JSLA, Volume 9 (2017), 123–133

英語学習者コーパスを用いた基本動詞を含む定型表現の習得に関する基礎調査

Basic Study for Acquisition of Formulaic Sequences Including Verbs Using Japanese EFL Corpus

村 木 恭 子*
Kyoko Muraki

The formulaic sequences are supposed to be affective for language learning from both cognitive and fluency point of view. The aim of this study is to explore basic study of how Japanese English Learner acquire formulaic sequence which include main verb using corpus. In this study, two types of corpus had been used; JEFLL, The ICNALE. The finding showed some formulaic sequences are acquired in earlier stage but its amount is limited. This corpus study still remain some problem which are learner's specific English level, effect by the topics. It is also unclear whether if student understand the collect meaning of formulaic sequences. However findings from this study suggested that it is required to gain more learning environment including material for formulaic sequences.

キーワード：英語学習者コーパス、基本動詞、定型表現、習得

1．はじめに

英語学習を進めていく上で、学習者の目の前には様々な壁が立ちはだかる。語彙を増やしたからといって、直ちに流暢な英語が使えるわけではない。語彙をつなげれば意味が通じる英文になるとは限らず、むしろ英文は語彙が横並びに並んでいるのではなく、階層構造を成している点が学習者の理解を困難にしている一つの要因である。

語彙より大きく、文より小さい言語単位として、チャンク（定型表現，コロケーション）が存在する。言語研究におけるチャンクとは、言語を処理する上で外界から入力された情報を意味のあるまとまりのある単位にまとめる過程のことを指す（e.g. 谷口，1992; 高梨・卯城，2000）。門田・玉井（2004）では、最初から英語の文法を学ぶのではなく、まずは語の

* 桜花学園大学学芸学部非常勤講師

まとまりを意識することを身につけることによって、文法学習が効果的になると指摘している。田中 他（2003）では、このチャンクをどのように作るか、名詞・動詞・形容詞・副詞・前置詞と文法項目ごとにまとめ、効果的な学習法を提案している。投野（2004）では、コーパス（電子化された言語データ）を用いた研究データより動詞を中心とした100語の高頻度語を抽出し、コロケーション情報をもとに学習表現を提案している。

言語研究において、チャンク・コロケーションを用いた研究は、コーパスを用いた研究や心理学実験の手法を用いた研究が行われてきた（e.g. Underwood et al., 2004; Schmitt and Underwood, 2004; 村木, 2006）。中でもコーパスを用いた研究は、英語母語話者及び学習者が脳内に保持している知識を探求する方法として有効であり、定型表現の使用状況を数と種類の面から比較し、使用する定型表現の種類や数は母語話者に比べて劣るが、学習者も定型表現を使用していることが明らかになっている（e.g. DeCock et al., 1998; Milton, 1998）。

國分（2012）では、７つの基本動詞（have, take, get, make, go, come, give）を含む定型表現を大規模コーパスである BNC（The British National Corpus）及び WordBanks を用いて調査した上で、語彙指導法への有効性を検証し、語彙知識の強化・記憶の保持、及び語彙使用の再生を促したと示唆している。また、学習者の事後アンケートより基本動詞を含む定型表現の学習を通して語彙知識のサイズの強化のみでなく、語彙知識の深さを身につけることの重要性、また使用頻度を意識したことで、より自然な英語表現を発信することができたという結果を報告している。

本研究では、学習者が基本動詞＋名詞から成る定型表現をどのように使用しているかを調査することで、國分（2012）における英語母語話者の使用と比較し、学習者がどの程度基本動詞＋名詞から成る定型表現を定着しているか調査することを目的とする。

２．英語学習者コーパス

最新の英語学習者コーパスとして JEFLL（Japanese EFL Learner）Corpus 及び The ICNALE（The International Corpus Network of Asian Learners of English）を用いる。JEFLL Corpus は、約１万人の中学校１年生から高等学校３年生までの日本人英語学習者が６種類のテーマについて書いた英作文をコーパス化したものであり、約70万語から成り立つ。各々が、20分間、辞書無しで作文を書く事になっている。学習者の英語レベルについては、特に記述はない。

一方 The ICNALE は、母語話者と、アジア圏の英語学習者を対象とした会話及び作文コーパスから成り立ち、中でも作文コーパスについては2800人が２つのテーマ（「大学生のアルバイト」「レストランでの禁煙」）について書いたものである。本研究では、その中の日本人英語学習者400人が書いた800の作文コーパスを調査対象とした。学習者の英語レベルについては、TOEIC や TOEFL などの習熟度テスト及び語彙サイズテストにより Common

European Framework of Reference（CEFR）としてA2, B1−1, B1−2, B2+の4段階に分類されている。算出条件の統制として、Microsoftの Word上で作業を行い、辞書の使用は認められていないが、スペルチェックは使用することができる。また、時間は20〜40分間で、200〜300語（±10％は許容）で作成することになっている。

本研究では作文のテーマ及び、学習者のレベルについては問わずに日本人英語学習者の中高校生コーパス（JEFLL）及び、大学生コーパス（ICNALE）としてデータの比較を行った。

3．結果と分析

まず、國分（2012）の結果と比較するために、それぞれのコーパスを用いて3つの基本動詞（have, take, get）を用いて共起検索（collocation search）を行い、結果より基本動詞＋名詞の共起頻度リストを作成した（表1，2，3）。またその後の分析では、語句検索（KWIC）を用い、検索表現を中心とした前後の語彙を表示することで、文全体を通して語句使用の特徴を観察した。尚、今回は代名詞についても、名詞として結果に含めることとした。

表1　have＋名詞の共起頻度

	BNC	%	Wordbanks	%	JEFLL	%	ICNALE	%
time	16	0.18	10	0.11	73	0.93	11	0.44
effect	12	0.13			0	0	0	0
right	10	0.11	9	0.10	0	0	0	0
idea	8	0.09			0	0	0	0
lot	8	0.09	7	0.08	0	0	0	0
chance	7	0.08	8	0.09	0	0	0	0
look	7	0.08			0	0	0	0
problems	7	0.08			0	0	0	0
money	6	0.07	9	0.10	92	1.17	10	0.40
place	6	0.07			0	0	0	0
kind			9	0.11	0	0	0	0
people			8	0.09	0	0	0	0
problem			9	0.10	0	0	0	0
years			9	0.10	0	0	0	0

表1から、“have”に対しては、中高生・大学生コーパス共に“time”“money”との共起が見られ、割合についても両コーパス共に母語話者コーパスに比べて高い結果であった。

“time”に対して例文を検索したところ、中高生のコーパスからは、“I don’t have time”のような否定表現と組み合わせた使用が73件中59件と大半を占めた。また、“have time”に続く語彙についても前置詞“to”＋不定詞が続く表現が73件中29件であった。一方大学生コーパ

スでは、中高生コーパスに比べて出現頻度自体が低いものの、“We have time to do something”, “they should have time to prepare” のような表現のバラエティーが見られた。また、“have time” に続く語彙について前置詞 “to”＋不定詞は、11件中９件であった。

“money” に対しては、中高生コーパスでは、否定表現 “don’t” “didn’t” との使用が92件中45件と約半数みられ、一方大学生コーパスでは10件中３件であった。

この結果から、“have time” “have money” 共に否定表現と組み合わせた定型表現、また “have time＋to 不定詞 ” が中高生の段階で身についているのではないかと考えられる。

表２　take＋名詞の共起頻度

	BNC	%	Wordbanks	%	JEFLL	%	ICNALE	%
place	163	1.86	118	1.35	9	0.55	1	0.71
account	86	0.98	32	0.37	0	0	1	0.71
part	69	0.79	70	0.80	53	3.24	12	8.51
advantage	56	0.64	60	0.69	0	0	0	0
care	56	0.64	61	0.70	0	0	0	0
time	50	0.57	55	0.63	4	0.24	2	1.42
action	44	0.50	55	0.63	0	0	1	0.71
look	31	0.35	38	0.43	0	0	0	0
steps	22	0.25			0	0	0	0
responsibility	21	0.24	20	0.10	0	0	2	1.42
notice			19	0.22	0	0	0	0

次に、表２より “take” に対しては、両コーパス共に “part” “place” “time” との共起が見られた。特に “part” との共起が目立ち、中高生のコーパスからは “take part” に続く語として53件中50件で前置詞 “in” の使用が見られた。一方大学生のコーパスでは、12件中９件であった。また、“take part in” の前に前置詞 “to” を付けた表現 “to take part in” については、中高生コーパスで50件中14件、大学生コーパスで12件中５件見られた。さらに “take part in” の頭に否定表現 “not” を付けた表現については、中高生コーパスで50件中17件、大学生コーパスでは12件中１件であった。

この結果から、中高生の段階で、“take part in” あるいは “to take part in” という定型表現が身についていると考えられる。

次に表３より、“get” に対しては、両コーパス共に “money” との共起が見られ、母語話者コーパスよりも学習者コーパスの方がより両語の結び付きが強いとみられる。これについては、学習者コーパスがテーマを持って書いた作文であるのに対し、母語話者コーパスが新聞、学術書、手紙など多様なテキストから構成されているという違いも影響していると考える。

“get money” に続く語について注目したところ、中高生コーパスでは、前置詞 “from” が３件、

一方大学生コーパスでは、8 件、更に “by” が14件 “for” が 9 件、“to” が 7 件であった。さらに “get money” の前に前置詞 “to” を付けた表現については、中高生コーパスで22件中 4 件、大学生コーパスで108件中45件であった。

“get money” の前後に続く前置詞を含んだ定型表現については、大学生の方が中高生に比べて定着がより見られる。しかし、作文のテーマの影響や、働くということに対してより身近な世代であるという社会的な条件も影響していることが示唆される。

表 3　get＋名詞の共起頻度

	BNC	%	Wordbanks	%	JEFLL	%	ICNALE	%
job	14	0.48	25	0.23	0	0	4	0.69
money	13	0.45	21	0.19	22	1.64	108	18.59
chance	10	0.34	14	0.13	0	0	0	0
lot	8	0.28	20	0.18	0	0	0	0
people	7	0.24	24	0.22	0	0	0	0
information	6	0.21	12	0.11	2	0.15	0	0
message	6	0.21			0	0	0	0
bed	5	0.17	13	0.12	0	0	0	0
bit	5	0.17			0	0	0	0
car	5	0.17			2	0.15	0	0
things			21	0.19	0	0	0	0
time			13	0.12	0	0	0	0
help			12	0.11	0	0	0	0

さらに表1, 2, 3で示していない名詞との共起についても観察するために、中高生・大学生コーパスの基本動詞＋名詞の共起頻度それぞれ上位15語について、結果を表に示した（表4, 5, 6）。尚、それぞれのコーパスより上位15語であるため、表1, 2, 3に示した結果と一部重複している。

表 4 では、どちらのコーパスにおいても “breakfast” “smoking” のような作文のテーマに影響を受けた語彙との共起が多く、特に中高生コーパスの上位 3 語は顕著である。大学生コーパスからは、“responsibility” “experience” などの比較的テーマの影響を直接は受けず、またより難易度の高い語彙の使用が見られた。中高生コーパスからは、“nothing” “something” “anything” などの比較的テーマに影響を受けない語彙も複数見られた。

表 5 では、“money” “time” については、“have” の結果に引き続き “take” でも共起が両コーパスで見られたが、全体的に “have” に比べて名詞との共起頻度が低いことがわかる。また、“water”, “glasses”, “smoke” などのテーマを連想させるような語彙も多く見られるが、“take part” “take place” のようなテーマに依存しない定型表現の使用が、他の動詞に比べて両コーパス共に複数見られた。

表 4　have＋名詞の共起頻度　上位15語

	JEFLL	%		ICNALE	%
breakfast	999	12.67	smoking	12	0.48
bread	676	8.57	time	11	0.44
rice	652	8.27	responsibility	10	0.40
money	92	1.17	money	10	0.40
time	73	0.93	experiences	7	0.28
break	57	0.72	lunch	7	0.28
nothing	52	0.66	meals	6	0.24
lunch	47	0.60	cancer	6	0.24
something	20	0.25	experience	6	0.24
milk	17	0.22	rights	6	0.24
japanese	15	0.19	trouble	5	0.20
anything	13	0.16	dinner	4	0.16
food	13	0.16	friends	4	0.16
toast	13	0.16	lung	4	0.16
school	12	0.15	freedom	3	0.12

表 5　take＋名詞の共起頻度　上位15語

	JEFLL	%		ICNALE	%
part	53	3.24	part	12	8.51
money	26	1.59	smoke	3	2.13
pictures	19	1.16	responsibility	2	1.42
breakfast	12	0.73	time	2	1.42
nothing	12	0.73	money	1	0.71
something	12	0.73	smoking	1	0.71
bread	9	0.55	people	1	0.71
place	9	0.55	smoker	1	0.71
anything	8	0.49	nonsmoking	1	0.71
water	6	0.37	pay	1	0.71
glasses	5	0.31	action	1	0.71
this	5	0.31	harm	1	0.71
food	4	0.24	place	1	0.71
time	4	0.24	work	1	0.71
foods	3	0.18	objection	1	0.71

表 6　get＋名詞の共起頻度　上位15語

	JEFLL	%		ICNALE	%
money	22	1.64	money	108	18.59
something	9	0.67	cancer	11	1.89
foods	5	0.37	job	4	0.69
food	3	0.22	jobs	4	0.69
imformation*	3	0.22	lung	3	0.52
power	3	0.22	work	3	0.52
anything	2	0.15	something	3	0.52
bread	2	0.15	skills	3	0.52
breakfast	2	0.15	nicotine	2	0.34
car	2	0.15	chances	2	0.34
dream	2	0.15	confidence	2	0.34
everything	2	0.15	diseases	2	0.34
gole*	2	0.15	smokers	2	0.34
home	2	0.15	smoke	2	0.34
information	2	0.15	sick	2	0.34

* コーパス結果上の誤用は、そのまま残してあるため、実在しない語彙も含まれている。

表 6 より、“get” と “money” の共起が他の名詞に比べて顕著に高いことが明らかである。中高生コーパスでは、“get” 1338語中、“money” が続くのが22語、一方大学生コーパスでは “get” が580語出現し、その中で “money” が続くのが108語であった。この差は、作文のテーマの影響の 1 つと考えられる。その他の名詞については、他の動詞で示された結果と同様、テーマを連想させる語彙が多く含まれていた。

最後に、國分（2012）では示されていなかった、動詞 “make, go, come, give” についても、名詞との共起についても観察するために、両コーパスで検索し、基本動詞＋名詞の共起頻度それぞれ上位15語について、結果を表に示した（表7, 8, 9, 10）。

表 7 では、“make” との共起は、中高生コーパスからは “rice” “food” などの作文のテーマから連想される語彙も見られるが、それよりも “me” “my” などの代名詞との共起が多く見られた。一方大学生コーパスからは、“money” “smoking” などのテーマに依存する語彙が多く見られた。

表 8 では、“go” と共起する名詞は、両コーパス共に “shopping” “there” が多く見られた。“abroad” についても、割合に差が若干見られるが両コーパス共に使用が見られることから、中高生・大学生に共通した興味関心の影響であることが考えられる。

表 7　make＋名詞の共起頻度　上位15語

	JEFLL	%		ICNALE	%
me	68	6.79	money	50	11.60
my	36	3.60	friends	41	9.51
us	24	2.40	smoking	13	3.02
our	21	2.10	them	11	2.55
breakfast	18	1.80	their	10	2.32
bread	11	1.10	our	10	2.32
them	11	1.10	us	8	1.86
rice	9	0.90	people	7	1.62
foods	8	0.80	me	6	1.39
people	8	0.80	you	5	1.16
him	7	0.70	dishes	4	0.93
lunch	6	0.60	foods	3	0.70
money	6	0.60	matters	3	0.70
movie	5	0.50	mistakes	3	0.70
this	5	0.50	smokers	3	0.70

表 8　go＋名詞の共起頻度　上位15語

	JEFLL	%		ICNALE	%
back	54	3.13	shopping	18	5.20
there	51	2.96	abroad	11	3.18
shopping	32	1.86	there	9	2.60
home	28	1.62	outside	8	2.31
skiing	28	1.62	anywhere	3	0.87
anywhere	13	0.75	smoking	2	0.58
school	9	0.52	restaurants	2	0.58
sking	8	0.46	college	2	0.58
my	8	0.46	outdoors	1	0.29
fishing	7	0.41	travel	1	0.29
other	7	0.41	bowling	1	0.29
somewhere	7	0.41	everywhere	1	0.29
abroad	6	0.35	driving	1	0.29
tokyo	5	0.29	society	1	0.29
karaoke	4	0.23	lunch	1	0.29

表９　come＋名詞の共起頻度　上位15語

	JEFLL	%		ICNALE	%
here	36	5.70	there	1	0.87
our	25	3.96	people	1	0.87
home	10	1.58	home	1	0.87
my	8	1.27	bar	1	0.87
school	5	0.79	restaurants	1	0.87
me	3	0.47			
someone	2	0.32			
spring	2	0.32			
there	2	0.32			

* JEFLL では、10位以下、ICNALE では６位以下は存在しなかった。

“come” は出現回数が中高生コーパスで632回、大学生コーパスで115回みられたが、名詞との共起は全体的に少なかった。“come” 自体が名詞との共起よりも、“across” などの副詞を含む句動詞としての使用が多いためであると考える。

表10　give＋名詞の共起頻度　上位15語

	JEFLL	%		ICNALE	%
me	102	24.76	them	22	17.32
us	23	5.58	us	14	11.02
money	21	5.10	money	4	3.15
you	20	4.85	you	4	3.15
children	7	1.70	me	4	3.15
their	4	0.97	smokers	3	2.36
them	4	0.97	people	2	1.57
him	3	0.73	trouble	2	1.57
my	3	0.73	energy	1	0.79
something	2	0.49	lectures	1	0.79
this	2	0.49	time	1	0.79
your	2	0.49	harm	1	0.79
bank	1	0.24	my	1	0.79
breakfast	1	0.24	nonsmokers	1	0.79
child	1	0.24	students	1	0.79

“give” については、代名詞との共起が多く目立ち、大学生コーパスでは、中高生コーパスよりも “money” “smoker” などの作文のテーマに依存する名詞が複数含まれていた。

4．まとめ

本研究では、基本動詞を含む定型表現の習得について、2種類の日本人英語学習者コーパスを使用して横断的に観察をした。その結果、“have” については、使用も多く見られ、否定表現との組み合わせての使用や、“have time＋to 不定詞 ” が中高生の段階で身についているのではないかと示唆された。

“take” については、“part” との共起が最も多く、前置詞を含んだ “take part in”, “to take part in”、また否定表現を含む表現の使用が見られた。“take part in” が中高生の段階で身についていることを示唆するものであると考える。

“get” については、“money” との共起が見られたが、“get money” に続く語については、前置詞 “from”, “by”, “for”, “to” などの使用が見られたが、“have” “take” と比べると習得の状態はまだ浅いのではないかと考える。

“make, go, come, give” については、“money” “smoker” などの作文のテーマに依存する名詞との共起が見られるものの、特に “come” については語彙の性質上名詞との共起はあまり見られなかった。“go” については、“abroad” や “shopping” といった名詞との共起が共通して見られた。

学習者コーパスを使用した本研究では、学習者が実際に産出（使用）した表現に注目し、習得状況について検討を行ってきた。作文のテーマによる影響や、個々の学習者の習熟度がはっきりと分からないといった問題が残るため、学習者全体を通して、どの習熟度段階でどの表現が習得できているとは言い切ることはできないが、一定の指標にはなると考えられる。また、個々の品詞分類については、コーパスに既に付与された品詞情報も使用することは可能であるが、文全体の構造にも注目する必要があるため、さらに綿密な研究を行う必要がある。

今後の研究では、学習者の習熟度に合わせた分析及び文構造にも注目した分析を行うこと、さらに学習者が実際に定型表現を理解しているか語彙テストや、読みの実験など様々な視点から観察していく必要があると考える。また、学習者と同様に母語話者コーパスも同じような作文コーパスを用いることで、実際の使用を観察し、学習者との比較及び、学習者の弱点を知ることができるのではないかと考える。

謝辞

本研究において、広島大学外国語教育センターの阪上辰也先生よりご助言をいただきました。厚く御礼申し上げます。

注：

・BNC 及び WordBanks の検索結果は、國分（2012）より転記しました。

・本研究に使用したデータは、投野由起夫先生を中心として構成された構築された中高生の英作文コーパス Japanese EFL Learner (JEFLL) Corpus に基づくものです。検索ツールは、JEFLL Corpus の web 検索システム（小学館コーパスネットワーク）を利用しました。

参考文献

De Cock, S., Granger, S., Leech, G. and McEnery, T. (1998). An automated approach to the phrasicon of EFL learners. In Granger, S. (ed) *Learner English on Computer*. Addison Wesley:Longman. pp.67–79.

英辞郎 on the web　https://eow.alc.co.jp/

石川慎一郎（2015）.「The ICNALE: 国際中間言語対照分析研究のための新たな学習者コーパスの開発」『信学技報』115（361），13–18.

門田修平・玉井健（2004）.『決定版英語シャドーイング』コスモピア.

國分有穂（2012）.「日本人英語学習者の基本動詞と定型表現の学習：認知言語学及びコーパス分析に基づくアプローチの需要語彙知識への影響」『人間文化創成科学叢』（14），77–85.

JEFLL Corpus　http://scnweb.jkn21.com/JEFLL2/

Milton, J. (1998).　L1 and interlanguage corpora in the design of an electronic language learning and production environment. In Granger, S. (ed) *Learner English on Computer.* Addison Wesley:Longman. pp. 186–198.

村木恭子（2006）.「読解速度から見た日本人英語学習者の定型表現の処理に関する研究」名古屋大学大学院修士論文

Schmitt, N. and Underwood, G. (2004). Exploring the processing of formulaic sequences through a self-paced reading task. In N. Schmitt (ed.), *Formulaic Sequences*. Amsterdam: John Benjamins Publishing Company, pp. 153–172.

高梨庸雄・卯城祐司（2000）.『英語リーディング事典』研究社.

田中茂範・佐藤芳明・河原清志（2003）.『チャンク英文法』コスモピア.

谷口賢一郎（1992）.『英語のニューリーディング』大修館書店.

投野由紀夫（2004）.『NHK100語でスタート！　英会話　コーパス練習帳』NHK 出版.

The ICANE　http://language.sakura.ne.jp/icnale/index.html

Underwood, G., Schmitt, N. and Galpin, A. (2004). The eyes have it: an eye-movement study into the processing of formulaic sequences. In N. Shcmitt (ed.), Formulaic Sequences (pp. 153–172). Amsterdam: John Benjamins Publishing Company.