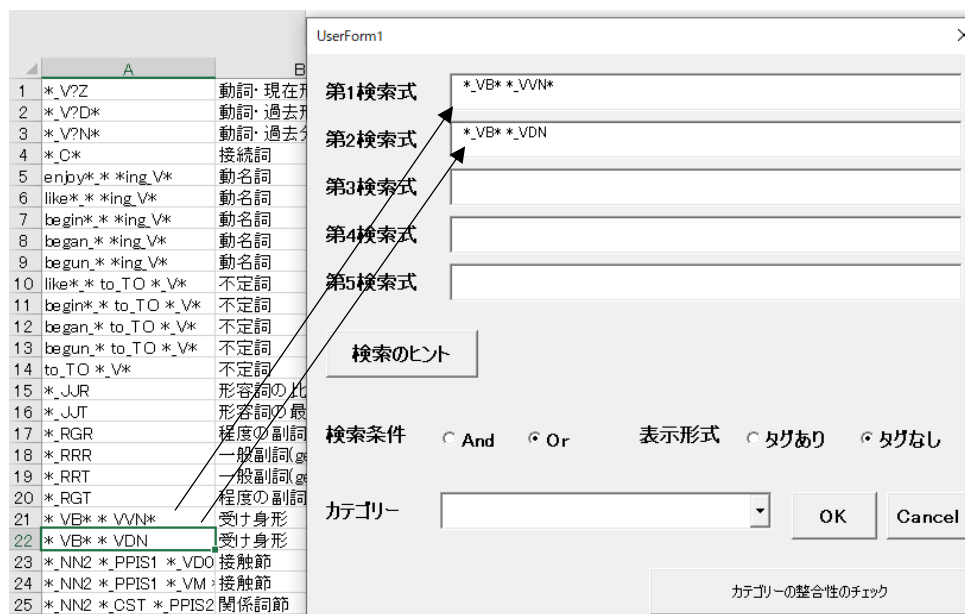


図7 EasyConc_tagged_for IzumiHS1.2.1.xlsm の文法用語の検索—「受身形」の検索結果

9	86	The contents of a shampoo bottle are refilled.	シャンプーボトルはそのまま中身を詰め替える。	12. 政治・社会問題
10	104	I wonder what can be created in the future.	将来どんなものが作り出されるか？	12. 政治・社会問題
11	151	People are trained to deal with the challenges of the new era.	新しい時代に対応できるようにする。	12. 政治・社会問題
12	157	The official mascot of Tessei was designed by an employee, who wanted to be an illustrator, with his colleagues' cooperation. It is used on social media by Tokyo station.	イラストレーター志望の人は皆の協力によりテッセイのイメージキャラクターをデザインしました。それは東京駅のSNSなどでも使われています。	12. 政治・社会問題
	176	I am reborn. / I am different now. / I am reborn as a new person.	新しい自分に生まれ変わる。	12. 政治・社会問題

(「受身形」の検索結果一部抜粋)

文法用語での検索ができることにより, 学習者がどのようなことを英語で表現したかったのか, そしてその英語表現の中で, 学習者がどのような文法的な知識があれば英語で表現することができるのかという情報を教師は得ることができる。例えば, 英語表現の授業では, 教科書で扱う文法事

項の補足をコーパスの中から生徒の目線で検索し活用することができるのである。コーパスの活用としてワークシートの教材作成例を示す。

筆者は、英語に苦手意識を持っていると思われる生徒にとって、不定詞を使いこなすことは容易ではない文法事項の一つと考えている。これについての実証的なデータの裏付けは今後の課題としたい。以下は、EasyConc_tagged_for IzumiHS1.2.1.xlsm を活用し、不定詞で検索し作成したワークシートの例である。

【不定詞の整序問題】

1. 最近の人は新しいものを買いすぎている。

Recently people tend (too / many / buy / new / to / things).

2. 環境のために何かやっていることはありますか。

Are you doing (to / the / environment / help / anything)?

3. 一人一人がごみを出さないように心がける。

Individual people need to try (to / too / much / garbage / not / produce).

4. このままだとごみが増え、私たちは生活しにくくなってしまいうだろう。

If we don't do anything, the amount of garbage will increase and (us / difficult / it / be / will / for / to / live).

他の文法用語についても EasyConc_tagged_for IzumiHS1.2.1.xlsm コーパスを活用し、文法事項の観点からワークシートを作成し、生徒からの質問をフィードバックすることができる。留意したい点として、使用する英文は学習者が表現しなかった内容という点で教師が文法の練習問題のために創作した英文とは異なるのである。

日臺(2003)では、英語を得意としていると思われる生徒は苦手意識を持っていると思われる生徒と比べ、接続詞の種類と使用頻度が高いこと、また日臺(2007)では、学年が進行するにつれて接続詞の種類が増えていくことを示している。筆者は、学習者が接続詞を使えるようになることで表現力を伸ばすことにつながると考えている。接続詞に着目し、コーパスから作成したワークシートの例を示す。

【接続詞の穴埋め補充問題】

Gap fill exercises: Fill in each bracket with a suitable linking word in the box.

1. () I have a lot of clothes, I can't stop buying more.
2. () I think of AI, I think of Doraemon.
3. () I write a little, I tend to throw the paper away immediately.
4. () we don't do anything, the amount of garbage will increase and it will be difficult for us to live.

when, if, although, after

【接続詞で文をつなげる問題】

Matching Exercises: Use a linking word in Box B to match the sentence parts in boxes A and C.

Box A	Box B	Box C
1. At first, I was afraid of AI, 2. Even though robbers come to the bank, AI can't run to catch them 3. The Edo Period was a good era	but because nor	it was peaceful and people were lively. now I think it will change our life for the better. deal with them.

また、学習者が仮定法の内容を理解するために、仮定法を文構造として学習する事に加えて、英語で言えなかった表現をフィードバックしたワークシートに取り組むことで、学習者の目線で文法概念を深めていくことができるのではないかと考えている。

【仮定法の構文理解のための整序問題】

1. If AI were used as a weapon, (be / it / dangerous / would) because it would hurt people.
2. If you could go back to the Edo Period and work there, (job / would / want / what / to you / do)?
3. Our wish would come true, (if / Doraemon's gadget / we / "What-If Phone Box" / had).

日臺(2003)、投野(2004)では、主語の位置に現れる後置修飾の関係代名詞(例えば、a person who …のような文構造)は難易度が高いことを示している。しかし、その場合に限らず学習者にとって後置修飾そのものが日本語構文にないだけに容易ではないと筆者は考えている。最後に、後置修飾の例として関係代名詞に着目し、コーパスを活用して作成したワークシートの例を示す。

【関係代名詞の先行詞と形容詞節の構文理解の整序問題】

1. I want (that / robots / people / take care of).
2. This is a story in (AI / humanity / destroys / which).
3. We should buy (can / that / use / we / for a long time / things).
4. The Goumet Table Cloth, one of Doraemon's gadget serves (any / eat / people / to / food

/ want).

本研究で構築したコーパスは学習者が英語で言いたかった表現が豊富なコーパスであり、そのコーパスから授業で補強する文法を含む表現を抽出し、授業で学習者にフィードバックする手法を授業に取り入れることにより生徒の表現力の向上に役立つことを示した。

現在, EasyConc_tagged_for IzumiHS1.2.1.xlsm は 2.1 までアップグレードし、下記の文法用語から検索することが可能となっている。

【検索可能な文法用語】

動詞・現在形・主語が 3 人称単数で現在るとき (is, does, has, -s form of lexical verb (e.g. gives, works)), 動詞・過去形 (were, was, did, had, past tense of lexical verb (e.g. gave, worked)), 動詞・過去分詞形 (been, done, had (past participle), past participle of lexical verb (e.g. given, worked), past participle catenative (e.g. bound in be bound to)), 接続詞, 動名詞, 不定詞, 形容詞の比較級, 形容詞の最上級, 程度の副詞の比較級 (more, less), 一般副詞 (general adverb) の比較級 (e.g. better, longer), 一般副詞 (general adverb) の最上級 (best, longest), 程度の副詞の最上級 (most, least), 受け身形, 受け身形, 接触節, 関係詞節

なお、関係詞節については、検索式の構文が非常に複雑である。図 6 の EasyConc_tagged_for IzumiHS1.2.1.xlsm は、27 の関係詞節の検索式を有しているが、さらにデータ数が増えると検索式が追加される可能性も予測されるだけに、身近な Excel を使用した検索方法に限界もあり、今後さらに研究する余地が残されている。

5. まとめ

本研究では英語のスピーキングテストの後に、振り返りの活動として、英語で言いたかった表現を日本語で集め英語に直し、ワークシートとして授業でフィードバックする一連の流れの中で、いかに英語で言いたかった表現を集めたコーパスが学習者の表現能力を伸ばすうえで可能性に富んでいるかについて研究したものである。

RQ1 では、構築した日英パラレル・コーパスから学習者ならではの興味関心を示す表現が見られること、そしてそれは大規模コーパス COCA には見られないことを示した。

RQ2 では、構築したコーパスに出現する語彙が意表を突くような話題の語彙ではなく、コーパスに出現する語彙の多くがもっともよく使用される語彙とかなり高い重なりが見られることを示した。また、学習者の質問の中に、チャンクとして同じ表現が繰り返し使用されていることも示した。

RQ3 では、文法標識のないコーパスでは日本語、英語からの検索ができるが、文法標識を付与したコーパスでは文法用語から検索することができることを示した。一方、関係詞節を検索式で表現することの難しさにも触れた。また、文法標識の有無にかかわらずコーパスからワークシートを作成することが容易であることを示した。

Nation (2008: 97) で指摘しているように、「語彙を教える際に効果が高いのは学習者が、語彙を必要だと感じるときに支援することである」とはその通りで、特にコミュニケーション活動のようにメ

ッセージに焦点をあてた言語活動において生徒が英語で言いたい表現が出てこないときに、その表現を教師が支援できることが極めて大切である。さらに一步進めて、学習者自らがそのようなツールを端末として持ち、言語各様において必要に応じて使用することができれば大変よいことである。現在、iOS 仕様の端末で利用できるソフトウェア EasyConc.fmp12 を開発し筆者のホームページで公開している。高校生が英語で言いたかった表現を集めたコーパスの開発とそれを用いた教材開発が一層望まれるのである。コーパスの一層の拡大がこれからの課題といえる。

謝辞

本研究で利用した高校 2 年生の英語で言いたかった表現の日本語データは、宮城県泉高校 2 年生、そして宮城県泉高校英語科教諭渡邊崇先生にご協力いただいた。

コーパスの英訳作業では、Steve Lia 氏(玉川大学)、また筆者の着想をもとに Windows 仕様の EasyConc for IzumiHS.xlsm の構築では大村あつし氏(オフィス不死鳥)、携帯可能な iOS 仕様の端末で利用可能な EasyConc for iPhone of Izumi.fmp12, EasyConc for iPad of IzumiHS.fmp12 の開発では内藤清志氏(玉川大学)のご協力をいただいた。御礼申し上げたい。

本研究は、科学研究費の支援を受けている。科学研究費基盤研究(C) 研究課題番号: 25370652, 研究代表者: 日臺滋之

引用文献

Anthony, L. (2021) AntPConc. [software]

<https://www.laurenceanthony.net/software/antpconc/>

Barlow, M. (2021) ParaConc.[software] <https://paraconc.com/>

Lynne, B., & Pearson, J. (2002). *Working with specialized language a practical guide to using corpora*. Routledge.

日臺滋之(2003)「4 英語の得意な生徒と苦手意識を持った生徒」『英語力はどうのように伸びてゆくか—中学生の英語習得過程を追う』(大修館書店), 125-131.

日臺滋之(2007)「接続詞の発達」『日本人中高生一万人の英語コーパス “JEFLC Corpus” — 中高生が書く英文の実態とその分析—』(小学館), 88-94.

日臺滋之(2013)「語彙指導で使う教材をオリジナルで作る:「英語で言いたかったけれど言えなかった表現・語句」を探り、授業に活かす」『英語教育』(大修館書店) 62(2), 20-22.

日臺滋之(2015)「日本人高大生の自己表現活動を支える日英パラレルコーパスの構築とその活用」. 科学研究費基盤研究(C) 研究課題番号:25370652, 研究代表者: 日臺滋之

<https://kaken.nii.ac.jp/d/r/60459302.ja.html>

日臺滋之(2017)「日英パラレルコーパスで補完する教科書の語彙指導」『英語教育』(大修館書店) 65(12), 18-19.

日臺滋之・Lia, Steve・大村あつし(2021) EasyConc for IzumiHS.xlsm, EasyConc_tagged for IzumiHS.xlsm [software]

<http://www.tamagawa.ac.jp/research/je-parallel/izumishs/>

日臺滋之・Lia, S.・内藤清志(2021) EasyConc for iPhone of IzumiHS.fmp12, EasyConc for iPad of IzumiHS.fmp12 [software] <http://www.tamagawa.ac.jp/research/je-parallel/izumishs/>

文部科学省(2019)『高等学校学習指導要領(平成 30 年告示)解説 外国語編 英語編』開隆堂.

Nation, I.S.P. (2008). *Teaching vocabulary: strategies and techniques*. Heinle.

Nation, I.S.P. (2013). *What should every EFL teacher know?* Compass Publishing.

Rayson, P. (2021). Free CLAWS web tagger.

<http://ucrel-api.lancaster.ac.uk/claws/free.html>. Lancaster University.

投野由紀夫編著 (2007)『日本人中高生一万人の英語コーパス “JEFLC Corpus” 一中高生が書く英文の実態とその分析』(小学館) 6-12.

投野由紀夫(2004)「The NICT JLE Corpus に見る英語学習者の発表語彙の使用状況」『日本人 1200 人の英語スピーキングコーパス』(アルク) 106-108.

内山将夫・高橋真弓(2003)「日英対訳文対応付けデータ」

<https://www2.nict.go.jp/astrec-att/member/mutiyama/align/download/index.html>