Kobe University Repository : Kernel

PDF issue: 2025-07-07

話し言葉コーパスと検定教科書に基づく日本人英語学習者の句動詞使用実態の分析

石井，康毅

(Citation)
Learner Corpus Studies in Asia and the World,3:101-119

(Issue Date)
2018-03-12

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/81010121

(URL)
https://hdl.handle.net/20.500.14094/81010121

話し言葉コーパスと検定教科書に基づく日本人英語学習者の句動詞使用実態の分析

石井　康毅（成城大学）
ishii@seijo.ac.jp

An Analysis of Phrasal Verbs Used by Japanese EFL Learners: Based on Spoken Learner Corpora and Authorized English Textbooks

ISHII Yasutake (Seijo University)

概要

日本人英語学習者は一般に句動詞をうまく使うことができない。この問題の原因と解決策を探るために，日本人英語学習者の話し言葉に焦点を当て，The NICT JLE Corpus と The ICNALE の Spoken Monologue モジュールを利用し，学習者と母語話者の句動詞使用実態を比較分析した。さらに，日本人英語学習者の句動詞学習状況も考慮するために，中高の英語検定教科書における句動詞使用状況も分析した。分析対象の句動詞は，句動詞辞典の収録項目に基づく約 7,000 項目である。その結果，学習者のレベルが上がるにつれて句動詞使用頻度は上昇し，母語話者に近づいていくが，使用する句動詞の項目は母語話者とは大きく異なり，教科書で出現する句動詞と学習者が使用する句動詞にも大きな違いがあることが明らかになった。

キーワード

句動詞，日本人英語学習者，話し言葉コーパス，検定教科書

1. はじめに

英語の句動詞（phrasal verb）は英語という言語に特徴的な構造のひとつであり，その使用頻度は決して低くない。Biber, Johansson, Leech, Conrad, and Finegan（1999, pp. 408–409, 415）は，フィクション・会話では 100 万語あたり 2,000 以上の句動詞（彼らの定義では動詞＋副詞パーティクル）が，そして5,000以上の前置詞付動詞（prepositional verb）が使われると指摘している。筆者が The British National Corpus における高頻度語（レマベース）と 5 つの句動詞辞典（3.3.1 参照）に収録されている句動詞の頻度を比較調査したところ，句動詞を単語と同様の語彙項目として扱うとすると，特に日常よく使う中頻度語彙（3,000～10,000 語程度の水準）で句動詞の占める割合は約 5%にも及ぶということが分かった。

このように，句動詞は英語の語彙を構成する重要な要素であり，学習者にとっても学習・習得の必要性の高いものであることは明白である。しかしながら，特に基本語彙を身に付ける重要な段階である初中級段階の英語教育において，句動詞は十分な指導がなされているとは言いがたく，日本人英語学習者は句動詞をうまく使うことができていないということが指摘されている（Yoshitomi, 2012; 前浜，2015）。この問題の原因と解決の方向性を探るためには，様々なコーパスにおける

句動詞の使用実態を分析する必要がある。2 節で見るように，母語話者コーパスや学習者の作文コーパスを対象とした句動詞の調査に関しては，ある程度の規模の調査に基づく研究が既に行われているが，学習者の話し言葉を対象とした，大規模でかつ網羅的な句動詞使用実態調査は，筆者の知る範囲では行われていない。そこで，本研究では日本人英語学習者の話し言葉コーパスを利用し，さらに母語話者のデータや検定教科書のデータも参照し，日本人英語学習者が学習・使用する句動詞の実態を分析する。

2. 先行研究

大規模コーパスから高頻度句動詞を抽出した先行研究には Gardner and Davies（2007）や Liu（2011）などがあるが，これらは母語話者コーパスを対象とした研究である。また，分析対象を動詞＋副詞パーティクルの狭い意味での句動詞に限定しているが，句動詞範疇の周縁は非句動詞と連続的であり（Ishii, 2009, pp. 126–128），前置詞付動詞や前置詞付句動詞（prepositional phrasal verb）も含めて広い意味での句動詞を分析し，高頻度の複合表現を明らかにする必要がある。

学習者コーパスと母語話者コーパスを対照し，句動詞の頻度を比較したものとしては Uchida（2012）や飯尾（2013）がある。前者は前置詞付（句）動詞も含めて抽出・集計しているが，検索対象の語形や補部のパターンが限定的である。後者は分析対象が動詞＋副詞パーティクルに限られている。さらに，Uchida（2012）は The JEFLL Corpus を，飯尾（2013）は日本人学習者のコーパスとしては The JEFLL Corpus と NICE を分析しており，いずれも日本人学習者の作文コーパスのみを分析していて，話し言葉コーパスは分析対象とはなっていない。

作文コーパスは一般にトピックの影響を強く受けた論述となることが多く，また，書き言葉と話し言葉とでは言語使用に大きな違いがあることから，言語使用の実態を多角的に検討するためには，話し言葉コーパスのデータも分析する必要がある。日本人英語学習者の話し言葉コーパスにおける句動詞使用を調査した研究としては Yoshitomi（2006），前浜（2015），石井（2017）がある。Yoshitomi（2006）はストーリーテリングにおける日本人英語学習者の句動詞使用を調査したが，分析対象の句動詞の定義は明示されていない。著者によると（A. Yoshitomi, personal communication, 2018 年 1 月 21 日），ある表現が句動詞であるかどうかの判断に当たっては，Wray（2002, p. 9）の定型表現（formulaic sequence）の定義（a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar）に従い，辞書における句動詞としての収録状況も勘案して判断したが，ストーリーテリングというタスクの性質上，ある程度決まった句動詞が使われ，判断に迷うようなことはほとんどなかったということである。また，Yoshitomi（2006）が分析対象とした学習者は中上級者 20 名に限られ，データサイズの小ささは否めない。前浜（2015）は The ICNALE の書き言葉と話し言葉データにおける句動詞を調査したが，「内外の 10 種類の TESOL 向け文法書を調査し，その中で 3 種類以上に共通して掲載されている句動詞 76 種を選

び，分析の対象とした」（p. 9）ということで，句動詞を網羅的に調査したとは言いがたい。石井（2017）は句動詞辞典の収録項目に基づく約 7,000 項目の句動詞項目を対象とし，各項目の可能な統語パターンに基づいて調査し，話し言葉コーパスとしては The NICT JLE Corpus を分析しているが，話し言葉コーパスに特化して分析を行っているわけではない。

本研究では，初級から上級までの日本人英語学習者の話し言葉における（広い意味での）句動詞の使用状況を母語話者と対照しながら網羅的に調査し，学習者にとって基本的なインプットである中学・高校の英語の検定教科書での扱いと比較することで，学習者が学習・使用する句動詞の実態を明らかにする。

3. 研究目的と手法

3.1 研究目的と研究設問

日本人英語学習者の句動詞の使用状況を，母語話者との比較と教科書での学習状況との比較を通して明らかにするために，本研究では次の 4 つの研究設問を設定した。

RQ1：日本人英語学習者の話し言葉における句動詞使用は，レベルが上がるにつれて，量と質の観点で母語話者の使用に近づいていくか

RQ2：日本人英語学習者が使用する句動詞は，量と質の観点で中高の検定教科書で学ぶものと類似しているか

RQ3：日本人英語学習者が多用する句動詞にはどのようなものがあるか

RQ4：日本人英語学習者がうまく使えない句動詞にはどのようなものがあるか

3.2 分析対象のデータ

本研究では 3 種のコーパスデータを分析した。それぞれのコーパスについて，3.2.1～3.2.3 で詳しく述べる。（以下に挙げる語数は TreeTagger による処理（3.3.2 参照）の結果に基づき，所有の“s”と句読点類を除き，TreeTagger が単語と認定したものを集計した結果である。）

3.2.1 中学校・高等学校用の英語の検定教科書コーパス（以降「教科書コーパス」と略記）

中学校用教科書については，平成 28 年度使用開始の 6 社による 18 点全てをコーパス化した。高等学校用教科書については，平成 25～27 年度使用開始のコミュニケーション英語 I・II・III と英語表現 I・II の計 75 点をコーパス化した。コーパス化に際しては，教科書中のリーディングの対象となる英語部分，ライティングのモデル文，文法提示用例文などを対象とした。“Read”などの見出し，英語で書かれた指示文，並べ替えや穴埋め練習問題とその選択肢，付録的なページに挙げられた語彙・表現・文法例文等は基本的に対象外とした。それぞれの教科書種別ごとの，コーパス化した冊数，コーパス化した教科書のシェアの合計（渡辺（2013, 2014, 2015）による各使用開始年の採用冊数に基づく），語数を表 1 に挙げる。（具体的な教科書名は表 1 の注に挙げる。）

表 1
学習者コーパスの内訳

教科書種別（使用開始年度）	冊数	シェア	語数
中学校用（平成 28 年度）	18	100%	124,658
コミュニケーション英語 I（平成 25 年度）	16	81.7%	144,600
コミュニケーション英語 II（平成 26 年度）	15	79.6%	180,198
コミュニケーション英語 III（平成 27 年度）	14	85.8%	204,991
英語表現 I（平成 25 年度）	17	100%	82,775
英語表現 II（平成 26 年度）	13	100%	102,258

注 コーパス化した教科書の略称を以下に挙げる。
中学校用（それぞれの 1 年生用・2 年生用・3 年生用）：COLUMBUS 21, NEW CROWN, NEW HORIZON, OEN WORLD, SUNSHINE, TOTAL ENGLISH
コミュニケーション英語 I：All Aboard!, BIG DIPPER, COMET, CROWN, ELEMENT, Grove, LANDMARK, MAINSTREAM, MY WAY, Perspective, Power On, PROMINENCE, PRO-VISION, VISTA, Vivid, WORLD TREK
コミュニケーション英語 II：All Aboard!, BIG DIPPER, COMET, CROWN, ELEMENT, Grove, LANDMARK, MAINSTREAM, MY WAY, Perspective, Power On, PRO-VISION, VISTA, Vivid, WORLD TREK
コミュニケーション英語 III：All Aboard!, BIG DIPPER, CROWN, ELEMENT, Grove, LANDMARK, MAINSTREAM, MY WAY, Perspective, Power On, PROMINENCE, PRO-VISION, Vivid, WORLD TREK
英語表現 I：BIG DIPPER, COSMOS, CROWN, Departure, Grove, MAINSTREAM, MONUMENT, MY WAY, NEW FAVORITE, New ONE WORLD, Perspective, POLESTAR, SCREENPLAY, UNICORN, VISION QUEST Advanced, VISION QUEST Standard, Vivid
英語表現 II：BIG DIPPER, CROWN, Departure, Grove, MAINSTREAM, MY WAY, NEW FAVORITE, New ONE WORLD, Perspective, POLESTAR, UNICORN, VISION QUEST, Vivid

3.2.2 The NICT JLE Corpus の受験者発話データ（以降「NICT JLE」と略記）

The NICT JLE Corpus は，英語インタビューテストである Standard Speaking Test（SST）の書き起こしデータである。データは，日本人 1,281 人のテストから成り，インタビュアーの発話も書き起こされている。さらに，比較用の母語話者のインタビューの書き起こしデータも含まれている。このデータに対して，インタビュアーの発話部分を取り除き，日本語部分と固有名詞と聞き取り不可能な語をそれぞれ固有の文字列に置換し，フィラー・ポーズ・非言語音を削除するという前処理を施した。The NICT JLE Corpus の全てのデータには各参加者の SST レベル（このテストで判定

される 1～9 のレベル）情報が付与されているが，本研究では，SST レベルと CEFR-J レベルの対応を調査した金子・和泉（2013）の対応付けに従い，CEFR-J レベル別にデータをまとめ直した。ただし，金子・和泉（2013）は，SST レベル 1 は CEFR-J レベルでは PreA1 に，2 と 3 の一部は A1.1 に，3 の残りは A1.2 に，4 は A1.3 と A2.1 に，5 は A2.2 に，6 と 7 は B1.1 に，8 は B1.2 に，9 は B2.1-C2 に対応するとしている。そのため，本研究では，SST レベル 9 を除いて複数の CEFR レベルにまたがる SST レベルのデータは扱わないこととし，SST レベル 2 と 3 を A1 に，SST レベル 5 を A2 に，SST レベル 6 と 7 を B1.1 に，SST レベル 8 を B1.2 に，SST レベル 9 を B2 に対応するものとして扱った。各レベルの語数は表 2 に示す。

3.2.3 The ICNALE Spoken Monologue 中の日本人大学生と母語話者のデータ（以降「ICNALE-S」と略記）

The ICNALE（International Corpus Network of Asian Learners of English）はアジアの 10 の国と地域で，大学生のアルバイトまたはレストランでの禁煙というテーマについての賛否を述べる大学生の発話・エッセイを集めたコーパスである。The NICT JLE Corpus と同様に，比較用に母語話者のデータも含まれている。本研究では，発話を電話録音した Spoken Monologue のモジュール（Version 2.0）を分析対象とし，日本人大学生と母語話者のデータのみを利用した。データは参加者の熟達度によって CEFR レベルに分かれているが，NICT JLE とは異なり，発話内容の評価ではなく，あくまで各参加者の英語力テスト（TOEIC・語彙テスト等）の結果に基づく熟達度である。また，The ICNALE における B1_1 と B1_2 の区分は英語力テストに基づく区分であり，CEFR-J レベルの B1.1 と B1.2 の区分と同一ではないが，それぞれ後者に該当するとみなした。さらに，The ICNALE の B2+も CEFR レベルの B2 に該当するものとみなした。各レベルの語数は表 2 に示す。

表 2
学習者コーパスの各サブコーパスの語数

	A1	A2	B1.1	B1.2	B2	母語話者
NICT JLE	131,676	242,759	250,884	77,561	60,411	95,054
ICNALE-S	NA	7,176	12,394	12,065	9,892	93,570

3.3 手法

3.3.1 句動詞の選定と生起パターンの定義

句動詞と非句動詞を明確に区別しながら定義することは容易ではない（Ishii, 2009, pp. 126–128）。そこで本研究では，句動詞辞典の収録項目に基づいて，分析対象とする句動詞を定めることにした。筆者が学習者向け英英句動詞辞典 5 点（*Cambridge Phrasal Verbs Dictionary*, 2006; *Collins COBUILD Dictionary of Phrasal Verbs* (2nd ed.), 2002; *Longman Phrasal Verbs Dictionary*, 2000; *Macmillan Phrasal Verbs Plus*, 2005; *Oxford Dictionary of*

Phrasal Verbs, 1993)の全収録項目を調査し，動詞と副詞パーティクルまたは前置詞から成る項目のリストを作成したところ，5 つの句動詞辞典のうちのいずれか 1 点以上に収録されている句動詞の異なり項目数は 6,940 であった。(各辞書の収録項目数の詳細については，石井(2017, pp. 2–3)を参照。)この句動詞リストの全ての項目には，各項目が句動詞として生起し得るパターンを品詞情報に基づく正規表現で記述した情報も付与している。(「正規表現」とは，通常の文字と特殊な意味を持つ「メタ文字」を使って，文字列を文字のパターンとして柔軟に表現し，通常の文字列検索では探し出せない文字列のパターンを検索・置換することを可能にする特別な記法のことである。)このリストにより，動詞と副詞パーティクルの間に名詞句の補部がある場合や受動態で用いられている場合なども含めて，一定の精度で大規模コーパスから各句動詞を抽出し，頻度を集計することが可能となっている。(各句動詞に対応する正規表現の具体例は，石井(2017, pp. 4–5)を参照。)

3.3.2 分析の手順

教科書コーパスと NICT JLE を TreeTagger(Schmid, 1994)で処理し，各語に品詞とレマの情報を付与した。ICNALE-S には同等の処理が予め施されている。これらのデータを句動詞リストの正規表現で検索できるように変換し，各句動詞のパターンにマッチするものの頻度を集計した。(処理の詳細は石井(2017, pp. 7-8)参照。)本研究で作成した頻度データには，TreeTagger の処理に含まれる誤りや各項目のパターン定義に起因する誤りも含まれるが，不正確さよりも大量のデータを自動処理することで得られる情報の価値を優先した。

4. 結果と考察

4.1 RQ1：日本人英語学習者の話し言葉における句動詞使用は，レベルが上がるにつれて，量と質の観点で母語話者の使用に近づいていくか

学習者コーパスの各サブコーパスにおける句動詞の頻度を表 3 に，それをグラフにしたものを図 1 に示す。

表 3 学習者コーパスのサブコーパスごとの句動詞の頻度(100 万語あたり)

	A1	A2	B1.1	B1.2	B2	母語話者
NICT JLE	21,264	23,509	24,163	26,057	31,749	30,046
ICNALE-S	NA	21,182	23,156	23,539	21,735	30,159

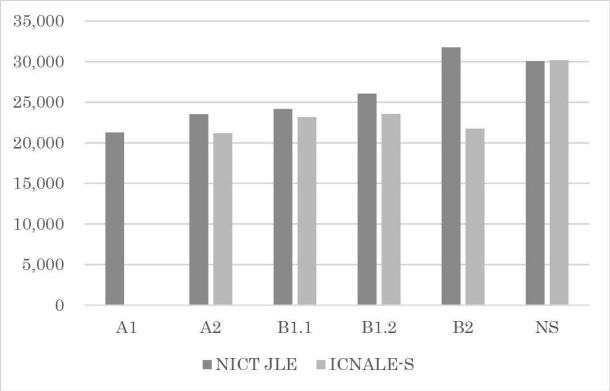

図 1. 学習者コーパスにおけるサブコーパス別の句動詞の頻度(数値は 100 万語あたり)

量の観点では, NICT JLE では, 学習者のレベルが上がるにつれて句動詞の使用頻度が上がって母語話者の頻度に近づいていく傾向が認められた。ICNALE-S はデータサイズが小さいため, 数値は慎重に解釈しなくてはならないが, B1.2 から B2 レベルにかけて頻度が低下している点を除き, NICT JLE と類似の傾向が認められる。

次に量と質の両観点からの分析として, コレスポンデンス分析(correspondence analysis)を行った。コレスポンデンス分析は対応分析とも呼ばれ, 「行と列の項目の相関 (正準相関) が最大になるように並べ替えて, 関連性が強いものやパターンが似ているものが近くになるような値を与える」(水本, 2009, p. 54), 多変量解析の手法のひとつである。本研究では, 全データでの合計度数が 100 万語あたり 30 以上で, 句動詞辞典の収録辞書数が 2 以上の句動詞 123 項目を分析対象として, R version 3.4.3 で, MASS ライブラリーに含まれる corresp 関数を利用して分析した。頻度データとしては 100 万語あたりの相対頻度を小数第 1 位で四捨五入した値を分析に用いた。なお, 視認性を高めるために, 描画には wordcloud ライブラリーに含まれる textplot 関数を利用した。

NICT JLE の各サブコーパスにおける 123 項目の句動詞の頻度をコレスポンデンス分析し, スコアに正準相関係数を掛けた値を布置した結果を図 2 に示す。分析に用いる因子数(number of factors; nf)は 6(行と列の項目数のうちの小さい方の数であるサブコーパスの数)として分析した。

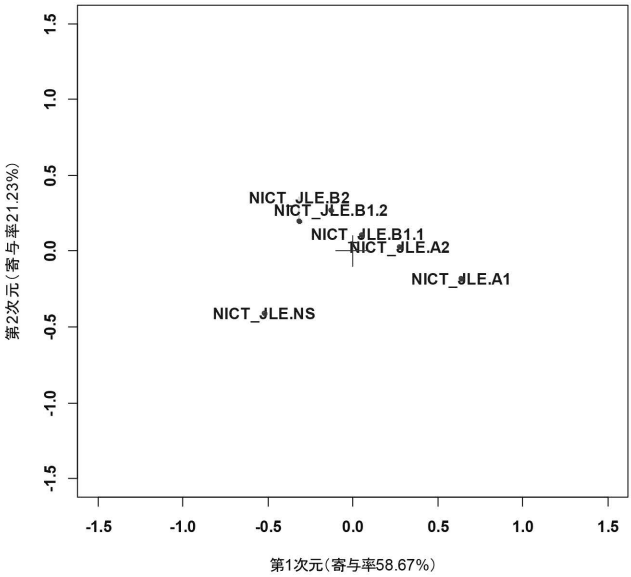

図 2. コレスポンデンス分析の結果（NICT JLE）

NICT JLE では第 1 次元にレベルが順に並んでいて，第 1 次元の寄与率も 58.67%あることから，レベルの上昇と句動詞の使用状況には密接な関係があるということが分かる。しかしながら，母語話者は大きく離れていて，B2 レベルでも母語話者とは大きな開きがある。

ICNALE-S では 20 項目の句動詞（以下，専ら受動態で用いられるものも含めて，動詞原形＋前置詞または副詞パーティクルの形で示す；arrive at, call in, come on, compare with, fall down, get off, go around, hear from, hear of, hold in, know as, look around, look up, return to, ride on, sit on, sleep on, try on, turn to, wait for）が 1 例も生起していないため，これらを除いた 103 項目の句動詞の頻度を対象とした。NICT JLE と同様に，各サブコーパスにおける 103 項目の句動詞の頻度をコレスポンデンス分析（nf=5）し，スコアに正準相関係数を掛けた値を布置した結果を図 3 に示す。

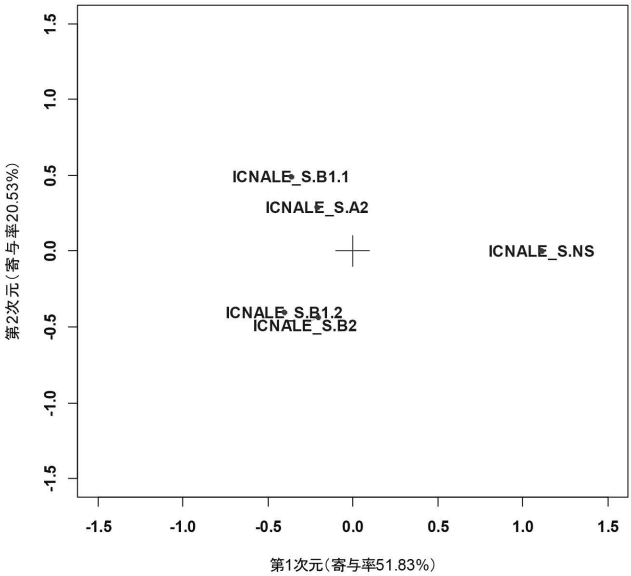

図 3. コレスポンデンス分析の結果（ICNALE-S）

ICNALE-S ではB1.1レベル以下とB1.2レベル以上の間に大きな差があるということが分かる。さらに母語話者は原点から大きく離れ，第 1 次元のマイナスとプラスで学習者と母語話者が明確に分かれていることから，NICT JLE の場合と同様に，ICNALE-S でも学習者と母語話者の間には極めて大きな差があることが示唆される。

これらの結果を総合すると，日本人英語学習者の話し言葉における句動詞使用状況は，量の観点ではレベルが上がるにつれて母語話者に近づいていくが，質の観点では上級の学習者であっても母語話者との隔たりが大きいと言える。

4.2 RQ2：日本人英語学習者が使用する句動詞は，量と質の観点で中高の検定教科書で学ぶものと類似しているか

各教科書種別における句動詞の頻度を表 4 に，教科書種別ごとの頻度と NICT JLE の各サブコーパスにおける頻度をグラフにしたものを図 4 に示す。

表 4
教科書種別ごとの句動詞の頻度（100 万語あたり）

中学	高校 C1	高校 C2	高校 C3	高校 E1	高校 E2
25,807	30,221	29,462	27,811	29,381	27,871

注 C はコミュニケーション英語，E は英語表現を表す。1・2・3 はそれぞれ科目名称の I・II・III を表す。

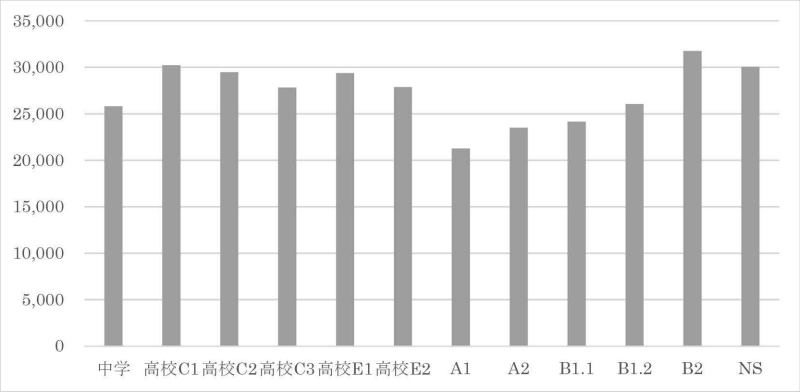

図 4. 教科書コーパスと NICT JLE の各サブコーパスにおける句動詞の頻度（数値は 100 万語あたり；NS は母語話者）

高校教科書では上位科目（英語表現 I に対する英語表現 II など）の方が句動詞の使用頻度がやや低いが，概して 100 万語あたり 28,000～30,000 の句動詞が使用されている。教科書と学習者コーパスにおける句動詞の使用頻度と比較すると，教科書における句動詞の使用頻度は学習者コーパスの B1.2～B2 レベルの頻度に近い。しかし，質の点でも学習者のインプットとアウトプットは近いのであろうか。これを明らかにするために，教科書コーパスと NICT JLE における 123 項目の句動詞の頻度をコレスポンデンス分析（nf=11）し，スコアに正準相関係数を掛けた値を布置した結果を図 5 に，学習者コーパスを ICNALE-S とし，その他は同様にコレスポンデンス分析（nf=10）した結果を図 6 に示す。

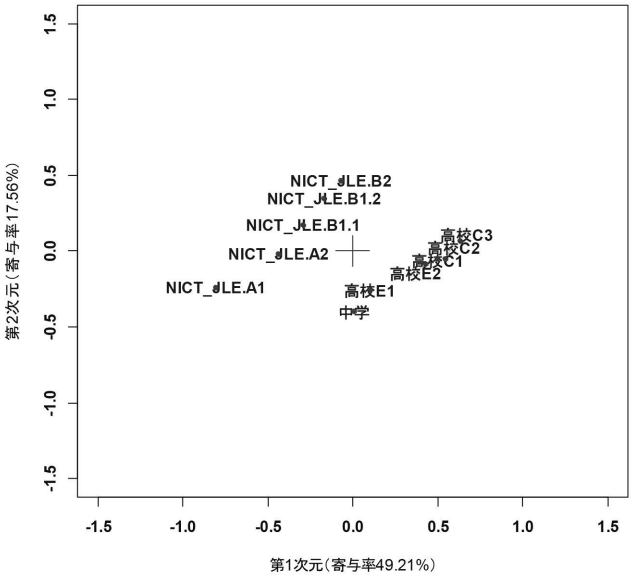

図 5. コレスポンデンス分析の結果（教科書コーパスと NICT JLE）

NICT JLE と ICNALE-S の両方で，学習者の発話と教科書の英文には隔たりがあることが分かる。各学習者コーパスのデータ収集時のタスクの影響も少なからずあるとは考えられるが，学習者が教科書で学ぶ句動詞項目と，話す場面で実際に使用する句動詞項目は異なるということが示唆される。また，4.3 と 4.4 で挙げるデータはこのことをさらに支持すると考えられる。ただし，当該の教科書で学んだ（あるいは学んでいる）学習者と NICT JLE・ICNALE-S のデータ提供者は異なるため，「日本人英語学習者は教科書で学んだ句動詞を使うことができていない」という一般化はできない。

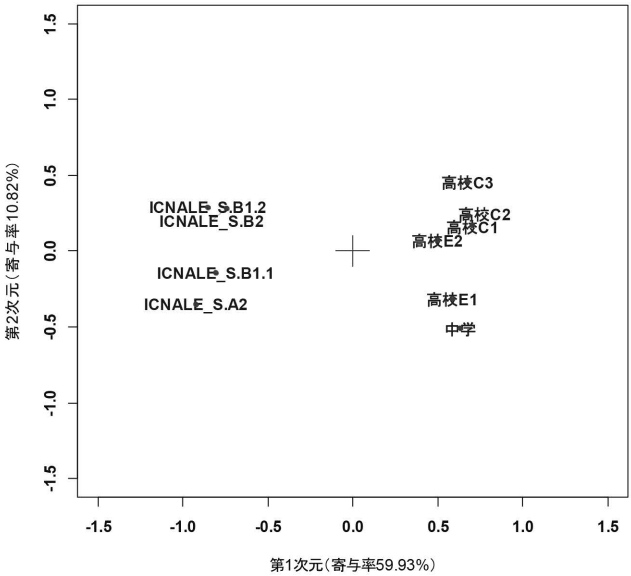

図 6. コレスポンデンス分析の結果（教科書コーパスと ICNALE-S）

4.3 RQ3：日本人英語学習者が多用する句動詞にはどのようなものがあるか

日本人英語学習者が多用する句動詞の特徴を調べるために，学習者データ全体において頻度の高い句動詞 25 項目と，各項目の教科書コーパス全体および母語話者データ全体における頻度を調査した結果を表 5 に示す。表 5 では，各コーパス・データにおける各句動詞の実度数と当該句動詞以外の全句動詞の実度数に基づいて行ったフィッシャーの正確確率検定（Fisher's exact test）の結果，$p < .05$ の水準で有意差が認められる場合には，過剰使用（＋）または過少使用（－）の情報も表示している。

学習者が多用する句動詞には，タスクの影響が大きいと思われる agree with や work in なども含まれてはいるが，教科書での頻度に比して過剰使用の傾向があるものが多い。他方，母語話者と比べると過剰使用しているものと過少使用しているものが混在していて，日本人英語学習者と母語話者の句動詞使用実態の違いがここでも浮き彫りになる。教科書と母語話者の両方に対して学習者が過剰使用している go to, live in, listen to などの句動詞は，初歩的で英語学習の初期段階から繰り返し使う練習をするものであると思われ，これらの句動詞を学習者が必要以上に繰り返し使用していると言える可能性がある。

表 5
学習者が多用する句動詞と教科書・母語話者データにおける頻度（100 万語あたり）

句動詞	学習者	教科書	教科書に対して過剰／過少使用	母語話者	母語話者に対して過剰／過少使用
go to	5,605	2,017	＋（$p < .001$）	2,900	＋（$p < .001$）
live in	1,216	562	＋（$p < .001$）	435	＋（$p < .001$）
have in	553	395	＋（$p < .001$）	1,193	－（$p < .001$）
come to	480	538		207	＋（$p < .001$）
listen to	406	349	＋（$p < .001$）	127	＋（$p < .001$）
go out	370	120	＋（$p < .001$）	498	
agree with	343	145	＋（$p < .001$）	329	
talk to	291	148	＋（$p < .001$）	286	
work in	240	105	＋（$p < .001$）	520	－（$p < .001$）
do in	217	226		530	－（$p < .001$）
play with	210	71	＋（$p < .001$）	58	＋（$p < .001$）
talk with	195	156	＋（$p = .001$）	37	＋（$p < .001$）
work at	174	38	＋（$p < .001$）	228	
live with	173	51	＋（$p < .001$）	85	＋（$p < .001$）
pick up	171	99	＋（$p < .001$）	106	＋（$p = .002$）
look at	163	519	－（$p < .001$）	254	
look for	160	194		191	
talk about	160	213		106	＋（$p = .006$）
belong to	159	88	＋（$p < .001$）	16	＋（$p < .001$）
get on	157	73	＋（$p < .001$）	85	＋（$p < .001$）
get to	153	187		414	－（$p < .001$）
wait for	148	172		42	＋（$p < .001$）
see in	139	317	－（$p < .001$）	276	－（$p = .007$）
get in	132	86	＋（$p < .001$）	414	－（$p < .001$）
stay in	119	93	＋（$p = .007$）	111	

4.4 RQ4：日本人英語学習者がうまく使えない句動詞にはどのようなものがあるか

日本人英語学習者がうまく使えない句動詞を見るために，まず教科書コーパス全体と各学習者コーパス（母語話者データを除く）全体におけるオッズ比（odds ratio）を算出した。ここでのオッズ比は，「教科書コーパスにおける特定の句動詞の実度数」を「教科書コーパスにおける当該項目以外の句動詞の実度数」で割った値と，「学習者コーパスにおける特定の句動詞の実度数」を「学習者コーパスにおける当該項目以外の句動詞の実度数」で割った値を求め，前者を後者で割ること

で得た。このオッズ比は，1 を超えて値が大きいほど教科書での使用頻度に比して学習者が当該句動詞を使用する頻度が低いということを示す。NICT JLE においてオッズ比が最も高い句動詞 20 項目を表 6 に挙げる。(ICNALE-S はコーパスサイズが小さく，学習者データに限ると 123 項目中およそ半分の 62 項目が度数ゼロであるため，ここでは NICT JLE に限って論じる。)

表 6
教科書データと NICT JLE の学習者データにおけるオッズ比(最も高い 20 項目)

句動詞	教科書コーパスにおける度数	NICT JLE における度数	オッズ比
lead to	97	1	75.48
believe in	245	9	21.31
disagree with	26	1	20.17
know as	55	3	14.24
look up	56	4	10.87
write in	125	11	8.85
show in	162	17	7.43
build in	47	5	7.30
base on	64	8	6.21
focus on	48	6	6.21
suffer from	65	9	5.61
live on	50	9	4.31
ask for	90	17	4.12
take out	57	13	3.40
return to	95	23	3.21
turn to	49	12	3.17
agree with	122	30	3.16
help to	77	19	3.15
keep on	51	13	3.05
hold in	77	20	2.99

これらの項目は，教科書でよく使われているが学習者が使えていない項目であると言える。ただし，lead to, know as(known as)，write in などの句動詞は，教科書が話すことだけでなく読み書きを含めた全てのスキルを扱っていて，特にデータサイズの大きい高校の教科書では書き言葉の占める割合が高いことを反映してオッズ比が高くなっていると考えられる。また，believe in, disagree with, look up などは，NICT JLE のインタビューテストというタスクの性質ゆえにオッズ比が高いと考えられる。しかしながら，そういったことを考慮しても，base on(based on)，focus on, suffer from, ask for, take out, return to などの話し言葉でも用いられそうな句動詞の中にも学

習者がうまく使えていない項目が含まれている点は注目に値する。

次に，母語話者データとの相関が強く，学習者がうまく使えていないと言える句動詞にはどのようなものがあるのかを見るために，図 2 のコレンスポンデンス分析の結果に注目する。NICT JLE の第 1 次元の行スコア（繁雑さを避けるために図 2 では示していない）の下位 20 項目の句動詞を表 7 に挙げる。

表 7
NICT JLE のコレスポンデンス分析の結果の第 1 次元の行スコア下位 20 項目

句動詞	スコア
end up	−2.930
look up	−2.777
live on	−2.405
come up	−2.159
take out	−2.090
disagree with	−2.069
hear of	−1.954
come out	−1.941
keep on	−1.907
focus on	−1.891
go into	−1.847
do with	−1.829
get along	−1.826
base on	−1.749
go on	−1.744
go in	−1.670
tend to	−1.663
come on	−1.600
set up	−1.589
get in	−1.558

これらの句動詞は同じタスクの中でも，母語話者が使い，日本人英語学習者があまり使わないものであると言える。基本的と考えられる項目が多いが，こういった母語話者がよく使う句動詞を学習者がうまく使えていないという事実は重く受け止める必要がある。教科書を含む教材や現場での指導の中で，こういった句動詞を実際に使う機会を学習者に与えるなどの対応を検討する必要があるかもしれない。

最後に，母語話者が高頻度で使用する句動詞と，各項目の学習者データおよび教科書コーパ

スにおける頻度を調査した結果を表 8 に示す。表 5 と同様に，各項目の実度数に基づいて行ったフィッシャーの正確確率検定の結果による，学習者による過剰使用（＋）または過少使用（－）の情報も付している。

表 8
母語話者が多用する句動詞と学習者・教科書データにおける頻度（100 万語あたり）

句動詞	母語話者	学習者	学習者が母語話者に対して過剰／過少使用	教科書
go to	2,900	5,605	＋（$p < .001$）	2,017
have in	1,193	553	－（$p < .001$）	395
believe in	705	14	－（$p < .001$）	292
do in	530	217	－（$p < .001$）	226
work in	520	240	－（$p < .001$）	105
go out	498	370		120
live in	435	1,216	＋（$p < .001$）	562
get to	414	153	－（$p < .001$）	187
get in	414	132	－（$p < .001$）	86
depend on	334	102	－（$p < .001$）	75
agree with	329	343		145
pay for	297	93	－（$p < .001$）	35
help to	297	26	－（$p < .001$）	92
talk to	286	291		148
see in	276	139	－（$p = .007$）	317
look at	254	163		519
work at	228	174		38
go into	223	35	－（$p < .001$）	89
end up	217	15	－（$p < .001$）	31
come from	212	83	－（$p < .001$）	225
come to	207	480	＋（$p < .001$）	538
deal with	196	35	－（$p < .001$）	35
look for	191	160		194
have on	191	97	－（$p = .025$）	162
go in	180	52	－（$p < .001$）	29

母語話者が高頻度で使用する句動詞の多くは，学習者の使用頻度が低いということが分かる。これらの句動詞のうち，教科書での使用頻度も低い depend on, pay for, end up などの句動詞

は，教育・教材での扱いが手薄である可能性があり，これらの句動詞を手厚く扱うことで学習者の句動詞使用状況を改善できる可能性がある。

5. まとめと今後の課題

日本人英語学習者の話し言葉における句動詞使用実態を明らかにするために，母語話者との比較と教科書における出現状況との比較を行った。その結果，学習者のレベルが上がるにつれて句動詞使用頻度は上昇し，母語話者に近づいていくものの，上級の学習者であっても使用する句動詞は母語話者とは大きく異なり，教科書で出現する句動詞と学習者が使用する句動詞にも大きな違いがあり，母語話者の使用頻度や教科書での使用頻度が高いにも関わらず学習者がうまく使えないものがあるということが明らかになった。

本研究では話し言葉に注目したが，書き言葉のデータも同様の観点で分析し，さらに話し言葉と書き言葉を詳細に比較することで，新たな知見が得られる可能性がある。また，学習者のインプットデータとしての教科書の分析に関しては，様々な要素が混在する教科書データを細分化し，各要素がどのスキル（リーディング，スピーキング，会話でのやりとりなど）を伸ばすことを目的としているのかを踏まえ，その区分に基づくサブコーパスを構築することにより，より精緻化した分析を行うことが可能になると考えられる。

今後，日本人英語学習者による句動詞の学習実態と使用実態を包括的に分析することで，日本の英語教育環境でより注意深く指導すべき句動詞項目を特定し，教育効果を高めることができると考えられる。

謝辞

本研究の一部は，JSPS 科研費「日本人英語学習者の学習・使用実態を反映した重要句動詞リストの作成」（26770201）・「多様なデータに基づく日本人英語学習者の文法項目学習・使用実態調査」（16K16882）の助成を受けたものである。本研究はまた，平成 26～27 年度成城大学特別研究助成「新課程用英語検定教科書における語彙使用状況の調査と分析」の成果を利用している。

引用文献

Biber, D., Johansson, S., Leech G., Conrad S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson Education Limited.

Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. *TESOL Quarterly*, *41*(2), 339–359.

飯尾豊（2013）「コーパスを用いた日本人学習者の句動詞の使用に関する研究」『熊本大学社会文化研究』, *11*, 35–53.

Ishii, Y. (2009). Making a list of essential phrasal verbs based on large corpora and phrasal verb dictionaries. In Y. Kawaguchi, J. Durand, & M. Minegishi (Eds.),

Corpus analysis and variation in linguistics (pp. 121–140). Amsterdam and Philadelphia: John Benjamins.

石井康毅(2017)「日本人英語学習者が学習・使用する句動詞の分析—網羅的な頻度調査に基づく考察—」統計数理研究所共同研究リポート 381『イベント・スキーマと構文に関する研究』, 1–20.

金子恵美子・和泉絵美(2013, 3 月)「学習者スピーキングデータにおける基準特性の候補—complexity, fluency, accuracy に着目して」平成 24-27 年度科学研究費補助金 基盤研究(A)「学習者コーパスによる英語 CEFR レベル基準特性の特定と活用に関する総合的研究」(代表:投野由紀夫)第 3 回公開会議. Retrieved from http://www.cefr-j.org/PDF/2012/Kaneko&Izumi_20130321.pdf

Liu, D. (2011). The most frequently used English phrasal verbs in American and British English: A multicorpus examination. *TESOL Quarterly*, *45*(4), 661–688.

前浜知味(2015)「話し言葉・書き言葉における句動詞使用—コーパスに見る日本人英語学習者と英語母語話者の比較研究—」『電子情報通信学会技術研究報告』, *115*(361), 7–12.

水本篤 (2009) 「コーパス言語学研究における多変量解析手法の比較—主成分分析 vs. コレスポンデンス分析—」統計数理研究所共同研究リポート 232 『コーパス言語研究における量的データ処理のための統計手法の概観』, 53–64.

Schmid, H. (1994). Probabilistic part-of-speech tagging using decision trees. *Proceedings of international conference on new methods in language processing*, 45–49.

Uchida, T. (2012). Use of multiword verbs by non-advanced EFL learners: Focusing on common verb + particle combinations.『コーパスに基づく言語学教育研究報告』(東京外国語大学大学院総合国際学研究院グローバル COE プログラム「コーパスに基づく言語学教育研究拠点」), *8*, 303–323.

渡辺敦司(2013, 1 月 29 日)「2013 年度高校教科書採択状況—文科省まとめ(下)」『内外教育』, *6221*, pp. 10–19.

渡辺敦司(2014, 1 月 28 日)「2014 年度高校教科書採択状況—文科省まとめ(下)」『内外教育』, *6307*, pp. 10–19.

渡辺敦司(2015, 2 月 3 日)「15 年度高校教科書採択状況—文科省まとめ(下)」『内外教育』, *6392*, pp. 6–13.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Yoshitomi, A. (2006). The use of phrasal verbs by Japanese learners of English: Implications from story telling data. In A. Yoshitomi, T. Umino, & M. Negishi (Eds.), *Studies in second language teaching and second language acquisition (Linguistic informatics V)* (pp. 201–225). Amsterdam/Philadelphia: John Benjamins.

Yoshitomi, A. (2012). The use of multi-word units in learner language narratives: Are

there qualitative and/or quantitative differences between Japanese ESL learners and EFL learners? In Y. Tono, Y. Kawaguchi, & M. Minegishi (Eds.), *Developmental and crosslinguistic perspectives in learner corpus research* (pp. 309–332). Amsterdam/Philadelphia: John Benjamins.