

〈論 文〉

英語のコミュニケーション活動で 日本人中学生が必要としている定型表現 — 検定教科書と日英パラレルコーパス *EasyConc* とを比較して

日 臺 滋 之

要 旨

本研究は、中学校検定教科書 *New Crown English Series 1, 2, 3* (以下, *03NC*) と、日英パラレルコーパス *EasyConc* とから抽出した定型表現とを比較し、英語のコミュニケーション活動で日本人中学生が必要としている定型表現を調査することである。

最初に *The first 100 collocations* (Shin & Nation, 2008) と *03NC* をもとに定型表現リスト *teikei5.txt* を作成する。*teikei5.txt* は見出し語とその変化形から構成されている。次に *teikei5_LR.txt* を作成する。これは *teikei5.txt* を用いて検索した結果を見出し語に戻す際に必要となる置換リストである。

続いて、*AntConc* で *teikei5.txt* を用いて *03NC* と *EasyConc* からそれぞれ定型表現を抽出する。次に秀丸エディタで置換リスト *teikei5_LR.txt* を用いて定型表現の各動詞等を見出し語に置換する。このようにして *03NC* と *EasyConc* の定型表現のリストを作成する。

この二つのリストを比較し、以下のことが分かった。(a) ピアソンの積率相関係数を算出したところ両者に有意な正の相関関係が認められた ($r = 0.475, p < 0.001$)。 (b) 定型表現とその頻度の分布グラフから少数の使用頻度の高い定型表現と多数の使用頻度の低い定型表現とが分布しているという類似点が両者に見られる。(c) 英語教育への示唆として、*03NC* と *EasyConc* の両方に出現する定型表現を意図的に指導することが望まれる。

キーワード：定型表現, 日英パラレルコーパス, *EasyConc*, 学習者のニーズ

1. はじめに

中学校の英語の授業でコミュニケーション活動として、週末の出来事についてクラスメートと対話する活動を行い、授業後に活動の振り返りを行うと、生徒の中には、振り返りシートに「自由が丘に買い物に行った、を英語でどう言いますか?」と書いてくる生徒がいる。

多くの生徒は、「買い物」は shopping で、「行った」は went と、単語は言えるが、「買い物に行った」という表現になると戸惑い会話が中断することがある。生徒によっては、「go shopping in 場所」というパターンが口をついて出てこない状況が見られる。教師は、学習者が必要とする表現のニーズを知り、表現に対応するパターンを教室で指導する必要がある。学習者は、学習者のニーズとしてこのような定型表現のパターンを知り、うまく使いこなすことで表現力が伸びていくものと予想される。これは、Martinez & Schmitt (2012) らが、定型表現 (formulaic language) は、第二言語学習者の言語生成の全体的な印象を向上させる可能性があることと述べていることと合致する。本稿では、検定教科書で提示される定型表現と比較しつつ、コミュニケーション活動における学習者のニーズとしてどのような定型表現を必要としているかについて論じる。

2. 先行研究

2.1 では、まず定型表現という用語をめぐってどのようなことが言われているか共通理解をしておきたい。2.2 では、精選された定型表現リストとして、The first 100 collocations (Shin & Nation, 2008) を採り上げる。2.3 では、学習者のニーズを捉えた先行研究として、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 *EasyConc* とその検索アプリ

EasyConc5.8.xlsm について扱う。

2.1 定型表現という用語について

『中学校学習指導要領解説外国語編（2017）』では、「(イ) 連語のうち、活用頻度の高いもの」を扱うとしており、「連語」とは、in front of, a lot of, look for などのように、二つ以上の語が結び付いて、あるまとまった意味を表すものを指している」と定義している。また、「小学校の外国語科においては、get up, look at などの活用頻度の高い基本的な連語を指導している」としている。同書において、「(ウ) 慣用表現のうち、活用頻度の高いもの」を扱うとしており、「慣用表現」とは、ある特定の場面で用いる定型表現を指している」と定義している。さらに小学校の外国語科で例示されている慣用表現として、excuse me, I see, I'm sorry, thank you, you're welcome, 中学校においては、first of all, on the other hand などを例示している。

学習指導要領の、連語、慣用表現、そして定型表現といった用語に代わるものとして様々な用語が使用されている。例えば、McCarthy (1990) は、multi-word units を使用し、Wray (2002) では、chunks, collocations, gambits, idioms, formulaic language, multiword units 等々、formulaicity (定型性) を説明するために 57 例もの用語を挙げているが、Wray (2002) は、formulaic sequence という用語を使用している。Schmitt & Schmitt (2020) らは、すべてのタイプの multi-word unit の包括的用語として formulaic language を使用し、個々のものを formulaic sequence として用いている。Martinez & Schmitt (2012) らは phrasal expressions という用語を使用している。さらに、Nation (2013) は、multiword units を使用し、Webb & Nation (2017) らは、'How are you?', 'of course' など一つのユニットとして機能する連続した単語の包括的用語として、multi-word combinations を用いており、用語が統一されているわけでは

ない。

本研究では、用語の分類の違いを目的としているわけではなく、Martinez & Schmitt (2012) が述べているように、「言語が個々の単語から構成されているだけでなく、かなりの量の定型表現 (formulaic language) で構成されている」という事実に目を向け、学習者が、英語の表現力を伸ばすうえで一体どのような定型表現 (formulaic language) を必要としているのかについて調査する。

2.2 精選された定型表現リストについて

— The first 100 collocations (Shin & Nation, 2008) について

Shin & Nation (2008) らは、コロケーションとは、頻繁に一緒に出現する2語あるいはそれ以上のグループを指し、中心となる語 (pivot word) とそれに伴う語句から構成されているとしている。そしてコロケーションは話し言葉では重要な役割を果たすことから British National Corpus (BNC) の話し言葉の1千万語を対象に、6つの基準を設定し、コロケーションを抽出した。6つの基準とは以下のとおりである。

- ・コロケーションの中心となる語 (pivot word) は異なり語であること。例えば、sigh は give よりも過去形の gave と典型的に共起する。
- ・コロケーションの中心となる語 (pivot word) は名詞、動詞、形容詞、あるいは副詞であること。
- ・コロケーションの中心となる語 (pivot word) は最も頻度の高い1,000語の内容語であること。
- ・コロケーションは1千万語中に少なくとも30回以上出現すること。
- ・文中で各コロケーションが直近のコロケーションの境界を越えるべきではないということ。例えば、I saw you at that place. の文では、

I saw you at that place, saw you at that place, saw you, at that place, that place と 5 つの直近のコロケーションの構成要素があるが、you at that place は直近の構成要素の境界を越えているので除外される。

- ・形は同じでも意味の異なるコロケーションは、別に数える。例えば、上向くという意味の look up と調べるという意味の look up は別に数える。

Shin & Nation (2008) による The first 100 collocations の最初の 100 語のうち上位 10 語のみ示すと以下ようになる。

表 1 The first 100 collocations の上位 10 番目までのコロケーション

RK	Collocations	FRE
1	<i>you know</i>	27348
2	<i>i think (that)</i>	25862
3	<i>a bit</i>	7766
4	<i>(always [155], never [87]) used to {INF}</i>	7663
5	<i>as well</i>	5754
6	<i>a lot of {N}</i>	5750
7	<i>{No.} pounds</i>	5598
8	<i>thank you</i>	4789
9	<i>{No.} years</i>	4237
10	<i>in fact</i>	3009

(以下、省略)

Shin & Nation (2008) は、コーパスの出所は話し言葉で、主として大人のイギリス人で、口語体の話し言葉とかなり改まった話し言葉が入り混じっていること、リスト中には、7 *{No.} pounds* といったイギリス英語の特徴の強い表現も含まれていることに留意するように述べている。

本研究の被験者は大人ではなく中学生であり、大人とは話題も異なる。したがって The first 100 collocations のリストだけでは不十分で中学生を対象とする令和 3 年度の中学校検定教科書 *New Crown English Series*

Book1, *Book 2*, *Book 3* (以下, 03NC) で使用されている連語, 慣用表現も加えて定型表現リスト (teikei5.txt) を構築する必要がある。

2.3 学習者のニーズを捉えたコーパスの活用について

— 英語で言いたかった表現や語句を収集し構築した日英パラレルコーパス *EasyConc* とその検索アプリ *EasyConc.xlsm*

中高大学生のコミュニケーション活動では, 対話の途中, 英語で表現できないために言葉に詰まり対話が中断され継続できない問題が起こる。この状況に着目し, コミュニケーション活動の後に, 活動の振り返りとして, 学習者が英語で表現したかったことを日本語で書いてもらう授業実践を積み重ねてきた。日臺, 佐藤, 茅野 (2021) では中高大学での授業実践が示されているが, ここでは中学校の授業実践例を紹介する。

中学校の実践例で, 中学2年生, 3クラス107名を対象に, トピックとして好きな教科, 家族とすること, 好きなスポーツ, 好きな動物, 好きな果物, 好きなテレビ番組, 好きな季節を設定し, 年4回 chat を実践した。活動後に, 英語で言いたかった表現を日本語で書いてもらい回収した。

これは, 本研究のデータが中学校の授業でどのように収集されたかを示す実例である。コミュニケーション活動の後には, 学習者が英語で言いたかった表現を英語母語話者教師と日本人英語教師とが英訳し, Excel 上で日本語と英語を一对一に対応させたデータベースを構築する。高校, 大学での授業実践のデータも合わせ, このデータベースを日英パラレルコーパス *EasyConc* と言い, コミュニケーション活動で日本人中・高・大学生が英語で表現したかった表現語句を知ることができる。*EasyConc* は学習者のニーズが反映されているコーパスである。この *EasyConc* を各種 OS 上で作成した検索プログラムに組み込み, アプリケーションソフトとして, Windows 用の *EasyConc.xlsm*, *EasyConc for Access.accdb*, iOS 用の *EasyConc for iPhone/iPad.fmp12*, Chromebook OS 用の *EasyConc for*

Chromebook を開発した（北村，内藤，Lia，日墓，2021）。

Windows 仕様の検索アプリ *EasyConc5.8.xlsm*⁽¹⁾ で使用した *EasyConc* の中・高・大学生からの日本語の質問件数，また，英単語の異なり語数，延べ語数を表 2 に示す。本研究は，中学生が必要としている定型表現に焦点を絞っているので中学生のデータを対象に扱う。

表 2 *EasyConc5.8.xlsm* の校種別の英語で表現したかったことの質問件数と英単語数

	中・高・大学生から 寄せられた日本語での 質問件数	教師による英訳の 英単語の異なり語数 (Word Types)	教師による英訳の 英単語の延べ語数 (Word Tokens)
中学	3,943	3,548	37,004
高校	467	1,166	5,295
大学	1,227	2,086	11,295
合計	5,637	4,646	53,594

3. リサーチデザインと手法

3.1 研究の目的

以上を踏まえ，下記の 3 項目を研究の目的として設定した。

- (1) The first 100 collocations のリストとインプットとしての中学校の検定教科書に現れる連語，慣用表現とから定型表現リスト (teikei5.txt) を作成する。
- (2) 定型表現リスト (teikei5.txt) を用いて，インプットとしての中学校の検定教科書に現れる定型表現とその頻度を調査する。また，定型表現リスト (teikei5.txt) を用いて，コミュニケーション活動で学習者が必要としている *EasyConc* の定型表現とその頻

度を調査する。

- (3) インプットとしての中学校の検定教科書に現れる定型表現とコミュニケーション活動で学習者が必要としている *EasyConc* の定型表現との間にどのような関係があるか調査する。

3.2 研究方法

研究の目的 (1) について、本研究で使用する定型表現リスト (*teikei5.txt*) の作成過程を示す (3.2.1)。

研究の目的 (2) について、研究の目的 (1) で作成した定型表現リスト (*teikei5.txt*) を用いて、インプットとしての中学校の検定教科書に現れる定型表現とその頻度の抽出過程を示す (3.2.2)。また、(1) の定型表現リスト (*teikei5.txt*) を用いて、コミュニケーション活動で学習者が必要としている *EasyConc* の定型表現とその頻度の抽出過程を示す (3.2.3)。

研究の目的 (3) について、インプットとしての中学校の検定教科書に現れる定型表現とコミュニケーション活動で学習者が必要としている *EasyConc* の定型表現との間にどのような関係があるか数量的な分析を行う (3.2.4)。

3.2.1 定型表現リスト (*teikei5.txt*) の作成過程について

研究の目的 (1) について、まず定型表現リストの作成過程について述べる。Shin & Nation (2008) の *The first 100 collocations* のリストはイギリス人の大人の英語母語話者の話し言葉に見られる使用頻度の高い上位 100 語の定型表現リストである。精選された定型表現であっても外国語として英語を学習している日本人中学生のコミュニケーション活動で使用する定型表現とは大きな開きがあることが予想される。そのために、検定教科書 *03NC* の各学年の本文や付録「単語の意味」に掲載されている連語

を抽出し、The first 100 collocations のリストと合体し、本研究で使用する定型表現リスト（teikei5.txt（保存形式 UTF-8））を作成する。

定型表現リスト（teikei5.txt）を用いて、AntConc の Advanced Search を利用し、検定教科書 *O3NC* から一気に定型表現を検索し、頻度を調査する。そのために、あらかじめ定型表現リストの定型表現を正規表現で書き直しておく必要がある。秀丸エディタを使い、行頭、行末に「\b」（語句の境界を示す記号）を入力した。秀丸エディタでの作業手順は以下のとおりである。

「検索」—「置換」を選択。続いて「正規表現」に✓を入れる。

行頭、行末に「\b」を入力。**は円マークを使用。**

検索：^

置換：\\b

行末に「\b」を入力。**は円マークを使用。**

検索：\$

置換：\\b

最後に、定型表現 teikei5.txt という名称で保存。正規表現で記載した定型表現の総数は、654 例となった。

```
\b(act|acts|acting|acted) out\b
\b(believe|believes|believed) in\b
\b(belong|belongs|belonging|belonged) to\b
\b(bend|bends|bending|bent) over\b
\b(blew|blow|blows|blowing) [a-z]+ up\b
```

（以下省略）

図1 定型表現 teikei5.txt のファイルに納められている文字列

The first 100 collocations に納められているコロケーションの中心となる語 (pivot word) は異なり語で表示されており, 03NC の連語は, 見出し語で記載されているので, teikei5.txt では, 見出し語とその活用形リストを作成した (図1を参照されたい)。

3.2.2 インプットとしての検定教科書に含まれる定型表現とその頻度を抽出する

研究目的 (2) について, 最初に, 定型表現 (teikei5.txt) を用いて, インプットとしての中学校の検定教科書に現れる定型表現とその頻度を明らかにする。本研究では, 既存のアプリを活用し, 大量のデータを対象に半自動化で定型表現を抽出し, 目視によって確認する手法を用いている。

一連の作業で必要となるアプリやファイルを以下に列挙する。

AntConc3.5.9, 秀丸エディタとそのマクロファイル ListReplace.mac⁽²⁾, 定型表現ファイル teikei5.txt, teikei5_LR.txt (秀丸エディタのマクロファイル ListReplace.mac を動作するときに必要な定型表現の置換ファイル)

上記のアプリ等を用いて, 検定教科書 03NC に出現する定型表現の文字列を抽出するための一連の作業は以下の (1) から (6) の手順に従う。

(1) 対象となるテキストの選択と整形

検定教科書 03NC を使用する。学年ごとに, 03nc1.txt, 03nc2.txt, 03nc3.txt (保存形式 UTF-8) とファイル名を付ける。

(2) 定型表現を含む英文の抽出

AntConc3.5.9 で, 03nc1.txt, 03nc2.txt, 03nc3.txt を開き, Concordance 機能で teikei5.txt を用いて, Advanced Search で定型表現を検索。

- ・ AntConc3.5.9 の Advanced Search では, Regex と Use search term(s) from list below の両方に ✓ を入れる。Load File をクリックし, teikei5.txt を選択。出力結果を 03nc_teikei5.txt (保存形式 UTF-8)

で保存。

(3) 定型表現の直前までの文字列の削除

03nc_teikei5.txt を Excel で開き、定型表現の直前までの文字列とファイル名 (03nc1.txt, 03nc2.txt, 03nc3.txt) とを削除する。定型表現とその直後に続く文字列を残し、03nc_teikei5_half.txt で保存。

なお、03nc_teikei5_half.txt は Shift-JIS で保存されているので、秀丸エディタで開き、UTF-8 で保存し直し、以下のテキスト整形をする。

- ・英文の前後に「”」が入ってしまう場合は、秀丸エディタで「”」を削除。検索一置換、「単語の検索」と「正規表現」に✓。

検索：”

置換：空白のまま

続いて、「全置換」―「一気」を選択。

- ・行頭に空白がないかチェックする。行頭に空白があれば、検索一置換で行頭の空白を削除。「単語の検索」と「正規表現」に✓。

検索：^[\t]+ ←[と \tの間には半角スペースが1つ入っている。

置換：←何も記入しない。

続いて、「全置換」―「一気」を選択。

- ・半角スペースが2つ続く箇所があれば、検索一置換で1つにする。

検索：←半角スペースを2つ入れる。

置換：←半角スペースを1つ入れる。

続いて、「全置換」―「一気」を選択。

(4) 定型表現の直後に記号 (>) を挿入

次に定型表現の文字列の直後に「>」の記号を秀丸エディタの置換機能で挿入する。

- ・03nc_teikei5_half.txt を秀丸で開く。秀丸の「マクロ」―「マクロ実行」で ListReplace.mac を選択。「変換リストによる連続置換」が表示されたら、「変換リスト」の「リスト選択」で teikei5_LR.txt（置

換りリスト) を指定して置換実行をクリックする。定型表現の直後に「>」の記号が挿入される。03nc_teikei5_LR.txt (保存形式 UTF-8) で保存。

- ・置換リストの teikei5_LR.txt には、以下の文字列が収められている。全部で 654 行の文字列で構成されている。

```
(act|acts|acting|acted) out,act out>,regular
(believe|believes|believed) in,believe in>,regular
(belong|belongs|belonging|belonged) to,belong to>,regular
(bend|bends|bending|bent) over,bend over>,regular
(blow|blew|blown|blows|blowing) [a-z]+ up,blow ... up>,regular
(以下省略)
```

図 2 置換リスト teikei5_LR.txt のファイルに納められている文字列

(5) 定型表現の直後の文字列の削除

Excel で 03nc_teikei5_LR.txt を開く。「区切り文字」で「その他」に「>」を入力。続いて、Excel 上で「>」の記号の直後の文字列を削除。残った文字列を目視で確認することにより精度が上がる。最後に、03nc_teikei5_LR.xlsx で保存。

(6) 定型表現の集計

03nc_teikei5_LR.xlsx の残った定型表現の文字列を Excel のピボットテーブルで集計する。

3.2.3 コミュニケーション活動で学習者が必要としている *EasyConc* の定型表現とその頻度を抽出する

研究目的 (2) について、定型表現 (teikei5.txt) を用いて、コミュニケーション活動で学習者が必要とする *EasyConc* の定型表現を抽出するために、学習者が英語で表現したかった表現や語句から構築した日英パラレル

コーパス *EasyConc* を使用する。学習者が表現したいと思う定型表現を知るには学習者が表現できたことを収集したコーパスではなく、英語で表現したかった表現や語句を集めた *EasyConc* が必要である。*EasyConc* の定型表現とその頻度を抽出する手順は以下のとおりである。

(1) 対象となるテキストの選択と整形

EasyConc の検索アプリ *EasyConc5.8.xlsm* の中学生の日本語による質問の英訳箇所をコピー & ペーストで Excel に張り付ける。これに *EasyConc.txt* とファイル名を付けて、「テキスト (タブ区切り) (*.txt)」で保存。続いて、*EasyConc.txt* をメモ帳あるいは秀丸エディタで開いて、UTF-8 で保存。

- ・英文の前後に「”」が入る場合、行頭に空白がある場合、半角スペースが2つ続く場合の対応については、3.2.2 の (3) を参照されたい。

(2) 定型表現を含む英文の抽出

AntConc3.5.9 で、*EasyConc.txt* を開き、Concordance 機能で *teikei5.txt* を用いて、Advanced Search で定型表現を検索。

- ・操作方法は、3.2.2 の (2) を参照されたい。出力結果を *EasyConc_teikei5.txt* で保存。

(3) 定型表現の直前までの文字列の削除

EasyConc_teikei5.txt を Excel で開き、定型表現の直前までの文字列とファイル名 (*EasyConc.txt*) とを削除する。定型表現とその直後に続く文字列を残し、*EasyConc_teikei5_half.txt* で保存。

なお、*EasyConc_teikei5_half.txt* は Shift-JIS で保存されているので、秀丸エディタで開き、UTF-8 で保存し直し、必要に応じて以下のテキスト整形をする。

- ・英文の前後に「”」が入る場合、行頭に空白がある場合、半角スペースが2つ続く場合の対応については、3.2.2 の (3) を参照されたい。

(4) 定型表現の直後に記号 (>) を挿入

次に定型表現の文字列の直後に「>」の記号を秀丸エディタの置換機能で挿入する。

- ・ EasyConc_teikei5_half.txt を秀丸で開く。秀丸の「マクロ」—「マクロ実行」で ListReplace.mac を選択。「変換リストによる連続置換」が表示されたら、「変換リスト」の「リスト選択」で teikei5_LR.txt（置換リスト）を指定して置換実行をクリックする。定型表現の直後に「>」の記号が挿入される。EasyConc_teikei5_LR.txt（保存形式 UTF-8）で保存。

(5) 定型表現の直後の文字列の削除

Excel で EasyConc_teikei5_LR.txt を開く。「区切り文字」で「その他」に「>」を入力。続いて、Excel 上で「>」の記号の直後の文字列を削除。残った文字列を目視で確認することにより精度が上がる。最後に、EasyConc_teikei5_LR.xlsx で保存。

(6) 定型表現の集計

EasyConc_teikei5_LR.xlsx の残った定型表現の文字列を Excel のピボットテーブルで集計する。

3.2.4 研究の目的「(3) インプットとしての中学校の検定教科書に現れる定型表現とコミュニケーション活動で学習者が必要としている *EasyConc* の定型表現との間にどのような関係があるか調査する」について

検定教科書では、定型表現のファイル (3.2.2(6)の 03nc_teikei5_LR.xlsx) を用い、*EasyConc* では、その定型表現のファイル (3.2.3 の(6)の EasyConc_teikei5_LR.xlsx) を用いて、以下 3 つの観点から分析する。

- ・ 検定教科書はインプットの側面を持ち、学習者が英語で言いたかった表現はアウトプットの側面を持つので、双方の定型表現リストとその

頻度の関係について Pearson の積率相関係数を算出する (4.3.1)。

- ・ Zipf's law において、少数の頻度の高い語と多数の頻度の低い語があるという現象が指摘されているが、定型表現についても同様なことが言えるか、双方の定型表現とその頻度をもとにグラフを作成し、視覚化することによって定型表現とその頻度についての特徴を調査する (4.3.2)。
- ・ 双方に共通して出現する定型表現から英語教育に示唆する視点を見いだせるか検討する (4.3.3)。

4. 結果と考察

4.1 研究の目的 (1) The first 100 collocations のリストとインプットとしての中学校の検定教科書に現れる連語、慣用表現とから定型表現リスト (teikei5.txt) を作成する。

- ・ Shin & Nation (2008) の The first 100 collocations と中学校の検定教科書 *O3NC* に出現する連語、慣用表現とを合体し、図 1 で示した定型表現リストの teikei5.txt を、作成することができた (3.2.1)。中学生を対象にした検定教科書も加えて作成したことで、BNC の成人の話し言葉から抽出した The first 100 collocations を補足できたと考えている。この teikei5.txt に収められた 654 の定型表現で学習者が言いたかったことをすべてカバーできているとは言い切れないが、コミュニケーション活動で学習者が必要とする定型表現を知る第一歩と考えている。
- ・ 置換作業で使用した置換リスト teikei5_LR.txt は、go skiing, go shopping はすべて go -ing として一括りにする設定となっているので個々の実例が見えにくいことに留意したい。

なお、表中の調整頻度については、双方のコーパスサイズが異なるため、ともに1万語中に出現する定型表現に調整するため、検定教科書 *03NC* では、go to の頻度 $42 \div 20,880$ (*03NC* の延べ語数) $\times 10,000$ で示した。また、*EasyConc* では、go to の頻度 $168 \div 37,004$ (表2の *EasyConc* の中学の延べ語数) $\times 10,000$ で算出した。

双方ともに、順位1位として go to, 順位2位として want to が占めているが順位3位以降については、定型表現の顔ぶれも異なっている。

・表現を難しくする要因——定型表現とその後に続く表現や語句に着目

頻度順位の最も高い go to, want to については、小学校から学習する表現だけに go to, want to は学習者に定着しやすいと考えられる。英語で言いたかった表現として学習者が求めているのは、実は go to, want to の後に続く表現や語句であると考えている。例えば、中学生が将来の夢を語るチャットなどの活動では、「want to be + 職業名」で表現する学習者が多い。その際、職業名が英語で表現できないために会話が中断してしまうケースが考えられる。go to や want to の後に続く語句についてはさらに調査する必要がある。

4.3 研究の目的 (3) インプットとしての中学校の検定教科書に現れる定型表現とコミュニケーション活動で学習者が必要としている *EasyConc* の定型表現との間にどのような関係があるか調査する。

4.3.1 *03NC* の定型表現と *EasyConc* の学習者の求める定型表現との相関関係

二つの定型表現リスト *03nc_teikei5_LR.xlsx* と *EasyConc_teikei5_LR.xlsx* を用いて、定型表現とその頻度を一つのワードリストにマージする機能を持つアプリ *wordlist_merge1.0.xlsm*⁽³⁾ を用いて合体する。表3において、定型表現 go to は、*03NC* での頻度が42、*EasyConc* では168、また want to は、*03NC* での頻度が39、*EasyConc* では165で、他の定型表現の頻度

と比べるとかなり高い。散布図を描いてみると、go to と want to の2つの定型表現は外れ値として認められるので除外する。03NC の定型表現と EasyConc の定型表現とでピアソンの積率相関係数を算出したところ、有意な正の相関関係が認められた ($r=0.475$. $p < 0.001$)。

表4 03NC と EasyConc との間の定型表現の相関

		03NC	EasyConc
03NC	Pearson の相関係数	1	.475**
	有意確率 (両側)		<.001
	度数	418	418
EasyConc	Pearson の相関係数	.475**	1
	有意確率 (両側)	<.001	
	度数	418	418

(度数 418 は、定型表現の総数 420 のうち、外れ値 go to, want to を除外した数。)

** : 1% 水準で有意 (両側)

しかしながら、EasyConc の英文は中学生が英語で表現できずに日本語で質問してきたことを中学校の検定教科書を熟知する日本人英語教師と英語母語話者教師とが中学生に代わり、教科書で学習するレベルの英文で表現したものである。もし中学生が英語で表現できたと仮定すれば、有意な正の相関が予想される可能性があるということに留意されたい。

4.3.2 少数の頻度の高い定型表現と多数の頻度の低い定型表現について

Nation (2013) では、Zipf's law を受け、George Orwell の小説 *Animal Farm* の頻度表で頻度順位 10 位, 20 位, 30 位, 40 位, 50 位, 60 位の各語とその頻度に対して $\text{rank} \times \text{frequency} = \text{a constant figure}$ を適用してみると、おおよそ同じ数値となることを示している。また、順位と頻度とのグラフを書いてみると曲線となり、少数の語の頻度がかかなり高く、大多数の語の頻度は低いことが分かっている。

03NC と EasyConc の順位付け頻度表で、定型表現の順位 \times その頻度を

算出したが一定数という結果は得られなかった。しかしながら，図3と図4の双方の曲線が示すように，共通の特徴として少数の使用頻度の高い定型表現と多数の使用頻度の低い定型表現とが曲線を描くことが示された。

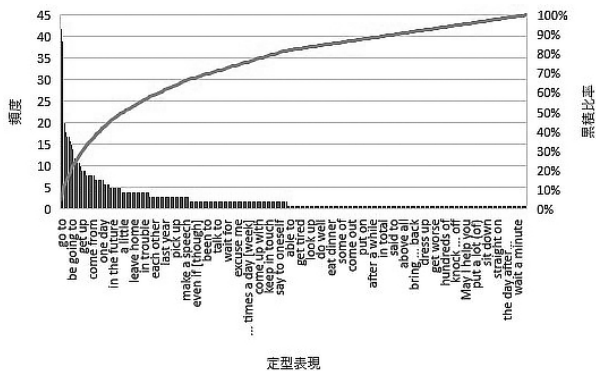


図3 検定教科書 03NC の定型表現とその頻度

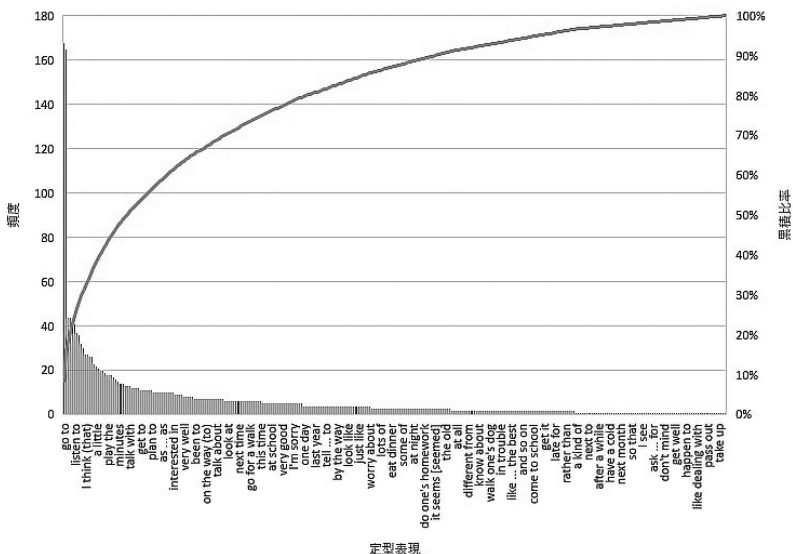


図4 EasyConc の定型表現とその頻度

4.3.3 中学校の検定教科書の定型表現リスト (03nc_teikei5_LR.xlsx) とコミュニケーション活動で学習者が必要としている *EasyConc* の定型表現リスト (*EasyConc_teikei5_LR.xlsx*) との重なりについて英語教育への示唆

検定教科書 *03NC* に出現する定型表現と *EasyConc* に出現する定型表現の重なり (Range) は以下のとおりである。

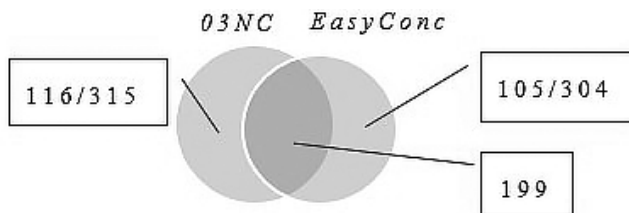


図5 定型表現の重なり

(図5には、表4で除外した外れ値の go to, want to が含まれている。)

Nation (2013) は、語彙学習や文法学習で費用対効果の原則 (cost/benefit principle) について述べており、学習者が最も使用頻度の高い語彙や文法をまず学ぶことが最も利益が高いとしている。この原則に従うと、図5の重なる部分の定型表現199は、*03NC*と*EasyConc*の両方に出現し頻度の高い定型表現であり、まず中学生が習得することが大切である。この定型表現199については付録を参照されたい。データ駆動型学習に見るように、定型表現の含まれる英文を生徒の目に触れさせる意図的な指導が必要である (日碁・仲・山田, 2016)。

5. まとめ

研究の目的 (1) では、3.2.1で定型表現リスト (teikei5.txt) とその作成過程を示した。3.2.2~3.2.3では、既存のアプリを使用し、半自動化する

ことで *03NC* と *EasyConc* から双方の定型表現を抽出し、その頻度を算出する方法を示した。

定型表現とその頻度を抽出するにあたっては、定型表現リスト *teikei5.txt*、置換リスト *teikei5_LR.txt* の役割が要となる。そのために双方のアップグレードに向けての課題がいくつか残されている。

- ・日墓, Lia, 北村 (2022) では、中学生が英語で言いたかった表現として *What else?*, *have nothing (special) to do*, *have something (else) to do*, *don't have anything to do* を挙げている。これらの日本人中学生のニーズとして必要な表現は、定型表現リストに組み込む必要がある。
- ・Martinez & Schmitt (2012) の Appendix: The PHRASE List に見られる頻度の高い語彙で構成される定型表現も利用する。
- ・*03NC* だけでなく複数の検定教科書に出現する定型表現も対象にし、分布の規準 (range criterion) を設定し定型表現を抽出する。
- ・置換リストでは、*go -ing* で一括りにせず、*go shopping* といった個別例で示すことも考えられる。置換リスト *teikei5_LR.txt* の設定方法については検討したい。

研究の目的 (2) では、4.2 の表 3 から、*go to*, *want to* の頻度は、*03NC* でも、*EasyConc* においても出現頻度が極めて高い。*go to*, *want to* の後に続く表現についてさらなる調査が必要である。

研究の目的 (3) では、4.3.1 において、検定教科書 *03NC* と *EasyConc* の定型表現との間に有意な正の相関関係が認められた ($r = 0.475$, $p < 0.001$)。ただし、*EasyConc* の英文は英語教師が学習者からの日本語の質問を受けて英訳したものであることに留意されたい。

4.3.2 の図 3 と図 4 から *03NC* も *EasyConc* も少数の使用頻度の高い定型表現と、多数の使用頻度の少ない定型表現とが出現することが分かった。

4.3.3 の図 5 から学習者のニーズを考慮し、表現力を伸ばすためには *03NC* と *EasyConc* とに共通する定型表現 199 をまず使えるようにすることが目標となる。

謝辞

EasyConc5.8.xlsm のアプリの開発では、コーパスの英訳作業で Steve Lia 氏（玉川大学）のご協力をいただいた。御礼申し上げます。

本研究は、科学研究費の支援を受けている。「日本人中高生のための英語表現コーパスの構築とその活用」JSPS 科研費 19K00771 令和元年度～令和 3 年度 研究代表者：日碁滋之

注

- (1) *EasyConc5.8.xlsm* は、日碁滋之・大村あつし氏による Excel 仕様の *EasyConc* の検索アプリ。
<http://www.tamagawa.ac.jp/research/je-parallel/>（最終閲覧日：2023 年 9 月 15 日）
- (2) *ListReplace.mac* は山紫水明氏による秀丸エディタ用のマクロファイル。変換リスト (*teikei5_LR.txt*) に従い連続して置換を実行する。
<https://hide.maruo.co.jp/lib/macro/listreplace307.html>（最終閲覧日：2023 年 9 月 15 日）
- (3) *wordlist_merge1.0.xlsm* は、日碁滋之・大村あつし氏による Excel 仕様の複数のワードリストを一つにまとめるアプリ。
<http://www.tamagawa.ac.jp/research/je-parallel/>（最終閲覧日：2023 年 9 月 15 日）

参考文献

- Martinez, R, & Schmitt, N. (2012). *A phrasal Expressions list*. *Applied Linguistics*, 33(3), 299-320.
- McCarthy, Michael. (1990). *Vocabulary*. Oxford University Press
- Nation, I.S.P. (2013). *Learning Vocabulary in Another Language, 2nd edn*. Cambridge.
- Schmitt, Nobert & Schmitt, Diane. (2020). *Vocabulary in Language Teaching 2nd Ed*. Cambridge
- Shin, Dongkwang & Nation, Paul. (2008). Beyond single words: the most

- frequent collocations in spoken English. *ELT Journal*, 62(4), 339-348.
- Webb, Stuart & Nation, I.S.P. (2017). *How Vocabulary is Learned*. Oxford
- Wray, Alison. (2002) *Formulaic Language and the Lexicon*. Cambridge.
- 北村広志・内藤清志・Steve Lia・日英滋之, 2021年, 「英語で言いたいことが言えるようになるために——日本人英語学習者のための日英パラレル・コーパス」
<https://sites.google.com/view/easyconc/for-access>（最終閲覧日：2023年9月15日）
- 日英滋之・仲圭一・山田洋, 2016. 『中学生のための英語表現 BINGO』. 明治図書.
- 日英滋之・LIA Steve 「日本人中高生のための英語表現コーパスの構築とその活用」JSPS 科研費 19K00771 令和元年度～令和3年度 研究代表者：日英滋之
- 日英滋之・佐藤浩希・茅野恵理香, (2021). 「中高大連携に基づく ICT 英語教材の開発——日英パラレル・コーパス *EasyConc* の構築と発信語彙の習得を促す *EasyConc for FlashCard.fmp12* の開発——」『玉川大学教師教育リサーチセンター年報 第11号』
- 日英滋之・Lia, Steve・北村広志, (2022). 「中・高・大学生のニーズを反映した日英パラレルコーパス *EasyConc* の検索ソフト *EasyConc for Access.accdb* の開発」『玉川大学教師教育リサーチセンター年報 第12号』
- 文部科学省 (2017) 『中学校学習指導要領 (平成 29 年告示) 解説 外国語編 英語編』 開隆堂

(原稿受付 2023年10月23日)

付録

03NC と EasyConc とに出現する定型表現の重なり 199
(順位は, EasyConc の頻度順で表示。)

1	go to	51	need to	101	leave home	151	in the hospital
2	want to	52	one of	102	look like	152	in time
3	a lot (of)	53	get tired	103	one day	153	in trouble
4	kind of	54	how to	104	out of	154	know about
5	like to	55	in particular	105	over here [there]	155	like ... the best
6	listen to	56	very well	106	ready for [to]	156	like this
7	like -ing	57	a member of	107	shall we	157	pick up
8	have to	58	been to	108	tell ... to	158	put on
9	go -ing	59	for the first time	109	these days	159	See you
10	because of	60	help ... with	110	wait for	160	such as
11	at home	61	how long	111	... old	161	take a walk
12	I think (that)	62	in the future	112	all over the world	162	that one
13	able to	63	talk about	113	at night	163	this (...) one
14	be going to	64	around the world	114	by oneself	164	very nice
15	on TV [video]	65	begin to	115	decide to	165	walk one's dog
16	a little	66	belong to	116	depend on	166	walk to
17	go out	67	famous for	117	do one's homework	167	... times a day [week]
18	good at (-ing)	68	get in	118	eat dinner	168	a couple of
19	look for	69	grow up	119	eat lunch	169	a fan of
20	so ... that	70	look at	120	find out	170	a kind of
21	do one's best	71	middle of	121	full of	171	after a while
22	play the	72	next time	122	get married	172	agree with
23	get home	73	years ago	123	give up	173	come across
24	get up	74	a few	124	go home	174	covered with
25	brush ... teeth	75	anything else	125	have a great time	175	go away
26	every day [morning, night, year]	76	at school	126	important to	176	have a cold
27	in the morning	77	at that time	127	lots of	177	How about you
28	talk with	78	come from	128	made of	178	I see
29	watch TV	79	for example	129	many kinds of	179	I'd love to
30	each other	80	I'm sorry	130	next week	180	in total
31	even if [though]	81	know that	131	some of	181	like ... better than
32	take a bath	82	last week	132	surprised (at [to])	182	look around
33	think about	83	look up	133	think of	183	next month
34	after school	84	take care of	134	a hundred [a few, many, million] times	184	next to
35	arrive in [at]	85	talk to	135	wash one's face	185	not ... any
36	from ... to	86	used to	136	a (little) bit	186	not only
37	get to	87	very good	137	at all	187	rich in
38	stop -ing	88	Why don't you	138	at first	188	said to
39	as ... as	89	afraid of [that]	139	at last	189	so that
40	for a long time	90	all the time	140	at least	190	suffer from
41	I hope (that)	91	as soon as	141	come back	191	take a [some] photograph(s) [photo(s), picture(s)]
42	interested in	92	based on	142	come in	192	take out
43	Thank you	93	by the way	143	come out	193	the number of
44	this year	94	do well	144	different from	194	the other (one)
45	try to	95	due to	145	eat breakfast	195	this morning
46	very much	96	for a while	146	excuse me	196	this summer
47	What time	97	go back	147	for one's birthday	197	turn off
48	would [d] like to	98	have fun	148	make [be, stay] friends with	198	What about
49	come home	99	last night	149	get off	199	what to
50	more than	100	last year	150	in order to		