

日本人英語学習者の語彙サイズと コロケーション性判断の関係

羽 山 恵

1. はじめに

第二言語習得において、目標言語の語彙知識を身につけることの重要性が強調され始めたのは、1980年代と言われている (Hashemi & Gowdasiaei, 2005)。1990年代以降現在に至るまで、機械可読な大量の言語データであるコーパス (corpus) の構築とその処理技術の発達によって、より言語使用の実態を反映した辞書や学習語彙リストの編集が進められてきた (代表的なものとして、*Collins Cobuild: Learner's Dictionary*, *Collins Cobuild: English Grammar*, *Academic Word List*, *JACET8000*)。そして近年では、より深い語彙知識としてコロケーション (collocation) の重要性に対する認識が広まっている (Walker, 2011)。Sinclair (1991) に従うと、コロケーションはイディオム原理と自由選択原理の中間に位置する。つまり、単語と単語の結びつきにはその共起制約の厳しさによって、(1) イディオム (idiom), (2) コロケーション (collocation), (3) 自由結合 (free combination) というカテゴリーが存在する。イディオムがもっとも共起制約が厳しく、たとえば 'blow your own trumpet' (自慢する) というイディオムは、他の単語、たとえば 'play' では同じ意味を持たない。もっとも制約の緩い自由結合の場合は、'trumpet' という名詞に対して、動詞 'blow' が共起することも、'play' が共起することも可能である。しかしコロケーションの場合、'blow a fuse' (ヒューズを飛ばす) の 'blow' をたとえば 'chop' に置き換えることはできない (以上例は Howarth, 1998 より)。これらコロケーションの不可には規則性がなく、「任意の結びつき」である ("collocation is an arbitrary linguistic phenomenon," Lewis, 1997, p. 32) とされている。

コロケーションが任意でありながら制限を持っているということは、学習者にとって非常にやっかいなものとなる。イディオムのように個々を記憶するにはその数は膨大で、しかし目標言語で許容されないコロケーションを用いれば、学習者特有の不自然な表現とみなされてしまう。Yamashita & Jiang (2010) は、コロケーションの習得には学習者の母語が影響していると主張している。2つの単語からなるコロケーションにおいて、それら語の直訳が母語における表現としても自然である場合 (congruent collocation, たとえば 'kill animals'), 不自然な場合 (incongruent collocation, たとえば 'kill time') よりも受容性判断 (phrase-acceptability judgment) にかかる時間が短く、正解率が高くなるという。この場合、コロケーションが持つ特性が、学習の成功／不成功を決定する要因であるといえる。一方でより教育的立場からは、学習者特性とコロケーション学習の関連性を明らかにする必要があると思われる。それにより、コロケーション学習に困難を感じる学習者グループを判別することが可能になり、さらには習得へと導く効果的な指導方法の提案にも結びつくと考えられるからである。石川 (2008) は英語習熟度 (TOEIC® スコアによる) とコロケーション性判断の処理時間の関係を調べたが、本研究は学習者の知識・能力をより限定的にとらえる。ここでは学習者が知っている語彙数、つまり語彙サイズとコロケーション性判断の正確性の関係を調査する。

2. 研究の方法

2.1. 目的

本研究は、学習者の語彙サイズとコロケーション性判断の正確さの相関を調べることにより、コロケーション知識の獲得に必要な学習者要因の特定を目指すものである。

2.2. 使用データ

コロケーション性判断テストでは、72 のコロケーションを用いた。これらコロケーションは全て「形容詞＋名詞」の組み合わせで、研究協力者にとって親密性が低いという前提で科学技術論文から成るコーパス (「Web を源とした質情報付き英語科学論文コーパス」田中・柴田・富浦, 2001) より抽出した。このコーパスは英語母語話者による論文と日本人による英語論文から構成されており、当該の「形容詞＋名詞」のコロケーションは、科学技術論文に通じた英

語を母語とする英文校正者によって「自然である」または「不自然である」との判断を受けている。72 のコロケーションのうち 33 が「自然である」もの、39 が「不自然である」ものである（付録を参照のこと）。

2.3. 研究協力者

15 名の大学生が協力者として参加した。彼らは英語を専攻する 3 年生（5 名）と 4 年生（10 名）であり、第二言語習得研究を専攻している。そのため「コロケーション」とは何かを十分に理解しており、本研究においては「次のコロケーションが自然か不自然か、直感的に判定しなさい」という指示を与え、こちらの意図する通りに振る舞った。

2.4. 語彙サイズ

研究協力者の語彙サイズは、日本人英語学習者用に開発された望月テスト（望月、1998）のコンピューターアプリケーション版である PVST7K を使用した。このテストでは、レベル 1 からレベル 7 までの単語が 25 問ずつ出題され、受験者は画面に現れた和語に相当する英単語を 3 つの選択肢の中から選ぶ。正誤により受験者の推定語彙サイズが結果として表示される。全ての問題に正解すると、推定語彙サイズは 7000 語となる。

2.5. 手順

英語科学論文コーパスから予め抽出され、英語を母語とする校正者によって自然さの判定を受けたコロケーション群の中から、本研究に用いる 72 の「形容詞＋名詞」の組み合わせを選択した。コロケーション群は形容詞ごとにまとめられており、本研究では基本的に 1 つの形容詞から任意の自然なものとは自然なものを 1 つずつ選択した。ただし、形容詞によっては抽出されているコロケーションが全て不自然と判定されているものもあった。その結果、自然なものが 33 問、不自然なものが 39 問となった。また選択に際しては、科学技術分野の専門的なコロケーションと校正者に判断されたものは排除した（たとえば ‘continuous node’, ‘empty event’ など）。

協力者には、語彙サイズを測定する PVST7K のプログラムとコロケーションをリストにしたファイルをメールに添付して配布した。その際、コロケーション性の判断には直感を用い、辞書やインターネットなどで使用の正誤を調査し

ないように注意した。結果はファイルに記入し、メールにより返送してもらった。コロケーション性判断の正誤を得点化し、個々人の成績を求めた後、推定語彙サイズとの相関を求めた。本研究では、協力者の人数が少ないこと、語彙サイズの散らばりが小さいこと（下記表 1 を参照）を踏まえ、スピアマンの順位相関を用いた。

3. 結果

協力者の語彙サイズテストの記述統計量は以下表 1 の通りである。また、図 1 は 15 名の語彙サイズを得点順にグラフ化したものである。これらからわかるとおり、協力者グループは比較的高い語彙サイズの結果を得ており、またグループ内の得点の散らばりはそれほど大きくはない。

表 1 語彙サイズテストの記述統計量

人数	最大値	最小値	平均値	標準偏差
15	6330	4935	5450	419.4

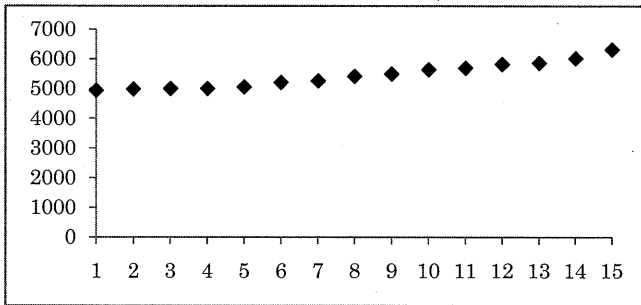


図 1 各協力者の推定語彙サイズ

一方、表 2 はコロケーション性判断テストの記述統計量、図 2 は 15 名の得点を順番にグラフ化したものである（図 1 と図 2 の番号が同一協力者を表しているわけではない）。満点が 72 点なのに対し平均値が 40.4（56.6% の正解率）というのは、決して高いとはいえない。最大値が 50 点（69.4% の正解率）ということからも、コロケーション性判断が当該の協力者たちにとって難しいものであったことがうかがえる。最小値が 32（44.4% の正解率）であることや図 2

日本人英語学習者の語彙サイズとコロケーション性判断の関係

の分布からも、得点の低い層と比較的高い層、そして平均値付近に集まっている層がいることがわかる。

表2 コロケーション性判断テストの記述統計量

人数	最大値	最小値	平均値	標準偏差
15	50 (69.4%)	32 (44.4%)	40.4 (56.6%)	4.3

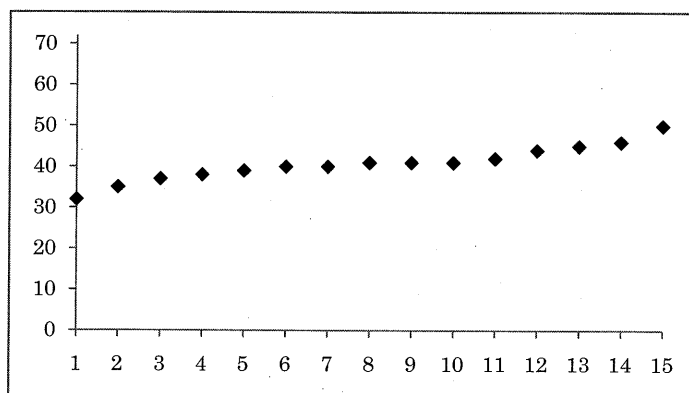


図2 各協力者のコロケーション性判断テストの得点

次に下記表3では、各協力者の語彙サイズとコロケーション性判断テストの得点、および順位を記す。スピーアマンの順位相関によると有意確率は0.36、相関係数は-0.25で、語彙サイズとコロケーション性判断テストの得点に有意な相関があるとはいえないことがわかった。語彙サイズのもっとも大きい協力者Aのコロケーション性判断得点は35点(72点満点中)で順位が15人中の14位であること、語彙サイズが5000語で12位でありながらコロケーション性判断得点は50点で1位の協力者Lなどからも、両者に正の相関がないことがわかる。一方で、語彙サイズの順位が3～6位の協力者たちがコロケーション性判断においても比較的上位(2～6位)にいることから、負の相関があるわけでもない。したがって、ここで得られた結果の限りにおいては、たくさんの単語を知っているからと言って、コロケーション性の判断が的確にできるようになるわけではないと言える。

表3 各協力者の語彙サイズとコロケーション性判断テストの得点および順位

協力者	語彙サイズ	順位	コロケーション性 判断	順位
A	6330	1	35	14
B	6020	2	37	13
C	5860	3	41	6
D	5820	4	44	4
E	5700	5	46	2
F	5640	6	41	6
G	5500	7	38	12
H	5425	8	40	9
I	5260	9	39	11
J	5220	10	42	5
K	5060	11	40	9
L	5000	12	50	1
M	5000	12	32	15
N	4980	14	45	3
O	4935	15	41	6

4. 結論

本研究は、日本人学習者(大学生)の英語語彙サイズとコロケーション性判断の相関を求めたものである。ここでは、協力者にとって馴染みのない「形容詞＋名詞」の組み合わせによるコロケーション群を用いて、コロケーションの知識ではなく、自然さ／不自然さに対する直感的な判定力を対象とした。両者の間に有意な相関が認められなかったという結果から言えるのは、知っている単語数が多いという語彙知識の広さ(vocabulary breadth)によって、コロケーションという語彙知識の深さ(vocabulary depth)に建設的な影響が起こるわけではないということだ。学習者が目標言語のコロケーションを適切に使用できるようになるには、語彙を覚えるように個々のコロケーションを覚える地道な努力を積み重ねる必要があるかもしれない。しかし本論冒頭で述べたように、膨大な数の組み合わせがあり得るコロケーションの記憶には大きな負担がかかる。やはり、コロケーション性を直感的に判断できる能力を身につけることが効果的であろう。語彙サイズの増加がその能力の習得につながらないのであれば、他の変数を候補として挙げ、その特定を試みる努力を続ける必要がある。

たとえば知っているコロケーションの数(コロケーション知識の広さ)や、個々の単語のより深い知識(類義語との区別, 用法, その単語のコロケーション例など, 多面的な単語知識の深さ)があり得る。また, コロケーション性判断の熟達を促す効果的なトレーニングを探求することも, 今後の課題となるだろう。

謝辞

本研究は科学研究費補助金(基盤研究B(一般))『Web上からの母語話者/非母語話者英語論文コーパスの作成・公開とその利用』(研究代表者: 富浦洋一(九州大学))による。

参考文献

- Hashemi, M. R. and Gowdasiaei, F. (1995) An Attribute-Treatment Interaction Study: Lexical-Set versus Semantically-Unrelated Vocabulary Instruction, *RELC* 36, 341-361.
- Howarth, P. (1998) The Phraseology of Learners' Academic Writing, in A. P. Cowie (ed.) *Phraseology: Theory, Analysis, and Applications*. Oxford: Oxford University Press.
- 石川慎一郎 (2008) 「学習者による英語語彙の反義性およびコロケーション性判断: L2 習熟度と判断時間の関係」『中部地区英語教育学会紀要』第 38 号, 55-62.
- Lewis, M. (1997) *Implementing the Lexical Approach*. Hove, England: Language Teaching Publications.
- 望月正道 (1998) 「日本人学習者のための英語語彙サイズテスト」『語学教育研究所紀要』第 12 号, 27-53.
- Sinclair, J. (1991) *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- 田中省作・柴田雅博・富浦洋一 (2011) 「Web を源とした質情報付き英語科学論文コーパスの構築法」『英語コーパス研究』第 18 号, 61-71.
- Walker, C. P. (2011) A Corpus-Based Study of the Linguistic Features and Processes Which Influence the Way Collocations Are Formed: Some Implications for the Learning of Collocations, *TESOL Quarterly* 45, 291-312.
- Yamashita, J. and Jiang, N. (2010) L1 Influence on the Acquisition of L2 Collocations: Japanese ESL Users and EFL Learners Acquiring English Collocations, *TESOL Quarterly* 44, 647-668.

付録

コロケーション一覧

不自然なコロケーション	自然なコロケーション
estimated system	estimated performance
estimated feature	estimated period
critical dynamics	critical accuracy
entire subsection	entire relation
constant time	constant period
short data	short signal
top term	top point
inner application	inner part
big amount	big element
effective information	effective combination
exact estimation	exact characteristics
negative increase	negative curve
rough result	rough information
rough locations	rough structure
natural policy	nature method
recent computer	recent concern
nearer place	near solution
close distance	
close area	
approximate estimation	approximate expression
complicated discussion	complicated program
worse change	
worst improvement	
later processing	later argument
Japanese problems	Japanese side
opposite discussions	opposite decision
precise characteristic	precise procedure
old period	old manner
more amount	more properties
remarkable impact	remarkable observation
potential possibility	potential demand
great trend	great feature
feasible tool	feasible task
major progress	major research
easiest situation	easy manner
little value	little variation
little difference	
little influence	
various information	various places