様　式　Ｃ－１９、Ｆ－１９、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成 ２８ 年 １０ 月　 ６ 日現在

機関番号：３４３１５
研究種目：挑戦的萌芽研究
研究期間：2013～2015
課題番号：２５５８０１２６

研究課題名（和文）平易な語から成る自然な英語表現の体系的抽出方法の研究と大学生用英作文教材の開発

研究課題名（英文）A study of how to extract commonly used English expressions consisting of basic words and an attempt to develop materials for improving the writing ability of Japanese university students

研究代表者
滝沢　直宏（TAKIZAWA, Naohiro）

立命館大学・言語教育情報研究科・教授

研究者番号：６０２５２２８５

交付決定額（研究期間全体）：（直接経費）　　1,900,000円

研究成果の概要（和文）：長年にわたる大学での英語教育の経験から、大学生たちは、中学校の時に習い覚えるような平易な語を有効に活用できないことに気付いた。そこで、大学生にとって平易な語で構成されている表現を大規模な言語資料（コーパス）から抽出することで、英語表現力向上に資する表現を可能な限り機械的に抽出する方法を模索した。成果として、コロケーションの認定や、近年の英語の変化を高速且つ正確に捉えるコンピュータプログラムの開発を行い、それを用いた表現抽出を行った。

研究成果の概要（英文）：Judging from my own experience of teaching English to university students in Japan, it can be said that there are a lot of English expressions which are easy to understand but difficult for students to use, even though they may consist of basic words. (For example, "Nobunaga was born in what is now Aichi," is not easy, even though students are familiar with the relative pronoun "what." Likewise, students can readily use "fresh" but they cannot appropriately use "freshly," although the meaning is clear to them.) This project has focused on the methodology of extracting as exhaustively as possible "easy-to-understand-but-difficult-to-use" expressions from large-sized corpora and on the development of teaching materials for university students. This research has led to the development of computer programs which can be used for identifying useful expressions in the sense described above. Using these programs, we succeeded in extracting a large variety of useful expressions.

研究分野：英語学

キーワード：中学必須単語　ly副詞　コーパス　コロケーション　歴史的変遷

１．研究開始当初の背景
研究代表者は、前任校において英語の授業を２０年半にわたって担当してきた。その際、英語の表現力向上にも留意してきたが、学生たちは、中学校の時に習い覚えるような平易な語を有効に活用できないことに気付くことが頻繁にあった（具体例は下記の「研究の目的」で記す）。そこで、大学生にとっては比較的平易に思われる英文を大量に読み、その際、「意味はよく分かるが、自分ではなかなか書けない」表現に留意するように促してきた。難しい単語や熟語を覚えるだけではなく、平易な語の「意外な」組み合わせに留意することは重要であると確信した。

そうした表現は、読書によって（つまり手作業で）収集することは研究代表者自身、長年行ってきたが、可能な限り、機械的に抽出する方法はないものかと考えてきた。

本研究では、網羅的とはいかないまでも、これまであまり気付かれていなかった表現を可能な限り抽出し、それを反映した辞書・語法書の類を作成することは有意義なことであると考え、研究を進めることとした。

２．研究の目的
英語には、読めば自信をもって解釈できても、自分から用いることが困難な表現が多々ある。例えば、「全ての人間は、生まれながらにして一つの言語を習得する能力をもっている。」All humans are born with the ability to acquire a language.という文には、難しい語は全く使われておらず、高校生であっても意味は容易に分かるが、be born と with 句を組み合わせて、このように書ける大学生はそれほど多くない（ことが経験上、分かっている）。「信長は今日の愛知県で生まれた」Nobunaga was born in what is now Aichi Prefecture.も同様で、関係代名詞 what は習っていても、このように what を用いることができる大学生は少数である。本研究では、この種の「書けそうで書けない英文」を大規模コーパスから半自動的に抽出する方法を提案することを第１の目的とした。そして、その方法を用いて、実際に同種の表現を見つけ出し、自然な英文から成る例文データベースを構築すると同時に、それを駆使した辞書あるいは語法書を作成し、英語学習者（主として大学生レベル）の表現力の向上を図ることを第２の目的とした。その際、既存の辞書で必ずしも十分な記述がなされているとは言いがたい-ly 副詞が含まれる表現（freshly baked bread「焼きたてのパン」のような表現）を特に重視することとした。

３．研究の方法
（１）まずは「書けそうで書けない英文」がどのようなものであるかを具体的且つ詳細に知ることが重要であると判断した。研究代表者は、英文として興味深い文・表現を読書によって見つけ出し、それをデータベース化するという作業を 1995 年から 15 年以上にわたって行ってきている。その例文の数は、科研研究開始時において 35,000 を超える数にのぼっていた。その例文を仔細に再検討すると同時に、日頃の読書等によってデータベースの拡張を進めることとした。データベースには、単に出典と例文を記載するのみならず、どのような点で興味深いのかを示すタグを付与してきたが、そのタグ付けを精緻にすることも同時並行で行った。

（２）その上で、「書けそうで書けない英語」とはどのような特徴をもつものであるのかを検討した。具体的には、作業仮説として、(i)中学生でも知っている平易な語、あるいはそれに接辞が付加された語で構成された表現（パターン）で、(ii)give up「諦める」などと異なり、熟語化しておらず、全体の意味が部分の意味の総和から得られる表現で、(iii)辞書に頼ることなく平易に意味が理解できるにもかかわらず、大学生英語学習者がなかなか使うことができない表現」とした。例えば、「研究の目的」の節で述べたものや、be fresh out of school「学校を出たばかり」などの表現（パターン）がそれに該当する。これらの語には難しい語は全く使われていない。また fresh を使って何らかの表現を作成するように言うと大学生であれば、fresh fruits のような表現は即座に作れるものの、freshly を使った表現となるとかなりの上級者でもなかなか自然な表現を作ることができない。この難易の差に注目することとした。

（３）自作のデータベースには、可能な限り、自然な日本語訳を追加することにより、パラレルコーパスにあたるものを自作した。例えば、"be born with 「能力・病名」"を含む例文には「生まれながらにして・・・できる能力をもっている/・・・という病気を煩っている」という訳文を与えることで、日本語と英語の間に見られる「ずれ」が明らかになるように試みた（英語では述語となっているものが、日本語では副詞的に表現されている、など）。freshly baked bread であれば「焼きたてのパン」であって、日本語では非自立要素である「たて」が freshly に対応することで、この表現を学習者自らが使うことが困難になっていることなどを見た。

（４）（３）で見たようなパターンを大量に得るためには、大規模コーパスの整備が重要である。表現パターン抽出の試みの一環として、勤務校のコーパスサーバー上に Corpus of Contemporary American English (COCA), Corpus of Historical American English (COHA), Corpus of Global Web-Based English (GloWbE)のテキストデータを整備し、上記で同定した「書けそうで書けない英語」の実態を詳細に検討することとした。実際に検討する過程で、いずれのコーパスにも公開されて

はいないコーパスの問題点（コーパスを構築する際に生じたと思われる欠陥や不統一）が意外に多いことが判明した。そのため、コーパスを利用する前にコーパス自体を詳細に検討することが不可欠であると考えた。また、British National Corpus などにも同様に公にはなっていない問題があることが判明したので、その問題を可能な限り網羅的に同定し、問題を回避しつつ、有意義な情報を抽出する方法を模索することとした。

（５）本研究では、英文の日本語訳が関わるため、既に勤務校で整備している日本語のコーパスを用いて、頻出パターンの同定に努めることとした。

（６）整備したコーパスからは、有益な情報を高速に抽出するシステムを構築する必要があると実感し、システムエンジニアの西村祐一氏に依頼して、大規模なコーパスからコロケーションの認定やパターン抽出に有益なシステムの構築を依頼した。平成 26 年度において完成を見、更に平成 27 年度にもその性能の向上に努めることとした。このシステムは、大規模なコーパスを瞬時に処理できることを目標に構築した。

更に、同じく西村祐一氏に依頼して、COHA という大規模な通時コーパス（アメリカ英語に限定されるものの、過去 200 年のデータが１年単位で記録されているコーパス）から頻度変化に関する情報を得るシステムを構築することとした。というのも、近年、急にその使用頻度を増した語については、既存の文献ではきちんと記述されていないとの認識による。このシステムは平成 27 年度中に完成を見、またその能力の更なる向上を図ることとした。

４．研究成果

（１）「書けそうで書けない英文」としてどのようなものがあるのか、またそれがなぜ重要であるかについては、中学・高校の英語教員を対象にした講演会において披露する機会を得た。愛知県私学協会教科等研究部英語部会における講演「中学必須単語を有効に活用した英語表現力向上に向けて」(平成 25 年 6 月 21 日：桜花学園高等学校（愛知県名古屋市）)、英語語法文法学会・第 10 回英語語法文法セミナー「使える英文法：語彙・構文研究を現場にいかす」における「文法と表現の接点：中学必須単語を有効に使う」(平成 26 年 8 月 4 日：関西学院大学・梅田キャンパス（大阪府大阪市）)、そして岐阜県英語の教え方研究会での講演「平易な語の有効活用：英文法と英語表現の接点を探る」(平成 27 年 1 月 24 日：岐阜県各務原市立中央図書館（岐阜県各務原市）)である。

（２）コーパス利用の方法論に関しては、英語コーパス学会の春期シンポジウムにおいて２度発表した。１度目は、平成 25 年度英語コーパス学会主催シンポジウム「私のコーパス利用」での「コーパスの（消極的）利用：語法・文法・構文研究のために」(平成 25 年 4 月 27 日：大阪大学・豊中キャンパス（大阪府豊中市）)という発表である。２度目は平成 27 年度英語コーパス学会春季シンポジウム「コーパス関連専門科目の授業内容について」での「コーパス関連専門科目の内容：言語を探るためのコーパス利用」（司会・発表）(平成 27 年 4 月 25 日：関西大学・千里山キャンパス（大阪府吹田市）)という発表である。前者では、コーパス利用の方法を、後者では、従来、あまり披露されることのなかった主として大学院レベルでの「コーパス教育」のあり方を議論することを目的とした（後者は研究代表者による企画である）。

（３）「研究の方法」の（４）で示した通り、コーパス利用環境を整備する過程で、様々な問題に遭遇し、コーパスを利用する前にコーパス自体の構成などについて詳細に検討する必要を痛感した。そのため、それに関連する内容をまとめ、英語コーパス学会・第 41 回大会(愛知大学（愛知県名古屋市）)（平成 27 年 10 月 4 日）において「コーパスからの情報抽出と抽出データの意味づけ：具体的表現からの抽象化」というテーマで発表した（講演）。その内容は、英語コーパス学会の学会誌『英語コーパス研究』(平成 28 年 3 月)において活字の形で公表した。

この講演では、当該コーパスに関する説明書・マニュアルには全く明記されていない問題点があることを明らかにした。この問題点を回避することなしに、コーパスの精緻な利用は不可能であるので、本科研研究の中心課題とはずれるものの、コーパス利用の前提として重要であると判断し、平成 27 年度はこの問題について、多くのエネルギーを費すこととなった。

（４）平易な語の使用に関する表現については、名詞 age に着目した論文を書き、また英語のコーパス利用の方法論については、研究代表者と深谷輝彦氏が編集した『コーパスと英文法・語法』をひつじ書房から出版した。（同書には、語法文法に関する多くの論文が収録されている。）

（５）本科研は、英文の日本語訳も重視しているので、日本語のコーパスを使った日本語のパターン研究も重要である。そこで、國立高雄第一科技大學應用日語系・台灣應用日語學會からの招待を受けたのを契機として、日本語のコーパス処理も行い、「日本語におけるコロケーションと語彙的パターンについて」というテーマで講演を行った（平成 26 年 5 月 24 日（台湾・高雄市））。

（６）大規模なコーパスからコロケーション

の認定やパターン抽出を行う西村祐一氏開発のシステムを myco と命名し、その設計原理などを同氏との共著で『言語科学論集』において公にした（平成 26 年度）。myco とは、MySQL を用いたコーパス分析システムという意味である。また通時的コーパスから頻度変化に関する情報を得るシステムは、WLP_dfc と命名した。これは、WLP 形式（word/lemma/pos が１行に納められている形式）から通時的頻度変化（diachronic frequency change）を見るためのシステムという意味である。（いずれも特定のコーパスに特化されたシステムというわけではない点が重要である。）WLP_dfc については、そのシステムについてだけではなくそれを使った ly 副詞の頻度変化について平成 28 年 3 月刊行予定であった『現代英語談話会論集』に掲載することとした。

（７）myco や WLP_dfc を用い、また自作のデータベースを駆使して、多くの「書けそうで書けない表現」の同定は既に済んでいる。現在は、これをもとに大学生向けの語法書などの出版を企図している。

（８）最終年度の年度末である平成 28 年 2 月には英作文教育に関するシンポジウム「英語表現および英作文教育に関するシンポジウム」を行った（立命館大学（京都府京都市））。このシンポジウムでは、myco の開発者である西村祐一氏、高校の英語教員、英作文関係の科研で研究を進めている研究者をパネリストとした。研究代表者は司会をすると共に、発表も行った。

５．主な発表論文等
（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計５件）
① 滝沢直宏・西村祐一. 2016.「WLP_dfc による通時的変化の記述 ― Corpus of Historical American English における ly 副詞を例に―」『現代英語談話会論集』11（印刷中）査読有.

② 滝沢直宏. 2016.「コーパスからの情報抽出と抽出データの意味づけに関わる諸問題」『英語コーパス研究』23: 45-60. 査読有.

③ 滝沢直宏. 2015.「age の意味と用法―「優先規則体系」からの考察―」『現代英語談話会論集』10: 37-54. 査読有.

④ 滝沢直宏. 2014.「日本語におけるコロケーションと語彙的パターンについて」『2014 年台灣應用日語教學國際學術研討會予稿集』pp. 15-28. 査読無.

⑤ 西村祐一・滝沢直宏. 2014.「リレーショナルデータベースを用いたコーパスからの情報抽出：その方法について」『言語科学論集』4: 87-111. 立命館大学言語教育情報研究科. 査読有.
URL:
http://r-cube.ritsumei.ac.jp/bitstream/10367/5524/1/LEIS_4nishimura.pdf

〔学会発表〕（計７件）
① 滝沢直宏.「英作文教育における中学基本単語の有効活用」立命館大学・言語教育情報研究科・国際言語文化研究所共催・シンポジウム「英語表現および英作文教育に関するシンポジウム」（企画・司会・発表）（2016 年 2 月 6 日：立命館大学・衣笠キャンパス（京都府京都市））

② 滝沢直宏.「ly 副詞にみる副詞の多様性：語法・文法・パターン」英語語法文法学会・第 23 回大会・シンポジウム「副詞を巡る諸問題：語法文法、辞書記述、談話、文体」（企画・司会・発表）（2015 年 10 月 24 日：龍谷大学（京都府京都市））

③ 滝沢直宏.「コーパスからの情報抽出と抽出データの意味づけ：具体的表現からの抽象化」（講演）英語コーパス学会・第 41 回大会(2015 年 10 月 4 日：愛知大学（愛知県名古屋市））

④ 滝沢直宏.「コーパス関連専門科目の内容：言語を探るためのコーパス利用」平成 27 年度英語コーパス学会春季シンポジウム「コーパス関連専門科目の授業内容について」（企画・司会・発表）（2015 年 4 月 25 日：関西大学・千里山キャンパス（大阪府吹田市））

⑤ 滝沢直宏.「文法と表現の接点：中学必須単語を有効に使う」英語語法文法学会・第 10 回英語語法文法セミナー「使える英文法：語彙・構文研究を現場にいかす」（2014 年 8 月 4 日：関西学院大学・梅田キャンパス（大阪府大阪市））

⑥ 滝沢直宏.「日本語におけるコロケーションと語彙的パターンについて」（招待講演）國立高雄第一科技大學應用日語系・台灣應用日語學會(2014 年 5 月 24 日：國立高雄第一科技大學（高雄市（台湾））

⑦ 滝沢直宏.「コーパスの（消極的）利用：語法・文法・構文研究のために」英語コーパス学会主催シンポジウム「私のコーパス利用」（2013 年 4 月 27 日：大阪大学・豊中キャンパス（大阪府豊中市））

〔図書〕（計 2 件）
① 前川喜久雄(監修)・砂川有里子（編）2016.『コーパスと日本語教育』pp. 1-34. 朝倉書店（総ページ：216）（滝沢直宏・千葉庄寿「第

１章　コーパス検索の方法」pp. 1-34 および滝沢直宏「付録　正規表現とコーパス」pp. 189-197）

②　深谷輝彦・滝沢直宏.（共編著）2015.『コーパスと英文法・語法』（英語コーパス研究シリーズ第４巻）ひつじ書房（総ページ：260）（滝沢直宏「コーパスを用いた英語語法文法の研究　その方法を中心に」（pp. 23-39））

〔産業財産権〕
○出願状況（計　　件）

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況（計　　件）

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕
ホームページ等

６．研究組織
(1)研究代表者
滝沢 直宏（TAKIZAWA, Naohiro）
　立命館大学・言語教育情報研究科・教授
　研究者番号：６０２５２２８５

(2)研究分担者
（　　）

　研究者番号：

(3)連携研究者
（　　）

　研究者番号：