

ISSN: 2436-6647

英語コーパス学会  
大会予稿集2024

**PROCEEDINGS OF THE  
JAECS CONFERENCE 2024**



**2024.10.5**

## 目次

田畑 圭介／山縣 節子 The Movie Corpusにおける強調のdoを用いた発語内行為 —学習英和辞典への活用—	... 1
川越 涼太 英語資格・検定試験のCEFRレベル推定 —CVLAを用いたリーディングとリスニングの難易度比較—	... 7
立野 寛太 Hurfordの制約に基づく“A or B”/“B or A”の形式差 —COCAデータとR/Stanによるロジスティック回帰分析—	... 14
竹下 綾音 小学校英語検定教科書中の基本動詞の分析とその出現傾向	... 23
TAKEBAYASHI Kohei Lexical Hierarchies in L2 Vocabulary Development —A CEFR-Based Analysis —	... 30
福田 航平 日本人英語学習者によるENABLE概念の表現形式	... 37
土屋 知洋 HAVE the advantageが従える同格節that —that節中の助動詞canは必須要素か—	... 43
前田 裕司 同一学習者の技能間における産出結果の違い —スピーキングとライティングの比較から—	... 49
久屋 愛実 Than I/meを伴う比較構文における格のゆれ	... 55

岡田 晃 否定接頭辞un-付加派生語の歴史的考察 —Unkempt の事例を中心に—	... 61
飯島 真之 現代英語の態度マーカー —COREコーパスを用いた33種のレジスター別比較に基づいて—	... 66
田中 悠介／瀬戸口 彩花／近 大志／神澤 克徳 日本人英語学習者の発話におけるフィラーの生起位置と習熟度の関係性	... 73

# The Movie Corpus における強調の do を用いた発語内行為 —学習英和辞典への活用—

田畑 圭介 (新潟県立大学)

tabata@unii.ac.jp

山縣 節子 (京都外国語大学)

s\_yamaga@kufs.ac.jp

## Emphatic “do” and its illocutionary act in the Movie Corpus —Application to English-Japanese Dictionary—

TABATA Keisuke (University of Niigata Prefecture)  
YAMAGATA Setsuko (Kyoto University of Foreign Studies)

### Abstract

Based on the Movie Corpus, this study argues that a speaker’s specific intention can be articulated with emphatic “do.” Stressed functional words primarily emphasize the speaker’s assertion or fact. However, there is a certain frequency of cases where the emphatic “do” conveys blame, praise, or appreciation. A speaker’s intent in stressing a functional word is to highlight the communicated content and to get it interpreted as illocutionary politeness. This study examines the categorization of the emphatic “do” based on illocutionary acts, which ought to be incorporated into English-Japanese dictionaries.

### Keywords

強調の助動詞 do, The Movie Corpus, 発語内行為, 学習英和辞典

### 1. はじめに

機能語には音声的に強形と弱形があり、話者の意図に応じて機能語にも強勢が置かれる。強勢を伴った語は音声的に特別な響きを持つが、その音調にはどのような種類の意図が込められるか、Movie Corpus のデータをもとに考察を行う。本稿で問題にする意図については福田 (2013, p.136) が述べる「コミュニケーション上の摩擦や衝突を避け、コミュニケーションをできるだけスムーズに展開しようとする意図」と同義となる。話者が機能語に強勢を置く意図は伝達内容の強調だけでなく、発語行為的ポライトネスと解釈できる実例があることを示し、発語内行為に基づく強勢を伴う機能語の分類と、英和辞典への関連記述が有用となることを本稿の帰結として論じる。

## 2. you really did VERB が示す発語内行為

Austin(1962)以来、発語内行為は語用論の中心的研究対象に据えられている。コンテキスト内で意味づけられる性質を持つ発語内行為は、日常生活で多用され、相手に行動を促したり、自身の思いを相手に伝え、時に心理的距離と関係性を良好に維持する機能も持ち合わせる。リーチ (2020, p.18) は、感謝と謝罪を発語内行為の代表的な事例と位置づけ、thank you と sorry が極めて頻度の高いポライトネス表現だと述べている。また、リーチ (2020, p.170) は、発語内行為をポライトネスを実現するための手段とし、一般ストラテジー (General Strategy of Politeness) の行動指針を pos-, neg-ポライトネスの指標とともに体系的にまとめている。

表 1 ポライトネスに関する発話事象の領域間の関係 (発話事象の応答部分は省略)

GSP の行動指針	典型的な発話事象タイプ	pos/neg-ポライトネス	典型的に S/O 志向
気前のよさ	申し出・誘い・約束	pos-ポライトネス	S 志向
気配り	依頼, 行為指示, 懇願	neg-ポライトネス	O 志向
是認	ほめ・賞賛	pos-ポライトネス	O 志向
責務 (S から O)	感謝	pos-ポライトネス	O 志向
責務 (O から S)	謝罪	pos-ポライトネス	S 志向
同意	同意	pos-ポライトネス	O 志向
共感	祝い・お悔やみ・祈願	pos-ポライトネス	O 志向

S = Speaker O = Other

人々の発話は、ポライトネスに関わる事象と密接に関係している。そのことは、私たちの発話の多くが表 1 の発話事象のいずれかに関連していることから裏づけられる。リーチ (2020) はそれぞれの発話事象を pos-/neg-ポライトネスの例証としてまとめ、何らかの肯定的な価値を聞き手に与えようとするものは pos-ポライトネス、不作法の原因を減じようとする控えめな表現を neg-ポライトネスに分類している。リーチは、強意修飾表現、例えば really, extremely, very much, ever so much などで意味内容が強調されるものを pos-ポライトネスの発語内行為とし、ヘッジや緩衝表現 (downgraders) によってポライトネスの程度が高まるものを neg-ポライトネスに属す発語内行為と規定している。

Movie Corpus では強調の do に really が付加した、you really did VERB は計 91 例検出される。各例を見ると、話者の主張や事実を強める例が 91 例中 40 例観察される。(本稿では該当例のうち、代表例として 3 例を提示する。)

- (1) 2015 US/CA *Weepah Way for Now* I never wan na go back. I can't believe you really did fall. What was captivity like? I don't wan na go into  
 2014 US/CA *Hit by Lightning* you prayed to God that she was still out there. Wow. You really did read my profile. It got to me. And I thought, no

2014 US/CA *The Rewrite* officer and her mother is a beautician and opera singer.  
- Wow, you really did read it. - Uh-huh. So, what did you think?

(1)は検出結果が最も多い割合のグループとなることから、do の強調用法の基本用例だといえる。次に多かったものは、(2)のほめ・賞賛の発話内行為と解釈できる例で、計 18 例見つかる。

(2) 2016 US/CA *Little Boxes* You got packers? - Packers. Okay! You really did get a good job. Congratulations!

2006 US/CA *Spymate* You fix the drill... I'll let you go home. You really did make it! Now, let's talk about how you got... - the drill

1990 US/CA *Tales from the Darks*. I don't want to know. My goodness, you really did keep the best one for last. No. I saved the really, really best

強調の do は、話者の伝達内容を強める用法が基本となるが、(2)のように単なる強調用法でなく、相手をほめ、賞賛する発話となる例が *Movie Corpus* の中に見られる。(3)の発話と比較すると、(2)がほめ・賞賛の発話内行為を形成していることがわかる。

- (3) a. You got a good job.
- b. You made it.
- c. You kept the best one for last.

(3)の場合は単なる事実を述べている解釈も可能であるが、(2)の You really did VERB の表現では、ほめ・賞賛の発話内行為としか解釈できないことから、really did 自体に pos-ポライトネスの発話事象の用法が内在すると結論づけられる。純粋な強調用法は、Brown and Yule(1983)が言及する報告的(transactional)機能として、ほめ・賞賛の発話内行為は相互作用的(interactional)機能の用法と捉えることができ、情報伝達の機能と社会関係を維持する相互作用的功能に二分できる。Brown and Yule の理論では、発話はこの 2 つの機能をもとに分類されるが、誉め・賞賛の発話内行為は、相手の資質や能力を評価するのに加え、社会的立場や役割を讃える機能も備えていることから、ラポール(rapport)の維持・管理につながる表現として捉えることもできる。

強勢を伴う機能語は、話者の主張や思いを強める傾向にあるが、(2)のように really did が相互作用的に働き、調和的な人間関係を維持する役割を果たすことがコーパスデータから明らかとなる。(2)の用法が特別なものであることは、Manes and Wolfson(1981)が示す、アメリカ英語におけるほめの定型パターンの中に(2)の他動詞構文が含まれていないことから裏づけられる。Manes and Wolfson はほめの表現の 97.2%が(4)の 9 つの表現パターンのどれかに属すと論じており、このことは(2)の特異性を示している。

- (4) a. NP {is/looks} (really) ADJ (53.6%)      f. You have (a) (really) ADJ NP (2.4%)  
 b. I (really) {like/love} NP (16.1%)      g. What (a) ADJ NP! (1.6%)  
 c. PRO is (really) (a) ADJ NP (14.9%)      h. ADJ NP! (1.6%)  
 d. You V (a) (really) ADJ NP (3.3%)      i. Isn't NP ADJ! (1.0%)  
 e. You V (NP) (really) ADV (2.7%) (Adapted from Manes and Wolfson 1981 p.121)

よって、強勢を伴った機能語は、語用論的に特別な役割を持ちえ、話者が通常とは異なる特有の意図を持って使用できることになる。

you really did VERB には、(5)のように感謝の発話事象と解釈できる例が 7 例見られる。(5)も really の付加により、文の意味内容が強調されていることから、pos-ポライトネスの発話行為だとわかる。

- (5) 1982 US/CA *Diner* Merry Christmas, baby You really did treat me nice. Bought me a hi-fi for Christmas. Now I'm living in  
 2015 US/CA *After Words* to have done this to myself while I'm with you. You really did show me the best week of my life. But unfortunately... I'd already screwed  
 2010 US/CA *Ramona and Beezus* It took me this long to believe it. You really did save us, kiddo. - Whoo!

You treated me nice/You showed me the best week of my life/You saved us は必ずしも感謝を表現するとは断定できず、単に事実を述べる解釈も可能である。しかしながら(5)は感謝の意図が含意され、thank you や thanks a lot のようなルーティン化された発話とは違い、特別な感謝の響きを生み出している。(5)は really did を伴っているが、did のみでも thank you 以上に話者の思いが聞き手に伝わる表現となる。強調の do 自身が感謝の発語内行為を生み出す要因となっていることがわかる。

Aijmer(1996, p.37)は London-Lund コーパスを基に感謝のストラテジーをリスト化している。(5)はリストに示されている、Expressing appreciation of the act: That is lovely. It is appreciated.に分類できるが、当該用例と比較すると You (really) did VERB では感謝の気持ちがより強く表現されている。強勢を伴った機能語である(really) did は、話者の賞賛や感謝の思いを明確に聞き手に伝達する発話を形成していると帰結できる。

Movie Corpus における you really did VERB にはその他、(6)のように相手を非難する例(22例)が見られる。

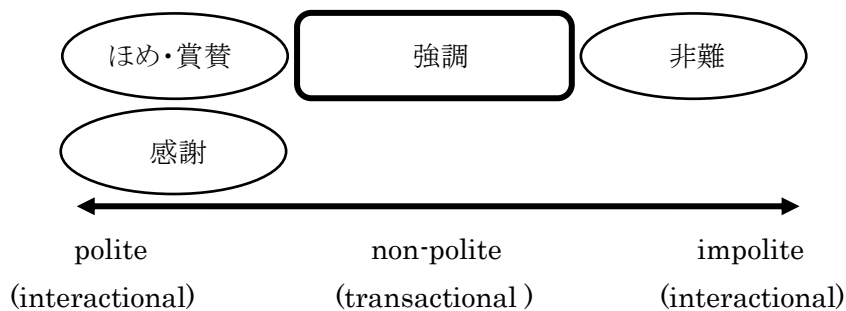
- (6) 2016 US/CA *The Dooms Chapel Horror* that... Shit. That's the one Jordan had branded on him. You really did do something terrible, didn't you? What? Mandy, what?

2009 US/CA *Wizards of Waverly* Let me use the wand or I'll tell Dad you really did take it without permission this time. Fine. Then I'll tell Dad

(6)は You should be ashamed of yourself. や That was a stupid thing to do. のような直接的な批判の表現と同質で、話者の非難の気持ちが強く響く表現となっている。(really) did は、非難を表す機能も備わっていると解釈でき、ポライトネスに関してはインポライトネスに分類される。

ここまでの議論で明らかになった強勢を伴う(really) did の機能をまとめると図 1 のようになる。

図1 強勢を伴った機能語(really) did の解釈スケール



(really) did が表す話者の意図は Movie Corpus のデータでは次の結果となる。

表 2 Movie Corpus における really did の用法

主張・事実の強調	非難	ほめ・賞賛	感謝	総計
42	22	18	7	88
47%	25%	20%	8%	

実際の検索では、really did は 91 例検出されるが、そのうち 2 例は同一のもので、2 例はインターフェイスでは表示されないことから、総数はそれらを除き 88 となる。表 2 の検索結果から、話者の主張・事実性の強調が基本となることがわかり、ついで、非難、ほめ、感謝の用法で用いられている。いずれの発語内行為も、日常表現を構成する中心的な相互活動の一つとなる。主張・事実の強調は、基本的な情報伝達の具現化であり、特別な思いを相手に伝えたい時には、非難、ほめ、感謝といった表現形態が選択されている。

### 3. 英語辞書編纂への活用

Movie Corpus では、ストーリーと当該場面に合わせて各発話が選択されており、表 2 における頻度差は、日常会話における頻度とは直接の関連性はないが、英語辞書編纂では、当該語の語法注記および用例作成に表 2 の頻度傾向と表現パターンを参考にできる。

Longman Dictionary of Contemporary English (LDOCE) (<https://www.ldoceonline>).



com/)では、強調の do は次の用例で構成されている。

(7) Longman Dictionary of Contemporary English (Online)

**used to emphasize the main verb in a sentence**

Do be careful. (強調)

You do look nice in that hat. (ほめ)

I do think she's behaved badly. (非難)

'You should have warned me.' 'But I did warn you.' (強調)

He owns, or did own (=emphasizing past tense), a yacht. (強調)

用例群には表 2 で示したパターンのうち、感謝の表現以外は含まれているので、(8)を追加することで、発話内行為のすべてのパターンを取り込むことができる。

(8) Friends did treat me better than I expected.

友人は思っていた以上に私に良くしてくれた (感謝)

Web でアクセス可能な Oxford Advanced Learner's Dictionary (<https://www.oxfordlearnersdictionaries.com/>)や Cambridge English Dictionary (<https://dictionary.cambridge.org/dictionary/english/>)での強調の do の用例はすべて強調用法で構成されている。Cambridge English Dictionary については、5 つの追加用例を表示することができるが、それらも強調の用法であるため、用例数が多くとも平板な印象を抱いてしまう。英語表現の発信面も考慮し、強調・非難・ほめ・感謝の 4 パターンすべての用例を示すことで、学習者に可能な表現形式の理解を促すことができる。ウィズダム英和辞典(三省堂)では、強調の do の項目で「話し手の主張を強調する場合と、ほめや賞賛、感謝、非難などの気持ちを表す場合がある」という注記を付記し、すでに記述のある強調・非難・ほめの用例に加えて、LDOCEと同様、(8)の感謝の用例を追記することで全 4 パターンの用例を学習者に示すことができる。

引用文献

Aijmer, K. (1996). *Conversational routines in English: Convention and creativity*.

Longman.

Austin, J. L. (1962). *How to Do Things with Words*. Harvard University Press.

Brown, G. & Yule, G. (1983) *Discourse Analysis*. Cambridge University Press.

福田一雄(2013)『対人関係の言語学: ポライトネスからの眺め』開拓社.

ジェフリー・リーチ(著), 田中典子(監修, 翻訳), 熊野真理(翻訳), 斉藤早智子(翻訳), 鈴木卓(翻訳),

津留崎毅(翻訳)(2020)『ポライトネスの語用論』研究社.

井上永幸・赤野一郎(2019)『ウィズダム英和辞典第 4 版』三省堂.

# 英語資格・検定試験の CEFR レベル推定 —CVLA を用いたリーディングとリスニングの難易度比較—

川越 涼太(九州大学 大学院生)  
kawagoe.ryota.165@s.kyushu-u.ac.jp

## Estimating the CEFR Levels of English Proficiency Tests —Comparison of Difficulty Using CVLA—

KAWAGOE Ryota (Kyushu University, Graduate Student)

### Abstract

This study investigated the characteristics and differences in the reading and listening sections of English proficiency tests using CVLA, an application that estimates the CEFR-J level of text inputs. This study focused on English tests that could potentially be used in university entrance examinations. The results of comparing the difficulty levels of both sections showed that the CEFR-J level was higher for the reading section in almost all tests, and the reading section tended to be one or two CEFR-J level above the listening section. In addition, the reading section had more words and a higher ratio of verbs per sentence compared to the listening section.

### Keywords

CEFR-J, リーディング, リスニング, 難易度比較, 英語資格試験

### 1. はじめに

これまでリーディングとリスニングの難易度について、個別では様々な研究が進められており、リーディングに関しては、大学入試や英語資格・検定試験の難易度比較をした研究が数多く存在する。しかしながら、リスニングで難易度比較をした研究は現状確認できず、リスニングスクリプトの難易度に関する研究は進んでいない。特にリーディングテキストとリスニングスクリプトの難易度の関係性を明らかにした研究は非常に少なく、現時点で確認できるのは齋藤(2022)の共通テストの比較のみである。

また、文部科学省及び大学入試センターは平成 29 年 7 月に策定した「大学入学共通テスト実施方針」に基づき、大学入学者選抜における英語4技能評価の推進のために、令和2年度から大学入試英語成績提供システムを導入することとして準備を進め(文部科学省, 2019)、令和 3 年 7 月 30 日に英語民間試験の導入断念を正式に表明した(文部科学省, 2021)。しかしながら、今後大学入試において、英検をはじめとする英語資格・検定試験が、何らかの形で大学入試における

合格判定の材料の一つとして利用される可能性は十分にある。

## 2. 先行研究

### 2.1 CEFR-J について

CEFR-J のウェブサイトによると、CEFR-J は欧州共通言語参照枠 (CEFR) をベースに、日本の英語教育での利用を目的に構築された、新しい英語能力の到達度指標である。CEFR-J の指標は、「言葉を使って何ができるか」ということを文章で明示する、can do という能力記述子 (descriptor : デスクリプタ) を用いて記述されている。CEFR-J は CEFR に準拠しつつも日本の教育環境における英語に関する枠組みに特化したもので、PreA1、A1.1、A1.2、A1.3、A2.1、A2.2、B1.1、B1.2、B2.1、B2.2、C1、C2 の 12 段階に区切られており、CEFR よりも A-B レベルを細分化した詳細な評価が可能な枠組みである (投野, 2013)。

### 2.2 CVLA について

その CEFR-J レベルを推定するオンラインアプリケーションとして、Uchida & Negishi (2018) によって開発された CEFR-based Vocabulary Level Analyzer (ver. 2.0) (以下 CVLA) がある。CVLA は入力された英文から算出した4つの統計量を CEFR レベル別に分類された英語教材のコーパスで測定した基準値と比較して、入力文の CEFR-J レベルを推定するものである (内田・根岸, 2021)。入力テキストの推定レベルは、4 つの指標 (ARI、VperSent、AvrDiff、BperA) に基づいて表示される。ARI は文字数や 1 文の単語数などから推定されるリーダビリティ指標、VperSent は 1 文に含まれる動詞の平均的な割合、AvrDiff は CEFR-J Wordlist にもとづく単語の難易度の平均、BperA は A レベルの内容語 (名詞、動詞、形容詞、副詞) に対する B レベルの内容語の割合を示している。CVLA にはリーディングテキストの CEFR-J レベルを推定する Reading Mode と、リスニングスクリプトの CEFR-J レベルを推定する Listening Mode の2つのモードが搭載されている。

畔元・内田 (2022) は CVLA における CEFR レベルごとの各指標の平均値を測定し、A1-A2 間と A2-B1 間では VperSent、B1-B2 間では AvrDiff と BperA、B2-C1 間では ARI が最も重要な変数であると報告している。また、内田・根岸 (2021) は CVLA の妥当性について検証しており、CVLA の一致率は約 53%、隣接レベルも含めた一致率は約 80% となり、400 語程度の入力があれば、大きくずれたレベルを判定することは少ないと主張している。

## 3. リサーチデザイン

### 3.1 研究目的と研究設問

本研究の目的は、入力テキストの CEFR-J レベルを推定するオンラインアプリケーションである CVLA を使用し、英語資格・検定試験のリーディングとリスニングにおける特徴と差異を明らかにすることである。そこで、以下のリサーチクエスチョン (RQ) を設定した。

RQ: 英語資格・検定試験において、リーディングとリスニングの難易度にはどのような違いがあるか。

### 3.2 分析対象

分析対象は令和 3 年度において、大学入試英語成績提供システムの参加要件を満たしていた全ての英語資格・検定試験と大学入試である。具体的には、①英検、②GTEC、③TEAP、④IELTS、⑤TOEFL、⑥TOEIC、⑦ケンブリッジ英検、新旧の大学入試である⑧大学入試センター試験と⑨大学入学共通テストの本文を対象とする。問題のテキストは過去問や公式問題集、インターネット上のサンプル問題から収集した。

リーディングは問題のテキストのみ(文法・語彙問題等を除く)、リスニングは問題に関連する音声(指示・説明の音声を除く)を対象とした。リーディングに関して、空所内に語句を挿入する問題については、正しい解答を入力し、テスト作成者が意図した本来の文章になるようにした。

### 3.3 手法

上記の英語資格・検定試験における、リーディングテキストとリスニングスクリプトの難易度について、CVLA を用いて CEFR-J レベルの観点から分析した。リーディングは Reading Mode、リスニングは Listening Mode でそれぞれ測定した。その際に、CVLA の 4 つの指標 (ARI、VperSent、AvrDiff、BperA) や延べ語数からも分析を行い、複数の指標から英語資格・検定試験の特徴と差異を明らかにした。

## 4. 結果と考察

### 4.1 RQ 英語資格・検定試験におけるリーディングとリスニングの難易度比較

#### 4.1.1 英検

表1 難易度平均(英検:2020 年第 3 回~2023 年第 3 回)

英検		サンプル数	ARI	VperSent	AvrDiff	BperA	Final
1級	Reading (2,592語)	10	C2	C2	C2	C2	C2
	Listening (3,593語)	10	A2.2	B1.2	C2	C2	C1
準1級	Reading (1,771語)	10	B2.2	C2	C2	C2	C2
	Listening (2,557語)	10	A1.3	B1.1	C1	C2	B2.1
2級	Reading (1,501語)	10	A2.2	B2.2	B2.1	B1.2	B1.2
	Listening (2,123語)	10	PreA1	A2.2	A2.1	A2.1	A2.1
準2級プラス	Reading (1,003語)	10	A2.1	B2.1	A2.2	A2.2	B1.1
	Listening (1,947語)	10	PreA1	A2.1	A2.1	A2.1	A1.3
準2級	Reading (1,166語)	10	PreA1	A2.2	A1.3	A1.3	A1.3
	Listening (1,668語)	10	PreA1	A1.2	A1.3	A1.2	A1.1
3級	Reading (739語)	10	PreA1	A1.3	A1.2	A1.2	A1.1
	Listening (991語)	10	PreA1	A1.1	PreA1	PreA1	PreA1

延べ語数はリスニングの方が多く、級が上がるにつれてその差は大きくなっている(準 2 級プラスを除く)。また、全ての級においてリーディングの方がリスニングよりも CEFR-J レベルが高い結果となった。全体として、リーディングはリスニングの1つまたは2つ上の CEFR-J レベルになる傾向が見られ、ARI や VperSent ではかなりの差が見られた。例えば、英検 1 級では、CEFR-J レベルで ARI の差が 6(C2-A2.2)、VperSent の差が4(C2-B1.2)となっている。この結果より、英検は難易度が高くなればなるほど、両者の差が激しくなると考えられる。

#### 4.1.2 GTEC

表 2 難易度平均(Basic、Advanced:第 24 回, 30 回過去問 CBT:公式問題集本試験 1, 2 回目)

GTEC		サンプル数	ARI	VperSent	AvrDiff	BperA	Final
Basic	Reading (1,871語)	2	A1.1	A1.2	A2.1	A1.3	A1.3
	Listening (1,234語)	2	PreA1	A1.3	A1.2	A1.1	A1.1
Advanced	Reading (2,994語)	2	A1.3	A2.2	B2.1	B2.1	B1.1
	Listening (1,383語)	2	PreA1	A2.1	A2.2	B1.2	A2.1
CBT	Reading (5,049語)	2	A1.3	A2.1	B2.1	B2.1	B1.1
	Listening (4,556語)	2	A2.1	B2.2	B1.2	B2.1	B1.2

全ての試験でリーディングの延べ語数がリスニングを上回っている。リーディングとリスニングで最も差が大きいのは ARI である。今回の調査において、リスニングの CEFR-J レベルがリーディングより高い結果を示したのは CBT のみであった。しかしながら、CBT は受験者のレベルによって問題が変化するため、正しい CEFR-J レベルを推定できているとは限らない。結果として、CBT を除くすべてのリーディングにおいて、リスニングよりも 1 つまたは 2 つ上の CEFR-J レベルを示した。

#### 4.1.3 TEAP

表 3 難易度平均(TEAP:見本問題)

見本問題		サンプル数	ARI	VperSent	AvrDiff	BperA	Final
TEAP	Reading (2,753語)	1	B1.2	B2.2	C1	C1	B2.2
	Listening (3,772語)	1	A1.3	A2.2	B2.1	C1	B1.2

リスニングの延べ語数が約 3,800 語とかなり多く、リーディングの延べ語数を上回っている。両者の間で最も差が大きいのは ARI と VperSent であり、リーディングはリスニングの 2 つ上の CEFR-J レベルになっている。

#### 4.1.4 IELTS

表 4 難易度平均(IELTS:無料サンプルテスト コンピューター版)

サンプルテスト		サンプル数	ARI	VperSent	AvrDiff	BperA	Final
IELTS	Reading (2,522語)	1	B2.2	C1	C2	C2	C2
	Listening (925語)	1	A2.1	B1.2	C1	C2	B2.2

リーディングの延べ語数がリスニングを上回っている。しかしながら、リスニングに関してはサンプル問題の数が少なく、実際の試験とは異なっている。最も差が大きいのは ARI であり、リーディングはリスニングの 2 つ上の CEFR-J レベルになっている。

#### 4.1.5 TOEFL

表 5 難易度平均(TOEFL:サンプル問題)

サンプル問題		サンプル数	ARI	VperSent	AvrDiff	BperA	Final
TOEFL	Reading (1,969語)	1	C1	C1	C2	C2	C2
	Listening (4,371語)	1	A1.3	B1.1	B2.1	B2.2	B1.2

リスニングとリーディングの延べ語数には非常に大きな差が見られ、リスニングの延べ語数はリーディングの倍以上となっている。最も差が多いのは ARI であり、その差は他の英検資格・検定試験よりも大きい。リーディングはリスニングの 4 つ上の CEFR-J レベルであり、両者の難易度にはかな

りの差が見られる。

#### 4.1.6 TOEIC

表 6 難易度平均 (TOEIC:公式問題集 8, 9, 10)

公式問題集		サンプル数	ARI	VperSent	AvrDiff	BperA	Final
TOEIC	Reading (3,296語)	6	A2.1	A2.1	C2	C2	B2.1
	Listening (2,888語)	6	A1.1	A2.1	B2.2	C2	B1.2

リーディングの延べ語数がリスニングを上回っている。最も差が大きいのは ARI であるが、AvrDiff にも同程度の差が見られる。リーディングはリスニングの 1 つ上の CEFR-J レベルであり、両者の難易度に大きな差はない。

#### 4.1.7 ケンブリッジ英検

表 7 難易度平均 (ケンブリッジ英検:テストサンプル)

ケンブリッジ英検		サンプル数	ARI	VperSent	AvrDiff	BperA	Final
A2	Reading (486語)	1	A1.2	C1	A2.1	A2.1	A2.2
	Listening (1,058語)	1	PreA1	A2.1	A1.2	PreA1	A1.1
B1	Reading (1,238語)	1	A2.1	C1	B1.1	B1.1	B1.1
	Listening (1,899語)	1	PreA1	B1.1	A2.2	A2.2	A2.1
B2	Reading (1,828語)	1	A2.2	C1	B2.2	B2.1	B2.1
	Listening (2,612語)	1	A1.3	B1.2	B1.2	B2.1	B1.1
C1	Reading (2,703語)	1	B2.1	C2	C2	C2	C2
	Listening (2,413語)	1	A2.1	B2.2	C1	C2	B2.2
C2	Reading (2,516語)	1	B2.1	C2	C2	C2	C2
	Listening (2,458語)	1	B1.1	C1	B2.2	C2	B2.2

A2～B2 はリスニングの方が延べ語数は多く、C1～C2 ではリーディングの方が多くなっている。また、C1 と C2の難易度はほとんど変わっていない。どのレベルでも VperSent の数値が非常に高い傾向が見られた。全体として、ARI と VperSent の差が激しく、リーディングはリスニングの1つまたは2つ上の CEFR-J レベルになっている。

#### 4.1.8 大学入試センター試験及び大学入学共通テスト

表 8 難易度平均 (センター試験:2017, 2018, 2019, 2020 共通テスト:2021, 2022, 2023, 2024)

センター試験・共通テスト		サンプル数	ARI	VperSent	AvrDiff	BperA	Final
センター試験 (本試験)	Reading (2,799語)	4	A2.1	B2.1	B2.2	B2.2	B1.2
	Listening (1,149語)	4	PreA1	A1.3	A2.2	B1.1	A2.1
共通テスト (本試験)	Reading (4,594語)	4	A1.3	A2.2	B2.1	B2.1	B1.1
	Listening (1,553語)	4	PreA1	A2.1	B1.2	B2.1	A2.2
共通テスト (追試験)	Reading (4,439語)	4	A1.3	A2.1	B2.2	B2.2	B1.1
	Listening (1,557語)	4	PreA1	A2.1	B2.1	C1	B1.1

全ての試験でリーディングの延べ語数がリスニングを上回っている。リーディングの延べ語数については、試験によって大きな差があり、共通テストはセンター試験の約 2 倍の語数となっている。リーディングはセンター試験の方が CEFR-J レベルが1つ高く、リスニングでは共通テストの方が CEFR-J レベルが 1 つまたは2つ高い結果となった。この結果より、共通テストはセンター試験と比較して、リーディングとリスニングの難易度の差が小さくなっていることが分かる。リーディングとリスニングで最も差が大きいのは、共通テストでは ARI、センター試験では VperSent であった。

また、共通テストの本試験と追試験では、リーディングとリスニングともにほとんど同じような結果が得ら

れた。つまり、少なくとも英語に関して、追試験の方が本試験よりも難易度が高いという事実はなかった。

#### 4.2 考察

表1から表8より、ほぼすべての試験でリーディングの方が CEFR-J レベルは高く、全体の CEFR-J レベルで、リーディングがリスニングの1つまたは2つ上のレベルになることが示された。また、試験によって多少の違いはあるものの、4つの指標のうち ARI や VperSent ではそれ以上の差が見られた。つまり、リーディングとリスニングの最も大きな違いは、1文の長さや複雑さであると考えられる。この結果より、たとえ同じ試験であってもリーディングとリスニングでは難易度が大きく異なっており、リスニングは1文の長さや複雑さで難易度を調節している可能性が示唆された。

#### 5. まとめ

本研究は、入力テキストの CEFR-J レベルを推定するオンラインアプリケーションである CVLA を使用し、英語資格・検定試験のリーディングとリスニングにおける特徴と差異を明らかにすることを目的として検証を行った。両者の難易度を比較した結果、ほぼすべての試験でリーディングの方が CEFR-J レベルは高く、全体の CEFR-J レベルにおいて、リーディングがリスニングの1つまたは2つ上のレベルになる傾向が見られた。また、4つの指標のうち ARI や VperSent ではそれ以上の差が確認された。

最後に、本研究の課題を2点示す。まず、テストの難易度にはタスク自体の難易度やテキストタイプ、トピック、タスクタイプなど様々な観点があるということである。CVLA はテキストの表面的な難易度を測定するものであるため、受験者が実際に感じる難易度とは異なる可能性があることを忘れてはならない。次に、リスニングとリーディングでは求められている技能が異なっており、単純に難易度を比較することはできないということである。根岸(2008)はリスニングの難易度について、「言語的難易度(語彙)」、「言語的難易度(文法・文構造)」、「トピックの馴染みの度合い」、「理解の詳細さ」、「議論の複雑さ」、「発話速度」などの要素があり、これらの要素がリスニングの難易度に恒常的に影響を及ぼすと述べている。それらを踏まえ、今後は様々な要素を組み合わせ、統合的にリーディングとリスニングの難易度について明らかにしていきたい。

#### 謝辞

本研究は JSPS 科研費 JP22H00677 の助成を受けたものです。

#### 引用文献

- 畔元里沙子・内田諭(2022)「CEFR レベル別英語教科書における基準特性の重要度—CVLA の指標を用いて—」『言語文化論究』, 48, 35-47.
- 内田諭・根岸雅史(2021)「英語読解教材の CEFR レベルの推定：CVLA の妥当性評価」『Journal of Corpus-based Lexicology Studies』, 3, 1-14.
- Uchida, S., & Negishi, M. (2018). Assigning CEFR-J levels to English texts based on textual features. In Y. Tono & H. Isahara (Eds.) *Proceedings of the 4th Asia Pacific*

*Corpus Linguistics Conference (APCLC 2018)*, 463-467.

斎藤典子(2022)「新大学入試問題のリーディング部分とリスニング部分の CEFR レベル推定による難易度比較」『日本英語教育学会・日本教育言語学会第 52 回年次研究集会』(パワーポイントスライド)。

投野由紀夫 (2013)『英語到達度指標 CEFR-J ガイドブック』大修館書店。

根岸雅史 (2008)「オーセンティック・リスニング・テキストの CEFR リスニングのレベル判断における諸問題」『ARCLE REVIEW』, 3, 100-109.

文部科学省 (2019)「令和 3 年度大学入学者選抜に係る大学入試英語成績提供システム運営大綱の廃止について (通知)」『元文科高』, 669, 1-3.

文部科学省(2021)「令和 7 年度大学入学者選抜に係る大学入学共通テスト実施大綱の予告」及び「令和 7 年度大学入学者選抜実施要項の見直しに係る予告について (通知)」『3 文科高』, 471, 1-2.



# Hurford の制約に基づく“A or B”/“B or A”の形式差 —COCA データと R/Stan によるロジスティック回帰分析—

立野 寛太(大阪大学 大学院生)  
u103982k@ecs.osaka-u.ac.jp

The difference between “A or B” and “B or A” based on Hurford's  
constraint: logistic regression using COCA and R/Stan

TATENO Kanta (Osaka University, Graduate Student)

## Abstract

This study investigates the factors influencing the order of “A or B” and “B or A” constructions in English, using data from the Corpus of Contemporary American English (COCA) and applying logistic regression analysis. Specifically, from a pragmatic perspective, the study explores English disjunction, focusing on forms that violate Hurford’s Constraint, and examines the factors affecting their occurrence through a statistical approach. The results did not provide significant evidence that the factors identified in previous research promote changes in word order; however, a significant effect was found for the parameter related to the structural position of the sentence.

## Keywords

語用論, ベイズ統計学 (Bayesian statistics), ロジスティック回帰分析, COCA

## 1. はじめに

本研究では, 英語の “A or B” もしくは “B or A” に, どのような要素が影響を与えるのかについて英語コーパスの一種である Corpus of Contemporary American English (COCA) のデータをもとにして, ロジスティック回帰を用いて分析を試みた。特に, 語用論の観点から英語の選言 (disjunction) に関して, Hurford の示した制約に違反する形式についての差が, どのような影響を受けて生起するのかを統計的アプローチに基づいて議論する。そして, 先行研究では述べられている, 形式差に大きく影響を与える要素をもとに “A or B” についての詳細で包括的な理解を目指す。

## 2. 先行研究

語用論において, 関連する語彙の意味にはスケールがあることを想定し, その意図を伝えるために適切な語を発話するというスカラー含意 (Scalar Implicature, 以下 SD) の推論が

議論の対象となっている。

- (1) a. Jack saw some horses.
- b. Jack saw a horse.

(Spector 2007: 254 を改変)

例えば「馬が 2 匹いる場面」では(1a,b)ともに状況を適切に描写していると思われるが, (1b)の文は「馬が 1 匹しかいない」という推論を導く。これは情報を可能な限り多く含まれている **some** が存在していることに起因して <some, all> のスケールが形成されており, これを可能にするオペレーター (**exh**) の役割が議論の対象になっている。

- (2) a. **exh** (John talked to Mary or Sue)  
    Equivalent to `John talked to Mary or Sue but not to both`.
- b. **exh** (some boys came)  
    Equivalent to `Some but not all boys came`.

(Fox and Spector (2018: (4))

以上のような先行研究の中で, オペレーターが文の中に埋め込まれるのかが議論の対象になっている。新グライス主義 SI 理論では, 常に発話レベルで SI が計算されるため, オペレーターが埋め込まれている例が問題となる。次節ではこのような例が英語の選言にみられることを示す。

## 2.1. Hurford の制約とその違反例

Hurford(1974)は以下のような制約を提案した。

### (3) Hurford's Constraint:

A sentence that contains a disjunctive phrase of the form 'S or T' is infelicitous if S entails T or T entails S.

この制約について, 以下のような例が不自然であることから成り立っていることが確かめられる。(2a,b) については, 「animal (動物)」の一部として「dog (犬)」が含まれている (すなわち含意関係が認められる)ことから不自然である。

- (4) a. # Mary saw a dog or an animal.
- b. # Mary saw an animal or a dog
- c. # Every girl who saw an animal or a dog talked to Jack.

(Chierchia et.al (2007: (1))

一方で、これらの制約に違反しているものの、自然であると認められる例文が存在する。(5a)では“both problems”が“first problem, second problem”を含意している。さらに(5b)では、“all”の一部には“some”が含まれていることから Hurford の提案する制約に違反している。

- (5) a. Mary solved the first problem or the second problem or both problems.
- b. Mary read some or all of the books.

(Gazdar (1979)に基づく例)

これらの例に基づいて、Gazdar (1979) は Hurford の一般化を以下のように修正して提案した。(7) のような図式化に基づく、(6) の S が“A or B”であり T が“A and B”であるとすれば、S の結合とそのスカラー含意が“A or B but not both”となるので T と矛盾する。よって(6)の一般化は(8)の適格性を予測できる。

**(6) Gazdar’s generalization:**

A sentence containing a disjunctive phrase ‘S or T’ is infelicitous if S entails T or if T entails S, unless T contradicts the conjunction of S and the implicature of S.

- (7) (A or B) or (A and B)

(Chierchia et.al (2007:(4))

- (8) a. John talked to Mary or Sue, or to both Mary and Sue.
- b. #John talked to both Mary and Sue, or to Mary or Sue.

しかしながら、Gazdar の提案する一般化にも当てはまらない例が存在している。(9a,b)では“both”が選言 or の前に生起している例であるが容認されている。

- (9) a. The students are required to do both the reading and the homework or they’re required to do one of them.
- b. Either everyone did both the reading and the homework or everyone did the reading or the homework.

また、Fox and Spector (2018) は表 1 に示すような傾向があることを示し、先行研究のような語順の変化が絶対的なものではないことを COCA のデータを用いて述べている。

表 1 COCA データ Fox and Spector (2018:21)から抜粋

	Canonical Order	Reverse Order
some or all	396	53
some or many	7	0
most or all	164	152
many or all	14	2
often or always	16	14
good or excellent	79	34

この語順の差の制約に頼らない説明が先行研究によって明らかになっている点が多いものの、これらの語順の違いが「どの要素の影響によるものなのか」を各要因の効果量に基づきパラメータを推定することで、統計的なアプローチで発展させていくことが期待できる。

### 3. 統計モデリング

#### 3.1 研究目的

本研究では先行研究に基づき、形式差に影響を与える要素について推測統計学の手法を用いて検討することを目的としている。さらに事例研究として、英語コーパスの一種である COCA によってデータを抽出しこれらの形式について着目する。

#### 3.2 データ

本研究では、文における機能を数量詞の役割に限定して、正規の語順と逆の語順が同じように出てくるデータをサンプルとして用いる。まず COCA を用いて 2024 年時点での粗頻度を以下の表 2 に示す。本研究ではパイロットスタディーとして (i) “some or all”, “all or some” と (ii) “most or all”, “all or most” の表現を用いる。

表 2 調査時における COCA データ

	Canonical Order	Reverse Order
some or all	864	113
most or all	370	298

本研究では、これらのカテゴリーにおける違いを細かな粗頻度の違いが結果に影響を与える可能性を考慮し、(i), (ii) のカテゴリーに含まれる正規語順・逆語順の中から 50 例ずつランダムサンプリングを行い、合計 200 例の中での形式差について分析を試みた。その際、箇条書きなどの特殊な形式で書かれた文章は取り除いた。そして 1 文単位（句読点から句読点まで）を抽出し、これらについて次節で述べた特徴に当てはまるかを調査した。

言語表現のバリエーションを統計的アプローチで解明することを試みる研究に山田 (2021) が挙げられる。山田 (2021)では、日本語の尊敬語のバリエーションについてグループレベルにおける固定効果をモデルに組み込むことを試みている。よってモデルの中では階層構造を想定した上で、交差分類型の混合効果ロジスティック回帰が行われていた。本研究が取り上げるデータについても、より広範囲の事例をカバーするサンプルを作成することで、このようなモデルに基づく詳細な分析を行うことは今後の課題としたい。

### 3.3. 形式差に影響を与えうるとされる要素 (変数の設定)

これらの容認度の差について、Fox and Spector (2018)をはじめとする先行研究に基づき以下のような要素が形式に影響を与えていることを仮定して統計分析を行う。

#### (A) これらの形式が生起する文中の位置について

まず、この表現が生起する文中の位置が埋め込まれているのか否かに基づいて形式が異なっていることをパラメータとして設定する。

#### (B) DE オペレーター的作用域内にあるかどうか

次に、否定環境をはじめとする下方含意 (Downward Entailing, DE) を許すオペレーターの存在が形式に影響を与えることを Fox and Spector (2018) は指摘している。“If” のような表現も DE オペレーターの一つである。(詳しくは Ladusaw (1980) を参照。)

また、以下のような例からも Hurford の制約について DE オペレーターが関与していることを示し、統計分析に用いるオペレーターの1つとして設定する。

- (10) a. \*I would go to the movies without John or Bill or both.  
b. I wouldn't go to the movies without John or Bill or both.

#### (C) Free Choice を引き起こす表現と共起しているかどうか

以下のような例では、選択の自由や任意性が、義務や命令と相入れないことから従来の選言の役割ではない特徴がみられており、語順を変化させられる。(Fox and Spector (2018))

- (11) a. # He did both the reading and the homework or he did one of them.  
b. The students are required to do both the reading and the homework or they're required to do one of them.

(D) 普遍量化詞によって量化されたコンテキストに生起しているかどうか

以下のように、Every のような普遍量化詞を含むコンテキストが語順の変化を許すことを Fox and Spector (2018) は示している。

- (12) a. \*The student solved all or some of the problems.  
 b. Every student solved all or some of the problems.

### 3.3 推測統計学：統計モデリング

#### 3.3.1 モデル式の記述

本研究では、先行研究での知見を反映し、以下の表 2 に示された変数を用いて統計モデルに組み込む。本研究では、簡易的なモデルを用いたパイロットスタディーとして位置付けているため、独立変数 x については違いに影響を及ぼさない変数と仮定した。

表 3 変数一覧

変数名	性質	レベル	内容
y	応答	事例	選択された形式を表す。Gazdar の提案に基づく正規の語順 (some or all) (= 1)か、逆の語順 (all or some)なのか(=0)に設定した。
x (A)	説明 (固定)	事例	語順の変化が発生する位置が埋め込まれた位置かどうかで、当てはまれば 1，それ以外は 0 として扱う。
x (B)	説明 (固定)	事例	取り上げる文が DE オペレーターの作用域内にあるかどうか。当てはまれば 1，それ以外は 0 として扱う。
x (C)	説明 (固定)	事例	対象とする文が義務・命令を含んでいるかどうか。当てはまれば 1，それ以外は 0。
x (D)	説明 (固定)	事例	対象とする文が普遍量化のコンテキストに生起しているかどうか。当てはまれば 1，それ以外は 0 として扱う。

(13) モデル式

$$q[i] = \text{inv\_logit}(b_1 \text{ 切片} + b_2 * A[i] + b_3 * B[i] + b_4 * C[i] + b_5 * D[i])$$

$$y \sim \text{Bernoulli}(q[i])$$

#### (14) Stan の実装 (以下, 抜粋)

```
transformed parameters{
  real q[I];
  for (i in 1:I)
    q[i] = inv_logit(b1 + b2*A[i] + b3*B[i] + b4*C[i]
    + b5*D[i]);
}
model {
  for (i in 1:I) {
    Y[i] ~ bernoulli_logit(b1 + b2*A[i] + b3*B[i] +
    b4*C[i] + b5*D[i]);
  }
}
```

#### 3.3.2 推定方法

本研究では, ベイズ推測の枠組みで (6) に含まれるパラメータの推測を行う。Stan で記述したモデルを R から呼び出し実行し, 推定を行った。(Stan のコードは馬場 (2019), 松浦 (2016) を参照。) Chain は 4 本として設定し, 各チェーンの反復回数 (イテレーション数) は 2000 回とした。その半分がウォームアップ区間として使われ, 残り 1000 回が結果に使用される。

### 4. 結果と考察

推定結果は以下ようになる。 $n_{\text{eff}}$  (有効サンプルサイズ) が十分におおきく,  $R_{\text{hat1}}$  をとっているため今回の統計において収束に問題はないと判断した。

#### 4.1 推定結果と事後分布について

推定結果は以下ようになる。

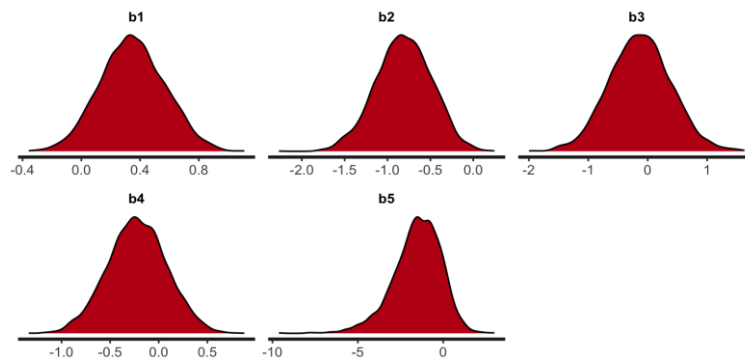
表 4 COCA に基づく推定結果の一覧 (q[2]以降は省略)

	mean	se_mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
b1	0.35	0.00	0.22	-0.07	0.20	0.34	0.50	0.78	2570	1
b2	-0.81	0.01	0.32	-1.47	-1.02	-0.81	-0.59	-0.20	3676	1
b3	-0.13	0.01	0.52	-1.13	-0.48	-0.13	0.21	0.91	3903	1
b4	-0.22	0.01	0.30	-0.83	-0.42	-0.22	-0.02	0.37	3038	1
b5	-1.53	0.02	1.41	-4.72	-2.35	-1.42	-0.55	0.87	3853	1
q[1]	0.39	0.00	0.07	0.26	0.34	0.39	0.44	0.53	4377	1

Stan の出力における信頼区間に着目すると **b2** の場合、2.5%の信頼区間が-1.47、97.5%の信頼区間が-0.20 で、範囲に 0 が含まれていない。**b2** は統計的に有意である可能性が高いと考えられる。一方で、**b3, b4, b5** に着目すると、範囲に 0 が含まれているため、統計的に有意ではない可能性が高い。

事後分布の可視化は以下ようになる。パラメータの事後分布がゼロを中心にして大きく広がっている場合、そのパラメータの影響は小さいか不明確である。一方、ゼロから離れて狭い分布の場合、そのパラメータが重要である可能性が高いと考えられる。

図 1 各パラメータにおける事後分布



#### 4.2 推定されたモデルの解釈

今回の検証では、Fox and Spector (2018) が示しているような Hurford の取り上げた制約に違反する”A or B”のような表現において、語順が “B or A” のように入れ替わる場面における決定的なパラメータは統計的に示すことが難しいことが明らかとなった。一方で、文構造の違いに基づくパラメータ (A) の違いは統計的に有意である結果となった。これは埋め込まれていない文に存在する SI を可能にするオペレーターと、埋め込まれている文に存在すると考えられているオペレーターは違う可能性が考えられる。そのオペレーターが語順を逆にする仮説は経験的・理論的側面から検証する必要がある。



## 5. まとめ

本研究では語用論の先行研究で取り上げられている“A or B”, “B or A”の形式差に影響を与えるパラメータについて、ロジスティック回帰分析を用いて検証した。その結果、先行研究で取り上げられた要因が語順の変化に影響を及ぼすことは示されず、文の構造が影響を与えている可能性が明らかになった。これはデータや理論に基づいて発展の余地が残されている。また統計分析に使用したモデルにおいては、各要因が独立していることやサンプルをとったグループごとに存在するパラメーターを仮定して階層的に分析を試みる懸念点を残しているが、今後の課題としたい。

## 引用文献

- 馬場真哉 (2019) 『ベイズ統計モデリングにおけるデータ分析入門』 講談社.
- Chierchia, G., Fox, D. & Spector, B. (2009) Hurford's Constraint and the Theory of Scalar Implicatures. In P. Égré. & G. Magri (Eds.) *Presuppositions and implicatures: Proceedings of the MIT-paris workshop*. MITWPL.
- Fox, D. & Spector, B. (2018). Economy and embedded exhaustification. *Natural Language Semantics*, 26, 1-50.
- Gazdar, G. (1979). *Pragmatics: Implicature, Presupposition and Logical Form*. Academic Press.
- Hurford, J. (1974). Exclusive or Inclusive Disjunction. *Foundations of Language*, 11, 409 – 411
- Ladusaw, W. (1980). *Polarity Sensitivity as Inherent Scope Relations*. Garland.
- 松浦健太郎 (2016) 『Wonderful R 2 : Stan と R でベイズ統計モデリング』 共立出版.
- Spector, B. (2007) Aspects of the Pragmatics of Plural Morphology: On Higher-Order Implicatures. In Uli Sauerland and Penka Stateva (Eds.) *Presupposition and implicature in compositional semantics*, chapter 9, 243–281. Palgrave Macmillan.
- 山田彬堯 (2021) 「江戸期から昭和期にかけての尊敬語のバリエーション：多項混合効果ロジスティック回帰によるベイズ統計学的アプローチ」『言語文化共同研究プロジェクト』, 5-25, 大阪大学大学言語文化研究科

## コーパス

Davies, Mark. (2008-) The Corpus of Contemporary American English (COCA). Available online at <https://www.english-corpora.org/coca/>.

# 小学校英語検定教科書中の基本動詞の分析とその出現傾向

竹下 綾音 (九州大学 大学院生)  
takeshita.ayane.760@s.kyushu-u.ac.jp

## The Analysis of Basic Verbs Appeared in Authorized English Textbooks for Elementary School Students in Japan

TAKESHITA Ayane (Kyushu University, Graduate Student)

### Abstract

The aim of this study is to examine the diversity of meanings of basic verbs presented in authorized English textbooks for elementary school students. The corpus used comprises twelve authorized English textbooks, manually annotated to classify verb meanings. The corpus was analyzed using EJWFTG (Ishikawa, 2024) to count vocabulary and parts of speech. The findings indicate that the range of basic verb meanings in the corpus is not particularly diverse, and that certain meanings for each verb tended to be presented repeatedly in authorized English textbooks for elementary school students.

### Keywords

小学校英語検定教科書, コーパス, 基本動詞

### 1. はじめに

ことばの意味の習得は、ことばと指示対象とを結びつけるのではなく、別の状況でもそのことばを使えるようになる(今井, 2024) ことである。換言すれば、語彙の意味習得は、学習者が文脈や多義性を踏まえ、適切な意味を選び使用ようになることである、と言える。しかしながら、実際の日本の英語学習者は、have=「持つ」のような英単語と日本語訳の一对一对応で英単語を学習する傾向にあり、英語の意味の広がり意識できていない (cf. 今井, 1993)。

語彙の意味の広がり理解は、外国語として英語を学ぶ学習者にとって難しい。特に、動詞 give や make 等基本動詞は、使用場面が多岐に渡り、多義性が高いゆえに L1 として英語を習得する子どもであっても習得に時間を要する (Gentner, 1978)。外国語として英語を学ぶ学習者にとっては尚更意味の習得が難しいことが想定される。Lahlou and Rahim (2023) は、多義語学習の障害として、英語学習者が学習の初期段階で単語の意味を 1 つしか習得せず、その単語には 1 つの意味しかないと思い込んでしまうことについて指摘し

ている。

2020年度から小学校5・6年生から教科としての英語が導入され、それまでは存在しなかった小学校用英語検定教科書が教室で使用されるようになった。児童への主なインプット元は、小学校英語検定教科書である可能性が高い。この教科書中の語彙記述は、日本の英語学習者が学習初期段階で、語の意味の広がり意識できるかに直結すると推察される。そのため本研究では、小学校英語検定教科書中に出現する基本動詞を分析し、出現する基本動詞の語義の多様性を明らかにすることを目的とする。これにより、小学校児童への語彙の意味指導やその後の学習段階における語彙の意味指導についての示唆を得られると考えた。

## 2. 先行研究

まず、動詞の基本的な特徴と、日本の英語学習者の動詞習得に関する先行研究を概観する。動詞は具体的なモノを表す名詞と比較すると、語と概念が一对一対応しておらず、外国語としての英語学習者にとっては、英語の動詞が持つ意味構造が第一言語と大きく異なることが (Gentner, 1982), 習得の難しさの一因となっている。

日本の英語学習者の多義動詞習得の特徴について、今井 (1993) は、日本の英語学習者と英語母語話者が持つ、動詞 **wear** に関する意味の広がり比較し、考察している。結果、日本の学習者は **wear** を日本語の「着る」と単純に対応させ、**wear** の持つ意味範囲を限定的に捉えていたと指摘している。また、動詞 **make** と **give** についても、英単語と日本語訳とが一对一対応している (Hayashi, 2008; Hayashi & Marks, 2013) という日本の英語学習者の特徴が報告されている。

次に、多義語に関し、日本の英語教育の初期段階の教材に見られる特徴についての研究を概観する。中学校英語検定教科書については、平野 (2001) や松久保 (2010) 等により、語彙の多義性の記述に関する問題点が指摘されている。一方、小学校英語教科書における語彙の多義性に関する研究は少ない。出現語彙については、後藤・眞野 (2022) により、出現する動詞の種類に限られていること、**think** や **ask** 等教室英語で多用されることが想定される動詞の提示がないという問題点が指摘されている。また、佐藤 (2021) は語彙数を分析し、どの出版社のものを使用するかにより提示される語彙数が大きく異なり、機能語の提示が不足していることを報告している。

## 3. リサーチデザイン

### 3.1 研究目的と研究設問

本研究は、小学校英語検定教科書に出現する基本動詞がどの程度多義的に使用されているか明らかにすることを目的としている。本研究の研究設問は以下の2つである。

RQ1: 小学校英語検定教科書に基本動詞はどのくらい出現するのか。

RQ2: 小学校英語検定教科書に出現する基本動詞の意味の広がりはどのくらいあるか。

### 3.2 データ

使用した小学校英語検定教科書は、『令和 5 年 4 月目録（令和 6 年度使用教科書）』に記載の小学校英語検定教科書 12 冊である。レッスン毎の見開き表紙以外の英文をテキストファイル化し、小学校英語検定教科書コーパスを作成した。総語数は 15,448 語である。

### 3.3 手法

#### 3.3.1 基本動詞の選定

本研究では、come, do, get, give, go, have, keep, let, make, put, say, see, take の 13 の動詞を基本動詞と位置づけ、語義の分類と分析を行った。これら 13 語は、Ogden (1986)に基本的な動詞として挙げられた動詞と、『小学生のための重要語リスト 1,000』(佐藤, 2017)に記載された動詞とを照らし合わせ、共通して挙げられている動詞である。

#### 3.3.2 基本動詞の語義リストの作成

各基本動詞の語義については、『英語多義ネットワーク辞典』(瀬戸ら, 2007)を主な参照元とし、この辞典に記載のなかった動詞 say の語義については『エースクラウン英和辞典第 3 版』(投野ら, 2024)を参照することで、表 1 のような基本動詞の語義一覧を作成した。

表 1 基本動詞の語義一覧の例

see の語義リスト (瀬戸ら, 2007)	
語義 0 (対象を目で見て捉える基本的な視覚作用)	a <人・目などが> 見る能力がある
	b とくに <人を> 約束 [計画] して目で見 る
	c とくに <名所・映画などを> 目で見 る
	d とくに <事件などを> 目で見 る
	e とくに <新聞など (の内容) を> 目で見 る (記事などを目で見 て理解すること まで含む)
語義 1 <人・物などを> 目で見 て確認する (自分の 目で確認する)	a <物事・状況などを> 目 で見 るよう に確 認す る
語義 2 <物事を> 目で見 るよう に経 験す る	
語義 3 <物事・状況などを> 目 で見 るよ うに 思い 描く (心 の中 で見 る)	
語義 4 <物事を> 目で見 るよ うに 理 解す る (see A as B)	
語義 5 <物事に> 目配 りす る	

#### 3.3.3 分析の手順

まず、基本動詞の語義の分類を行う。語義の分類には、基本動詞の語義一覧(表 1)を用い、筆者が手動で行った。分類の際には、『英語多義ネットワーク辞典』に記載されている例文を参照し、文脈がある場合には文脈も考慮することで、客観的な分類となるよう考慮して作業を行った。

次に、分類した語義の集計を行う。出現した語彙の数と品詞の集計は EJWFTG (石川,

2024) を用いた。品詞タグが誤っている語も見られ、それらについては筆者が確認と修正を行った。また、基本動詞直後の 4 語に関し Python 3.10.12 の spaCy3.7.5 パッケージ、辞書として en\_core\_web\_sm を使用し、その傾向を分析した。

#### 4. 結果と考察

##### 4.1 RQ1 小学校英語検定教科書に基本動詞はどのくらい出現するのか。

小学校英語検定教科書コーパス中には選定した基本動詞 13 語中 11 語、797 件の基本動詞が出現し、全体では 5.16% の出現率であった。また、6 年生用小学校英語検定教科書の方が、5 年生用小学校英語検定教科書よりも、動詞の出現率と基本動詞の出現率が共に高かった。

表 2 は、コーパス中に出現した基本動詞の頻度をまとめた表である。括弧内には出現した教科書の冊数を表記した。come, say keep は、小学校英語検定教科書 12 冊中、半数以下の教科書中でしか出現せず、基本動詞であっても教科書を通じて出会う機会は限られていることが分かる。

表 2 各学年用英語検定教科書とコーパス全体の基本動詞の出現頻度

基本動詞	5 年生用英語検定教科書 (出現教科書数)		6 年生用英語検定教科書 (出現教科書数)		コーパス全体
do	116	(6)	137	(6)	253
go	67	(6)	116	(6)	183
see	26	(6)	95	(6)	121
have	43	(6)	55	(6)	98
let	24	(5)	25	(5)	49
get	19	(3)	21	(6)	40
make	2	(2)	25	(5)	27
take	7	(3)	10	(4)	17
come	2	(1)	4	(4)	6
say	0	(0)	2	(2)	2
keep	0	(0)	1	(1)	1
総計	306		491		797

##### 4.2 RQ2 小学校英語検定教科書に出現する基本動詞の意味の広がりはどのくらいあるか。

小学校英語検定教科書コーパスに出現する語義の約半数は中心義であった。また、take 以外の基本動詞は、1 つの語義の出現割合が、出現語義の半分以上を占め、コーパス中に出現する語義は偏っている (表 3 参照) ことが明らかとなった。



るか明らかにすることを目的とし、出現する基本動詞の頻度、語義の頻度とその多様性の観点から分析を行った。その結果、小学校英語検定教科書中に出現する基本動詞は、その種類が限られ、中心義が主に提示され、基本動詞の多義性に児童が気付くには提示される中心義以外の語義の出現頻度が不十分であることが明らかになった。また、基本動詞に後続する英単語にも基本動詞ごとに偏りが見られたため、児童は、語義だけでなく基本動詞と共起する語についてもインプットが限られていると言える。これらのことから、小学校英語検定教科書を通じて、英語学習初期段階の学習者に対し、中心義以外の語義について暗示的に意味の広がり伝えることは難しいと推察される。

ただし、本研究には以下の制約と課題が存在する。基本動詞の語義分類は筆者一人で行っているため、分類の客観性に欠ける。また、小学校英語検定教科書はデジタル教材内の例文も豊富であるため、分析する必要がある。最後に、本研究で得た日本の小学校英語検定教科書中の基本動詞の多義性の特徴が EFL 教材として一般的であるのか、特有の改善すべき問題点なのかということについては、今後、他国の教材等と比較することで明らかにしていく必要がある。

本研究の結果は、小学校英語において、小学校英語検定教科書の内容から派生した、基本動詞の多義性に学習者が気づくための工夫が重要であることを示唆している。具体的には、出現頻度の低い語義を用いた例文や、未提示の用法を使った活動を取り入れること等である。分析の結果を踏まえ、日本の英語学習者が学習初期段階で基本動詞の意味の広がりにつくことができるよう、小学校教員が活用しやすいよう教材などを発案することは今後の課題である。

## 謝辞

本研究は、JSPS 科研費基盤研究 (B) 「コロケーション・チャンクを含む高頻度フレーズの CEFR レベル別リストの作成と公開」(課題番号: JP22H00677) の助成を受けたものです。

## 引用文献

- Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child development*, 49(4), 988-998.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *BBN report; no. 4854*.
- 後藤桜・眞野美穂(2022) 「小学校英語検定教科書の語彙に関する現状と今後の課題」 『鳴門教育大学学校教育研究紀要』, 36, 63-72.
- Hayashi, M., & Marks, T. (2013). On perception of the verb “give” by Japanese learners of English. 『立命館言語文化研究』, 24, 159-174.
- Hayashi, M. (2008). The acquisition of the verb “make” by Japanese learners of English.

*The annals of Gifu Shotoku Gakuen University. Faculty of Foreign Languages, 47, 25-47.*

平野洋造(2001). 「中学校教科書における多義語の扱われ方: 基本動詞に焦点を当てて」『中国地区英語教育学会研究紀要』, 31, 11-20.

今井むつみ(1993) 「外国語学習者の語彙学習における問題点 言葉の意味表象の見地から」『教育心理学研究』, 41(3), 243-253.

今井むつみ(2024) 「語彙習得理論は何を説明しなければならないのか: 30年の軌跡を振り返って」『認知科学』, 31(1), 8-26.

石川慎一郎(2024) 「『森を見ながら木を見る』コーパス研究の意義: 複数テキストから統合語彙頻度表を作成する EJWFTG の開発」『統計数理研究所共同研究レポート』, 469, 95-122.

Lahlou, H., & Rahim, H. A. (2023). The Inclusion of Polysemes in Non-native English Textbooks: A Corpus-based Study. *Arab World English Journal, 14* (2) 19-29.

松久保暁子(2010) 「英語基本動詞の提示順序と方法: 中学校英語検定教科書における効果的な語彙指導」『桜美林論考. 言語文化研究』, 1, 19-38.

Ogden C. K. (1968). *Basic English: International second language*. Harcourt, Brace & World.

佐藤剛(2021) 「小学生のための受容語彙リストの開発」『小学校英語教育学会誌』, 21(1), 54-69.

投野由紀夫 (2024). (編著) 『エースクラウン英和辞典 第3版』三省堂.

瀬戸賢一 (2007). (編著) 『英語多義ネットワーク辞典』小学館.

<sup>1</sup> NEW HORIZON Elementary English Course 5・6(東京書籍), Junior Sunshine 5・6(開隆堂), CROWN Jr. 5・6(三省堂), ONE WORLD Smiles 5・6(教科書出版), Here We Go ! 5・6(光村図書), Blue Sky elementary 5・6(啓林館)(順不同)



# Lexical Hierarchies in L2 Vocabulary Development —A CEFR-Based Analysis—

TAKEBAYASHI Kohei (Tokyo University of Foreign Studies,  
Graduate Student)

takebayashi.kohei.v0@tufs.ac.jp

## Abstract

This study explores the use of A-level single-word nouns, as defined by the CEFR, as hypernyms and their higher-level synonyms as hyponyms. For example, "food" is a hypernym of "nutrition" and "sustenance." Drawing on Crossley's (2013) exploration of lexical hierarchies, the study argues that L2 learners' vocabulary can support hierarchical lexical networks in the L2 lexicon. Single-word nouns from CEFR-aligned wordlists were matched with synonyms from the Oxford Thesaurus of English to examine semantic correspondence. Results show that 148 pre-A1 nouns link to 2,534 counterparts, comprising 20% of relevant nouns, while nouns up to B1+ have 6,943 counterparts, accounting for 54%.

## Keywords

lexical development, lexical network models, lexical proficiency

## 1. Introduction

Words are more easily recognized when integrated into a learner's mental lexicon, particularly when existing lexical networks are robust (Elgort, 2020). Beyond vocabulary size (breadth), the depth of lexical knowledge—understanding polysemy, collocations, and word relations—is crucial for advanced language. Studies show that L2 learners develop lexical competence, though their lexical networks are less hierarchical than those of L1 speakers (Crossley, 2013). This study explores how L2 learners can deepen their vocabulary by organizing A-level CEFR nouns as hypernyms and linking them to higher-level synonyms (B and C levels) using cognitive linguistic frameworks.

## 2. Literature Review

Research shows that L1 speakers process lexical relations bidirectionally, recognizing both hyponym-to-hypernym and hypernym-to-hyponym pairs, while L2 speakers process only in one direction (Crossley, 2013). This highlights the need to teach hierarchical relations in L2 learning. Rosch (1978) introduced cognitive economy and

perceived world structure as key principles in categorization, with basic-level categories (e.g., “dog”) being more efficient. Taylor (2008) adds that central prototypes, like a robin for "bird," shape how categories are processed, influencing linguistic understanding.

### 2.3 Current Study

The literature highlights the importance of hierarchical lexical networks and prototypes in language processing. This study explores how L2 learners can deepen their vocabulary knowledge by establishing semantic relations between words across lexical levels. By treating A-level CEFR vocabulary as hypernyms (or prototypes), the study aims to determine how pedagogically relevant items can be linked to these hypernyms to enrich L2 learners' lexical depth. The central research question asks:

1. To what extent can A-level CEFR vocabulary items serve as hypernyms, connecting higher-level vocabulary items to enhance learners' lexical networks?

## 3. Method

The study employed a three-step approach. First, CEFR-aligned wordlists were assembled to identify pedagogically relevant English nouns. Second, these nouns were used as search words to collect their semantically related synonyms from the Oxford Thesaurus of English (OTE). Finally, vocabulary levels were assigned to the relevant nouns to examine semantic relationships between A-level and B/C-level nouns as defined by the CEFR.

### 3.1 Wordlists and Vocabulary Levels

Three wordlists were used to identify relevant English nouns: *the Global Scale of English* (GSE), *the English Vocabulary Profile* (EVP), and *the CEFR-J Wordlist*.

- The GSE Vocabulary is a structured list for adult learners, categorizing English words by proficiency levels and aligning language learning with specific communication goals (Benigno & De Jong, 2017).
- The EVP catalogs common words and phrases acquired by learners, providing lexicographic insights based on the Cambridge English and Learner Corpora (Capel, 2015).
- The CEFR-J Wordlist adapts the CEFR for English education in Japan, featuring 7,823 items spanning levels A1 to B2, based on textbook data from China, South Korea, and Taiwan (Negishi et al., 2014; Tono, 2013).

The English nouns from these wordlists were compiled into one comprehensive list. Duplicates were removed by keeping the least difficult level assigned to each noun, prioritizing GSE over EVP and CEFR-J. This yielded 12,939 relevant nouns for analysis.

### 3.2 Collecting Synonyms from the OTE

The 12,939 nouns were used to search for synonyms in the OTE. Of these, 6,203 had synonyms, which were separated into single-word and multi-word groups using Python. This study focused on single-word synonyms, and the data was processed in R to create a comprehensive Noun List with vocabulary levels for each noun and its synonyms.

### 3.3 Counting Synonym Types

Using R, the study grouped nouns and their synonyms by vocabulary levels, comparing the number of synonyms in each level to the total nouns in the Noun List (12,939). This allowed for the calculation of synonym representation at each level, particularly focusing on elementary-level nouns.

### 3.4 Determining Basic-Level Categorization

To identify basic-level category items, the study referenced the defining vocabulary from the Longman Dictionary of Contemporary English and frequency data from COCA (Davies, 2008). Nouns were classified based on frequency (10 occurrences per million words or higher) and spoken-to-written usage ratio (75% or less), excluding C-level items. This ensured the selection of frequently used, highly recognizable nouns at the basic level.

## 4. Results

### 4.2 The Number of Synonyms Corresponding to Search Words

The compilation of a comprehensive list of nouns and their synonyms, each assigned vocabulary levels, resulted in over 145,000 word pairs. Ninety-six homonyms (e.g., *bank*, *capital*, etc.) were treated as distinct entries with unique vocabulary levels. Some search words lacked synonyms in the OTE but appeared as synonyms for others, leading to minor data gaps. To address this, the order of entries was reversed, listing synonyms before search words, which expanded the list to over 290,000 entries (Noun List) as shown in Table 1. By specifying vocabulary levels in the "WORD" column, synonyms were extracted, duplicates removed, and unique nouns with their levels were recorded. The process revealed an increase in the number of corresponding synonyms as the vocabulary levels advanced from pre-A1 to B2+ using all available senses, as

illustrated in Figure 1.

Table 1 *A Comprehensive List of Target Words and Their Synonyms (The Noun List)*

Entry	CEFR	WORD	MEANING	SYNONYM	CEFR
1	<A1	address	1 >>>	directions	A2+
2	<A1	address	1 >>>	mark	A2+
⋮	⋮	⋮	⋮	⋮	⋮
292360	Unlisted	zombie	<<< 3	jerk	B2+
292361	Unlisted	zygote	<<< 1	egg	<A1

Figure 1 *Number of Synonyms Corresponding to All Senses of Search Words*

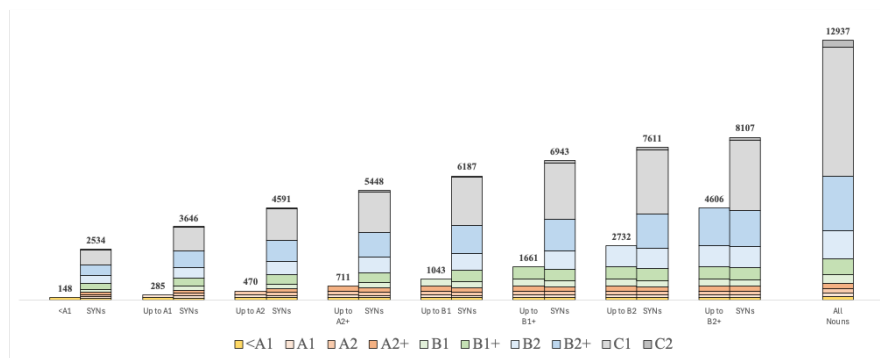
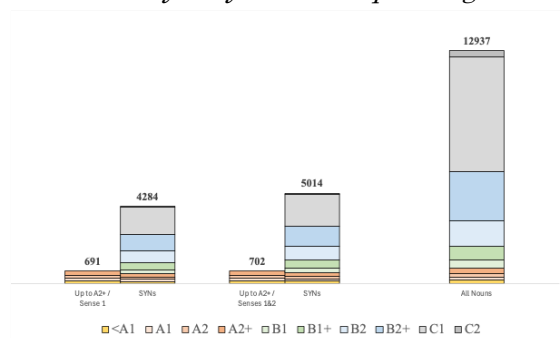


Figure 1 shows the increase in unique synonyms as vocabulary levels progress from pre-A1 to B2+. At pre-A1, 148 nouns correspond to 2,534 unique synonyms (20% of pedagogically relevant nouns), while at B1+, 1,661 nouns correspond to 6,943 unique synonyms (54%). The figure reflects form-based correspondence between search words and synonyms, without considering their semantic nuances.

#### 4.2 The Number of Synonyms Corresponding to A-level Nouns

Figure 2 shows the correspondence between A-level nouns and their synonyms, with senses capped at one or two. Around 40% of pedagogically relevant nouns can be linked to at least one A-level item. The analysis accounts for a maximum of two senses per noun to maintain practicality.

Figure 2 *Number of Synonyms Corresponding to A-level Nouns*



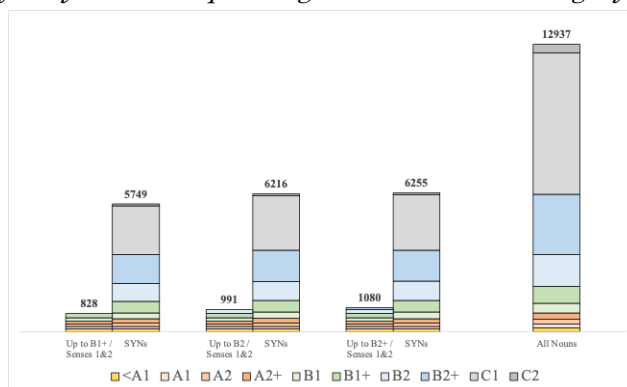
### 4.3 The Number of Synonyms Corresponding to Basic-Level Category Items

Table 2 shows the composition of basic-level category items, selected based on frequency and written usage criteria, with senses capped at two. Figure 3 indicates that basic-level items make up about half of all pedagogically relevant nouns as their synonyms. The 89 B2+ items contribute minimally to synonym counts, suggesting that items beyond B2 can be excluded from basic-level categorization.

Table 2 *The Composition of Basic-Level Category Items (Nouns)*

Basic Nouns	<A1	A1	A2	A2+	B1	B1+	B2	B2+	Total
	128	102	115	135	140	208	163	89	1080

Figure 3 *Number of Synonyms Corresponding to Basic-Level Category Items (Nouns)*



## 5. Discussion

The study suggests that learners can move beyond rote learning of form-meaning pairs by establishing hierarchical relationships between vocabulary items. For instance, understanding *food* as a hypernym for terms like *nutrition* and *sustenance* helps learners integrate new words into their mental lexicon. Approximately 50% of pedagogically relevant nouns have direct relationships with basic-level items, while the remaining require strategies like partonomy, where terms are linked through part-whole relationships (e.g., *engine* with *carburettor*, *ignition*, etc.). Using digital tools such as

spreadsheets or interactive apps can enhance learners' ability to visualize these relationships, helping them build stronger vocabulary networks. Visual aids, such as emojis, can represent semantic connections, making vocabulary more accessible. Exercises that encourage tracing advanced words back to simpler forms, such as linking *cacophonous din* to *noisy noise*, should deepen learners' understanding and retention of vocabulary. Prioritizing basic-level category items within each proficiency level can further support the development of a robust and interconnected lexicon.

This study focused solely on nouns, limiting its generalizability to other parts of speech like adjectives and verbs. It also broadly addressed polysemy, potentially oversimplifying word relationships, and relied on subjective basic-level categorization criteria. Expanding the dataset and including more parts of speech in future research could provide a more comprehensive understanding of lexical networks.

## 6. Conclusion

In conclusion, integrating hierarchical lexical networks into L2 vocabulary instruction should enhance learners' ability to connect new and existing vocabulary, facilitating deeper understanding and long-term retention. Using basic-level category items as anchors, supported by educational technology, can promote more effective and autonomous language learning, ultimately broadening learners' lexical proficiency.

## Bibliography

- Benigno, V., & De Jong, J. (2017). Developing the GSE Vocabulary. *Global scale of English research series*.
- Capel, A. (2015). The English vocabulary profile. *English profile in practice*, 5(1), 9-27.
- Crossley, S. A. (2013). Assessing automatic processing of hypernymic relations in first language speakers and advanced second language learners: A semantic priming approach. *The Mental Lexicon*, 8(1), 96-116.
- Davies, M. (2008). *The Corpus of Contemporary American English (COCA)*. Available online at <https://www.english-corpora.org/coca/>
- Elgort, I. (2020). Building vocabulary knowledge from and for reading: Improving lexical quality. In *Vocabulary and the Four Skills* (pp. 73-90). Routledge.
- Negishi, M., & Tono, Y. (2014, April). An update on the CEFR-J project and its impact on English language education in Japan. In *Language Assessment for Multilingualism, Proceedings of the ALTE Paris Conference* (pp. 113-133).
- Rosch, E. (1978). Principles of categorisation. In *Cognition and categorisation* (pp. 27-48). Routledge.

Taylor, J. R. (2008). Prototypes in cognitive linguistics. In *Handbook of cognitive linguistics and second language acquisition* (pp. 49-75). Routledge.

# 日本人英語学習者による ENABLE 概念の表現形式

福田 航平(東京外国語大学 大学院生)  
fukuda.kohei.r0@tufs.ac.jp

## Encoding the ENABLE Concept by Japanese Learners of English FUKUDA Kohei (Tokyo University of Foreign Studies, Graduate Student)

### Abstract

This research explores how Japanese learners of English and native English speakers encode the ENABLE concept, and explains the observed differences in terms of event construal. The data collected from the International Corpus Network of Asian Learners of English (ICNALE) reveal that native English speakers use the accusative-with-infinitive construction significantly more frequently, whereas Japanese learners use prepositions or conjunctions significantly more frequently. This research argues that this difference can be attributed to the tendency of Japanese speakers to recognize thematic relations as conceptually autonomous.

### Keywords

Crosslinguistic Influence, Learner Corpus, Causation, Enablement, Event Construal

## 1. はじめに

ENABLE 概念とは、例文(1)のように、使役主(the new technology)によって、被使役主(users)の行為や状態変化が可能になることを表す。

(1) The new technology allows users to choose exactly what they watch and when.

本稿では、日本人英語学習者が ENABLE 概念を英語でどう表現しているか、学習者コーパスを用いた調査により定量的に示し、その要因を、日英語の事態解釈の差異の観点から考察する。

## 2. 先行研究

### 2.1 使役概念の分類と不定詞付き対格構文

Wolff and Song(2003)は、使役主と被使役主の力関係(Force-dynamics)の観点から、使役概念を CAUSE・ENABLE・PREVENT・DESPITE の4つに分類している(後者2つは本稿では省略)。これは3つの基準、すなわち(a)被使役主が事象実現に向かう傾向を内在的に有しているか、(b)使役主と被使役主の力の方向が一致しているか、(c)結果的に事象実現に向かうか、という観点で分類している(表1)。例文(2a-b)は CAUSE 概念と ENABLE 概念の例文である。



(2) a. The blast caused the boat to heel. [CAUSE]

b. The vitamin B complex enables the body to make full use of the food. [ENABLE]

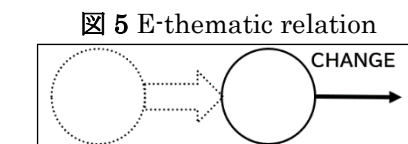
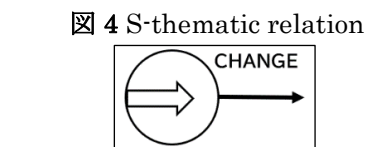
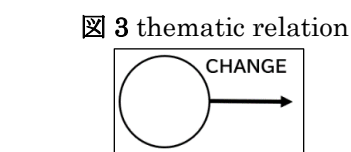
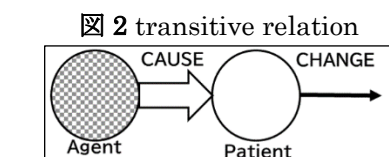
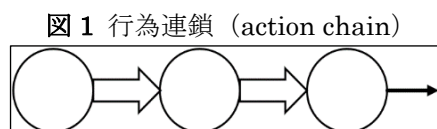
CAUSE と ENABLE の概念は、英語では(2a-b)のような、[V NP (to) V<sub>inf</sub>]の形式を持つ不定詞付き対格構文 (the accusative-with-infinitive construction; 以下 AwI 構文と表記) で表現することができる。AwI 構文で使用される動詞は make, force, cause のように CAUSE を表すものと、allow, let, enable, help のように ENABLE を表すものに分類することができる。

表 1 使役概念の分類

	(a) 被使役主の事象実現への傾向	(b) 使役主と被使役主の一致	(c) 事象実現に向かうか
CAUSE	なし	なし	向かう
ENABLE	あり	あり	向かう

## 2.2 日英語の事象解釈の差異

本節では、Langacker (1991) による事象概念の分類と、谷口 (2005) による日英語の事象解釈 (event construal) の違いについて述べる。Langacker (1991) は、事態が「個体である参加者 (participant) が相互作用を行うことによって成立する」とみなし、action chain (行為連鎖) という事態認知モデルを提唱している (図 1)。この行為連鎖モデルに基づいた事態概念の分類として、Agent が Patient の変化 (CHANGE) を引き起こす (CAUSE),



他動的な事態 (e.g. John broke the window.) を transitive relation (図 2) と呼び、ある単一の参加者が変化する自動的事態 (e.g. The ice melted.) を thematic relation (図 3) と呼んでいる。thematic relation はさらに 2 種類に分けられる。(a) 参加者自身がエネルギーを発して変化を起こしている場合が Self-induced thematic relations (以下 S-thematic relation) (図 4), (b) 外部の参加者からエネルギーを受けている場合が Externally-driven thematic relation (以下 E-thematic relation) (図 5) である。

この事態概念のうち E-thematic relation は背景に CAUSE を含意しているため (図 5 の破線部), transitive relation にも thematic relation にも解釈可能であるという曖昧性がある。谷口 (2005) は、概念的自律性の違いが E-thematic relation の記号化に影響を与えると述べている。E-thematic relation の参加者の CHANGE の部分 (図 5 の破線部を除いた部分) に概念的自律性が認められれば、thematic relation であると解釈され自動詞で記号化される。一方で、参加者の CHANGE の部分に概念的自律性が認められなければ、transitive relation として解釈され他動詞で記号化される。

この thematic relation の自律性が言語によって異なり、E-thematic relation の記号化も言語によって揺れが生じると考えられる。谷口 (2005) によれば、日英では特にその相違が対極的であると

述べ、以下のように日英語を特徴づけている：

- (3) a. 英語では、transitive relation が自律的な基本単位である。そのため E-thematic relation は transitive relation として解釈されやすい。  
 b. 日本語では、thematic relation が自律的な基本単位である。そのため E-thematic relation は thematic relation として解釈されやすく、CAUSE + thematic relation に分割した解釈をする傾向がある。

谷口 (2005) はこの特徴づけから、日英語の記号化の相違を広範に論じている。本稿ではこの (3a-b) の観点から、日本語母語話者による ENABLE 概念の記号化の傾向を考察する。

### 3. 方法

#### 3.1 研究目的と研究設問

本研究では、ENABLE 概念を記号化する際に用いられる形式である AwI 構文を、日本人学習者がどのように使用しているか、主節に生起する動詞の使用傾向から調査し、それを母語話者の使用傾向と比較する。その上で、ENABLE 概念を表す際に、AwI 構文に限らず、どのような形式を用いているか調査し、英語母語話者と中国人英語学習者の使用傾向と比較する。そして、日本人英語学習者による ENABLE 概念記号化の傾向を生み出す要因を、2.2 節で述べた、谷口 (2005) の日英の事態解釈の差の観点から考察する。以下研究設問である。

- RQ1: 日本人英語学習者が AwI 構文の主節動詞として使用しやすい／しにくい動詞は何か。CAUSE・ENABLE 概念を表す動詞を母語話者と比べてどの程度使用しているか。  
 RQ2: 日本人英語学習者は ENABLE 概念を英語で表現する際に、どのような表現形式を用いるか。そこに英語母語話者や中国人英語学習者との差は見られるか。

#### 3.2 手法

##### 3.2.1 AwI 構文の調査

表 2 AwI 構文の頻度

	素頻度	10 万語あたり頻度	タイプ頻度
The International Corpus Network of Asian Learners of English (ICNALE) Written Essays (v.2.6) (Ishikawa, 2023)	JPN 287	162.54	31
	ENS 239	256.72	28

(以下 ICNALE-WE) の日本人サブコーパス(JPN)と英語母語話者サブコーパス(ENS) 全てに目を通し、手作業で AwI 構文を抽出した。その上で、AwI 構文の主節に使用されている動詞をサブコーパス別に集計し比較した。表 2 は抽出された AwI 構文全体の頻度である。

##### 3.2.2 ENABLE 概念の表現形式調査

表 3 ENABLE 概念を表していると判断された文

ICNALE-WE の日本人 (JPN), 英語母語 (ENS), 中国人 (CHN) サブコーパスのうち、「アルバイト」をトピックとしたテキスト (PTJ) から、CEFR レベルごとにランダムに 450 文ずつ抽出した (B2 は 450 文に満たないので全て。ENS のテキストは全文)。そして、筆者を含む、英語教育学分野の修士号を持つ日本人 3 人で個別に、ENABLE 概念を表しているか、文ごとにアノテーションをした。今回は 3 人中 2 人が ENABLE 概念を表していると判断した文を分析対象とした。表 3 は ENABLE 概念を表している	A2	B1.1	B1.2	B2	合計
JPN	40	52	53	22	167
CHN	51	58	71	29	209
ENS		—			172

と判断された文数である。

#### 4. 結果と考察

##### 4.1 RQ1: AwI 構文の使用傾向

表 4 は JPN と ENS サブコーパスから AwI 構文を抽出し、主節動詞ごとに頻度を集計したものである(紙幅の都合上、合計頻度で上位 5 位まで)。JPN で最も頻度が高いのは **make** で、ENS は **allow** であった。**make** は CAUSE 概念を表す動詞であり、**allow** は ENABLE 概念を表す動詞である。表 5 は AwI 構文で使用された動詞のうち CAUSE 概念と ENABLE 概念を表す動詞を合計したものである。この結果から、日本人英語学習者が AwI 構文で CAUSE 概念を表すことができている一方、ENABLE 概念を表すことが少ないことが明らかである。

表 4 AwI 構文で使用された主節動詞ごとの頻度(正規化頻度は 10 万語当たり)

Verb	JPN 頻度	ENS 頻度	合計	JPN 正規化頻度	ENS 正規化頻度	対数オッズ比	p 値
<b>make</b>	99	27	126	55.87	30.02	1.417	< 0.001
<b>allow</b>	19	70	89	10.72	77.83	-1.762	< 0.001
<b>help</b>	30	44	74	16.93	48.92	-0.658	< 0.05
<b>want</b>	44	6	50	24.83	6.67	1.947	< 0.001
<b>let</b>	14	27	41	7.90	30.02	-0.908	< 0.01

表 5 CAUSE・ENABLE 概念を表す動詞をまとめた頻度

	JPN 頻度	ENS 頻度	JPN 正規化頻度	ENS 正規化頻度	対数オッズ比	p 値
<b>CAUSE</b>	126	56	71.11	62.26	0.937	< 0.001
<b>ENABLE</b>	78	151	44.02	167.88	-1.522	< 0.001

注 1: CAUSE を表す動詞は **make, force, cause, have, lead, get, require, prompt, inspire**

注 2: ENABLE を表す動詞は **allow, help, let, enable, teach, permit, assist, instruct** としている

注 3: オッズ比と p 値は AwI 構文の使用頻度を全体の頻度として計算している

##### 4.2 RQ2: ENABLE 概念の表現形式

###### 4.2.1 学習者コーパス分析の結果

表 6 は母語別(ENS, CHN, JPN)に ENABLE 概念をどのような形式で表現していたか集計したものである。AwI は不定詞付き対格構文を用いて表したも

表 6 ENABLE 概念の表現形式(括弧は列基準割合)

	ENS	CHN	JPN
<b>AwI</b>	54(31.4%)	60(28.7%)	16(9.6%)
<b>Conj</b>	27(15.7%)	18(8.6%)	33(19.8%)
<b>Prep</b>	43(25.0%)	74(35.4%)	93(55.7%)
<b>Other</b>	48(27.9%)	57(27.3)	25(15.0%)

の(4a)、Conj は **if** や **when** のような接続詞を用いたもの(4b)、Prep は **by** や **through** のような前置詞を用いたもの(4c)である。AwI と Conj と Prep が頻度的に主要な表現形式であったため、これらの頻度を集計し、その他の表現形式は Other にまとめた。

- (4) a. Having a part-time job allows you to meet people you might never have met if you do not have one. (ENS-PTJ-015)
- b. If you work at part-time job, you can talk with usual worker and see real work place and also notice the difficulty of working. (JPN-PTJ-012)
- c. By having a part-time job, we can communicate with people who we can't talk with in my usual college life. (JPN-PTJ-058)

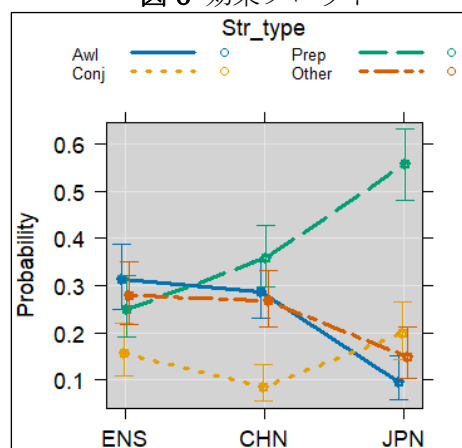
表7は母語別の表現形式の差が統計的に有意かどうか調べるために Region (ENS, CHN, JPN の3レベル)を説明変数に、表現形式(Str\_type) (AwI, Conj, Prep, Other の4レベル)を目的変数にして、Rで多項ロジスティック回帰分析(R core team, 2024)を行った結果である。図6はこの結果を効果プロットによって視覚化したものである。この結果から日本人英語学習者は母語話者と比較して、ENABLE 概念を表す際に、AwI を過少使用する一方、前置詞を用いた表現を過剰使用する傾向が示された。また中国人英語学習者は英語母語話者との形式においても有意な差がないことから、ENABLE 概念の表現形式の偏りは学習者一般の特徴ではなく、日本人英語学習者の特徴であることが示された。

表7 多項ロジスティック回帰分析の結果

	Estimate	Std.Error	zvalues	Pvalues
(Intercept):Conj	-0.693	0.236	-2.941	< 0.01
(Intercept):Prep	-0.228	0.204	-1.114	0.265
(Intercept):Other	-0.118	0.198	-0.594	0.553
RegionCHN:Conj	-0.511	0.357	-1.429	0.153
RegionCHN:Prep	0.451	0.268	1.683	0.092
RegionCHN:Other	0.049	0.272	0.18	0.858
RegionJPN:Conj	1.417	0.385	3.679	< 0.001
RegionJPN:Prep	1.988	0.339	5.861	< 0.001
RegionJPN:Other	0.564	0.377	1.498	0.134

$R^2 = 0.12$  LR = 62.134 df = 6 p < 0.001

図6 効果プロット



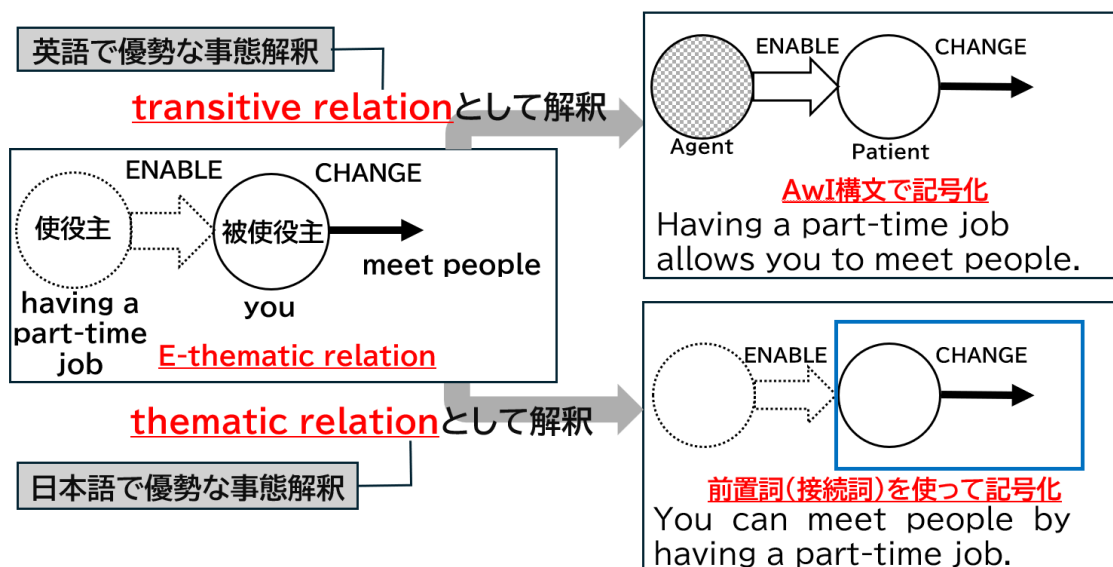
#### 4.2.2 事態解釈の日英差による考察

前節で示した日本人学習者の傾向を、事態解釈の日英差の観点から考察する。図7は「使役主である having a part-time job によって、被使役主 you が meet people することが可能となる」という事態を例にして、事態解釈による記号化の差異を図示したものである。ENABLE 概念は、被使役主 (Patient) に対して影響を与える使役主 (Agent) が存在するが、CAUSE 概念と比べて使役主の力が弱い (Kemmer & Verhagen, 1994)。従って Agent が背景化された E-thematic relation に対応し、曖昧性が生じる。そのため、transitive relation にも thematic relation にも解釈可能であり、2.2 節で述べたように、概念的自律性の差異の影響を受けやすい。CHANGE の部分 (図7では you + meet people の部分) に概念的自律性が認められなければ、transitive relation として解釈され、使役主 (Agent) を主語 (トラジェクター) としてプロファイルする構造である AwI 構文で記号化される。一方で、概念的自律性が認められれば、thematic relation として解釈され、使役主を前置詞や接続詞を用いて副詞的な付加要素として分割し、被使役主 (図7では you) を主語としてプロファイルする構造で記号化される。

(3a-b) の、事態解釈の基本単位の日英差が日本人英語学習者の ENABLE 概念の表現形式の特徴となって表れている。(3b) の特徴から、日本人学習者は使役主と thematic relation を分割して解釈する傾向があるため、使役主を主語として記号化するのではなく、副詞的要素として分割した表現形式を好むのである。同じ使役事態でも使役主の力が強い CAUSE 概念よりも、使役主の力が弱く、背景化されやすい ENABLE 概念の方が、この差異が表面化しやすいと思われる。その結果、ENABLE 概念において、日本人英語学習者と英語母語話者で構文選択に大きな差が出たと考えられる。また ENABLE 概念の表現形式の

差が中国人英語学習者と英語母語話者の間では有意でなかったことから、日本人学習者の特徴が、学習者一般のものではなく、日本語の事態概念解釈の影響であることが支持される。

図 7 ENABLE 概念の事態解釈の違いによる記号化の差異



## 5. まとめ

本研究では、ENABLE 概念を日本人英語学習者がどのように表現しているのか定量的なデータを示した。日本人学習者は ENABLE 概念を記号化する際に使役主を主語にした構文である AwI 構文を過少使用する一方で、使役主を副詞的要素として分割させた前置詞や接続詞を使った構文を好むことが示された。そして、この特徴は日本語の事態解釈の影響によるものであると論じ、母語の事態解釈の傾向が外国語の使用に影響を与えることを実証的に示すことができた。本研究では学習者の習熟度別の差異を検討するに足るデータ量がなかったため、今後はこの点を調査したい。また、紙幅の都合上、各構文でどのような語彙を使用していたのか(前置詞構文であれば, by や through)という点について述べられていない。この点に関しては別稿に譲りたい。

## 引用文献

- Ishikawa, S. (2023). *The ICNALE guide: An introduction to a learner corpus study on Asian learners' L2 English*. Routledge.
- Kemmer, S., & Verhagen, A. (1994). The grammar of causatives and the conceptual structure of events. *Cognitive Linguistics*, 5(2), 115-156.
- Langacker, R. W. (1991). *Concept, image and symbol: the cognitive basis of grammar*. De Gruyter Mouton.
- R Core Team (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- 谷口一美 (2005)『事態概念の記号化に関する認知言語学的研究』ひつじ書房.
- Wolff, P., & Song, G. (2003). Models of causation and the semantics of causal verbs. *Cognitive Psychology*, 47(3), 276-332.

HAVE the advantage が従える同格節 that  
—that 節中の助動詞 can は必須要素か—

土屋 知洋(芝浦工業大学)  
tomotsu@shibaura-it.ac.jp.

A Special Focus on the Appositive *That*-clause Following *HAVE*  
*the Advantage*  
—the Role of the Auxiliary Verb *Can*—

TSUCHIYA Tomohiro (Shibaura Institute of Technology)

Abstract

The main aim of this paper is to examine whether the abstract noun *advantage* can be followed by an appositive *that*-clause and whether the auxiliary verb *can* is necessary within the clause to explain *advantage*. Descriptive research, incorporating both corpus analysis and informant surveys, reveals two key findings. First, the construction where *advantage* takes an appositive *that*-clause tends to appear in the written register of British English. Second, the presence of the auxiliary verb *can* in the clause is not an essential element for *advantage* but is context-dependent.

Keywords

*advantage* の語法, 同格 *that*, 英米とレジスターによる相違, 意味と形式の関係性

1. はじめに

「～という強み[長所]を持つ」という表現 HAVE the advantage は, of を伴い(1)の名詞や動名詞を従える形が一般的であるが, イギリス人英語母語話者は, (3)の理由で(2a)は正しくなく(2b)が正しいという(以下, 下線及びイタリックは著者)。ここで問題となるのが, *advantage* が同格 *that* を従えることが可能なのか, 同格節内に“可能性や能力”を表す *can* が必要なのか, という点である。

- (1) a. She *had the advantage of* a good education. OALD 10 (s.v. *advantage*)  
b. Ken *has the advantage of* speaking Italian. W4 (s.v. *advantage*)  
(2) a. ? ... online classes *have the advantage of* being taken at any time.  
b. ... online classes *have the advantage that* they can be taken at any time.  
(3) I think the first version might be accepted in a casual email, but for exam purposes it is not correct. Because an advantage is being described, it is better to include “can”. (You can do this if you want.) Therefore I think the second version is correct.

## 2. 先行研究

### 2.1 advantage が従える同格 of

同格を表す後置修飾の方法には(4)の 3 つが存在する。実際, Biber et al. (1999: 654)は advantage を同格節の説明の一部で「抽象名詞+of+ing」にリストアップしている。また, (5)の辞書では「《◆of は同格の of》」と注付きで用例を上げ, COCA でも容易に検索 (HAVE the advantage of \*ing の形で 961 件)される。

- (4) ① to-infinitive のみ容認 ② to-infinitive と of +ing 共に容認 ③ of +ing のみ容認  
Quirk et al. (1985: 1272-1274) 安藤 (2007: 782-784)
- (5) He *had the advantage (over me) of knowing* the language. G 大 (s.v. advantage)
- (6) ... , he will *have the advantage of being elected* as a candidate of reform. COCA

### 2.2 advantage が従える同格 that

一方, advantage が同格 that を従えるという記述は, 綿貫・マークピーターセン (2006: 234) が「同格の that」を従える名詞としてリスト化し, (7)の例が見られる程度である。

- (7) He *had the advantage over me that he could* speak English.  
ランダムハウス (s.v. advantage)

先行研究が少ないため, 統語的振る舞いから同格の可能性を探る。2.1 節で触れたように advantage が同格 of を従えるという事実がある。一般に, 同格節を従える名詞は(8)のように be 動詞で結ばれる論理関係にあるといわれ, 名詞 advantage でも同形式の用例((9)参照)が検索される(COCA では 71 件)ことから, 同格の that 節を従える可能性は否定できない。

- (8) The *possibility is that he will resign as ambassador*. 安井 (1996: 72)
- (9) There are many ways of making thin film cells; the *advantage is that the amount of material needed is very small and that reduces the cost*. COCA

### 2.3 HAVE the advantage that 節内における can の必要性

Advantage が従える同格の that 節内に can が必要かという先行研究はないが, 辞書には could を含む用例が見られ((10)), コーパスでは advantage of に続く形に being able to V を検索するのは難しくない。先行研究を踏まえ, 次節から advantage が [1]同格の that 節を従える可能性, [2]従える同格 that 節内の can の必要性と有無の要因, の 2 点を実証的に調査する。

- (10) He *had the advantage over me that he could* speak English. (=7)

(11) a. Regular exercise appears to *have the advantage of being able to jump-start the immune system, and ...* COCA

b. He *has the advantage of being able to play in several forward roles.* WB (UK, W)

### 3. 量的調査

#### 3.1 Have the advantage of Ving の V の種類

コーパスの検索では, (12)から(16)の下線部のように of に後続する動詞の種類(状態・動作)は多岐に渡る。しかし, 表1と 2 の COCA, BNC, WordbanksOnline(以下, WB)を用いた量的調査の結果を見てみると, 圧倒的に being を従える形式が多いことが明らかとなる。

(12) Dana Perino *has the advantage of being an attractive female.*

(13) Made of carbon, these cells *have the advantage of being produced at low cost ...*

(14) They *have the advantage of being able to gain the trust of other residents, ...*

(15) I also *had the advantage of having gone to an experimental progressive school ...*

(16) ... George W. Bush will *have the advantage of knowing these lessons from the past.*

Being に続く形は, (12)のような名詞や形容詞が多いが, 「状況的可能性」「能力」を表す able to V も約 40%の出現率である。また, コーパスのサイズとレジスターを考慮するとイギリス英語の書き言葉で用いられる傾向にあることが分かる(以降, 表中の( )は話し言葉の件数;太字は重複語)。

表 1 COCA 及び BNC における HAVE the advantage of Ving の V の種類

順位	COCA				BNC			
	件数	V の種類	being + α		件数	V の種類	being + α	
1	246	<b>being</b>	名/形	150	88 (1)	<b>being</b>	名/形	54
2	44	<b>having</b>	受動	39	34	reading	受動	16
3	25	<b>allowing</b>	<b>able to</b>	<b>36(1)</b>	9	<b>providing</b>	<b>able to</b>	<b>12</b>
4	19	<b>knowing</b>	存在	19	9	making	存在	6

表 2 WB における HAVE the advantage of Ving の V の種類

順位	件数	V の種類	US : being + α		件数	V の種類	UK : being + α	
			名/形	5(2)			名/形	136(2)
1	16(1)	<b>being</b>	名/形	5(2)	214(2)	<b>being</b>	名/形	136(2)
2	6	<b>knowing</b>	受動	6	50	<b>having</b>	<b>able to</b>	<b>34</b>
3	5	<b>having</b>	<b>able to</b>	<b>5</b>	30(2)	<b>knowing</b>	受動	22(1)
4	2	<b>allowing</b>			24	playing	存在	21



### 3.2 HAVE the advantage that S' + (CAN) + V'の可能性

先行研究に同格の *that* 節を従える可能性と *can* の必要性を明示したものはないが、コーパスからは当該形式が検索される((17)(18))。表 3 から、同格 *that* を従える形が書き言葉で利用傾向にあること、*can* と共起する割合が 20%から 25%であることから、必須要素でないことが分かる。

(17) Local organizations *had the advantage that cooperation came about* quickly because people knew and trusted one another. COCA

(18) a. Wind generators also *have the advantage that they can* go on working 24 hours a day, where solar panels can only operate in sunlight. BNC (W)

b. ID card *has the advantage that it can* be used with more than one bank. WB (UK, W)

(19) Making a will with a solicitor also *has the advantage of that* you'll have a copy, he's likely to look after it for you if you want him to. BNC (S)

表 3 HAVE the advantage that S' (CAN) V'の実態

	COCA	WB [UKW]	WB [USW]	BNC
同格 <i>that</i> 節	83	未確認	未確認	95 (1)
+ CAN	20	12	1	16

尚、上記(19)は唯一の話し言葉の実例だが、HAVE the advantage of と言いかけ、*that* 節に言い換えていることから、この *that* が *of* と同じ同格用法として利用されている証拠ともいえる。

## 4. 質的調査

### 4.1 HAVE the advantage that S' (CAN) V'の生起条件

実例を精査すると、(20)のように“事実・状態・習慣”を表す場面では *can* は生起しない傾向にある。一方、(21a)では風力発電と太陽光パネルを比較し機能の高さを強調し、(21b)はモデルによる行為の可能性を強調するため *can* と共起していることが明らかとなる。

(20) a. Old photos and videos *have the advantage that* they often contain little gems that you had entirely forgotten about. COCA

b. HUCKABEE: Herman Cain does *have the advantage that* he talks the language of people who understand that he's talking to them ... . COCA (S)

(21) a. Wind generators also *have the advantage that they can* go on working 24 hours a day, where solar panels can only operate in sunlight. BNC

b. Software models *have the advantage that* you can make the models as complex as you need to, ... . COCA

#### 4.2 HAVE the advantage of Ving か being able to V の選択

HAVE the advantage of に後続する形を選択も、4.1 節の can の生起条件を適用することができる。(22b)の用例は(22a)の習慣と異なり([ ]は著者補足)、普通の選手と異なりフォワードの様々な役割を果たす能力の高さを強調しているため being able to V の形で表現している。

- (22) a. [Price] *has the advantage of playing with Brown, a strong point guard.* COCA  
b. ... he is very professional, a charming man and a good guy to have as a team-mate.  
He has the advantage of being able to play in several forward roles. WB (UK)

#### 4.3 質的調査のまとめ

実例から、HAVE the advantage of Ving / that S' (CAN) V' とともに“可能性や能力”を表す can は必須要素ではなく、[1] 行為・機能が他と比較して能力の高さや可能性を強調する場合に必要、[2]事実・状態・習慣的行為を示す場合は不要、な傾向にあることが明らかとなった。

### 5. インフォーマント調査

これまでの調査結果の妥当性をインフォーマント(米2名・英6名)で確認した。(23)は「比較・可能性」、(24)は「状態・習慣」、(25)は「理由・能力」という意味を含んだ質問文である。回答は【容認(○)–不自然(△)–非容認(×)】で示し、( )内はイギリス人の回答数である。

- (23) a. Compared with face-to-face classes, online classes have the advantage of *being taken* at any time. △ : 著者予想とインフォーマントの回答 【5(5)-2(2)-1】  
b. Compared with face-to-face classes, online classes have the advantage of *being able to be taken* at any time. ○ : 【2(2)-5(3)-1(1)】  
c. Compared with face-to-face classes, online classes have the advantage that they are taken at any time. × : 【1(1)-0-7(5)】  
d. Compared with face-to-face classes, online classes have the advantage that they can be taken at any time. ○ : 【5(4)-0-3(2)】
- (24) a. Old albums have the advantage of containing little gems that you had entirely forgotten about. ○ : 【5(4)-1-2(2)】  
b. Old albums have the advantage of being able to contain little gems that you had entirely forgotten about. × : 【0-4(4)-4(2)】  
c. Old albums have the advantage that they often contain little gems that you had entirely forgotten about. ○ : 【3(3)-2(1)-3(2)】  
d. Old albums have the advantage that they can often contain little gems that you had entirely forgotten about. × : 【2(2)-1(1)-5(3)】

- (25) a. They have the advantage of working overseas because they speak English fluently.  
 △ : 【5(4)-1(1)-2(1)】
- b. They have the advantage of being able to work overseas because they speak English fluently. ○ : 【7(5)-0-1(1)】
- c. They have the advantage that they work overseas because they speak English fluently. × : 【0-1-7(6)】
- d. They have the advantage that they can work overseas because they speak English fluently. ○ : 【5(4)-3(2)-0】

(23b)は being able to に受け身が続く形で容認度が低い、(23d)は「比較」により同格 that の容認度は高い。「状態」を表す(24)は“可能性や能力”と共起傾向になく、(25)は「理由」の明記で“能力”が強調され being able to や can と共起傾向にあり、調査結果を裏付ける傾向となった。

## 6. まとめ

先行研究で触れられてこなかった advantage が同格の that を従えるのか、また同格節内に can が必要なのか、についてコーパスとインフォーマントに基づく実証的調査を行った。結果、イギリス英語の書き言葉で同格節に従える傾向にある一方で、can は必須要素ではなく文意と密接に関係していることが明らかとなった。インフォーマント調査におけるバランスの問題で更なる調査は必要だが、本研究の新たな知見は辞書記述の改善に貢献できるものである。

## 引用文献

- 安藤貞雄. (2007). 『現代英文法講義』 開拓社.
- Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan. (1999). *Longman Grammar of Spoken and Written English*. Longman.
- Quirk, R., S. Greenbaum, G. Leech and J. Svartvik. (1985). *A Comprehensive Grammar of the English Language*. Longman.
- 綿貫陽・マークピーターセン. (2006). 『実践ロイヤル英文法』. 旺文社.
- 安井稔. (1996). 『コンサイス英文法辞典』. 三省堂.  
 [辞書・コーパスとその略称]
- G 大: 『ジーニアス英和大辞典』. (2001). 大修館書店.
- OALD10: *Oxford Advanced Learner's Dictionary* 10<sup>th</sup>. (2020). OUP.
- ランダムハウス: 『ランダムハウス英和大辞典』. (1973). 小学館.
- W4: 『ウィズダム英和辞典』第4版. (2019). 三省堂.
- COCA : Corpus of Contemporary American English
- BNC : British National Corpus
- WB : WorbanksOnline \*小学館コーパスネットワークで提供されていたデータ

# 同一学習者の技能間における産出結果の違い —スピーキングとライティングの比較から—

前田 裕司(九州大学 大学院生)  
yuji.maeda.275@s.kyushu-u.ac.jp

## The Differences in Output Results Across Skills within the Same Learners —From the Comparison of Speaking and Writing—

MAEDA Yuji (Kyushu University, Graduate Student)

### Abstract

本研究は、同一学習者のスピーキングとライティングの産出結果における使用語彙レベルやその他の言語的特徴にどのような違いがあるかを明らかにしたものである。31名の日本の高校生からスピーキングとライティングのデータを収集し、ミニ学習者コーパスを構築した。CVLAを用いて分析した結果、ライティングよりもスピーキングの方が、統計的に有意に使用語彙レベルが低いなどの結果が明らかになった。

### Keywords

L2 スピーキング, L2 ライティング, 学習者コーパス, CEFR

### 1. はじめに

日本人英語学習者のスピーキングとライティングという2技能の産出結果に関して、それぞれの技能に関連するコーパス研究は、近年盛んに行われている。しかしながら、同一学習者による2技能の産出結果をコーパス化したものは限られており、それに関する研究も、単一技能のデータから成るコーパスを基にした研究と比較して非常に少ない。

本研究では、日本の高校に通う高校1年生31名から、スピーキングとライティングのデータを取得、コーパス化し、技能間において、テキストの特徴を表す指標にどのような違いがあるかを明らかにする。

### 2. 先行研究

Vibar (2022)では、フィリピンの女子高校生1名のSpeakingとWritingの語彙密度について事例研究を行なった。共通テーマについてエッセイライティングとインタビューを実施し、その産出結果を比較したところ、Speakingの方が産出語数が多かったこと、Speakingの方が機能語の使用割合が多かったこと、Writingの方が独立節の数が多かったことなどが明らかになっている。

Abe (2007)では、中学1年生から高校3年生までの日本人英語学習者によるwrittenと

spoken の差を、コーパスを用いて分析した結果、異なる発達段階において、異なる産出の差があることがわかった。ライティングでは、時制、動詞に関する間違い、スピーキングでは、動詞の相、主語と動詞の一致、名詞の派生、これらのエラーが、発達段階が高くなるにつれ少なくなっていた。

さらに、野村(2009)では、中高生それぞれ 20 名分の written と spoken のデータをコーパス化し、単語および品詞 n-gram から分析した結果、産出モードの違いによる output の違いがあることが明らかになった。

### 3. リサーチデザイン

#### 3.1 研究目的と研究設問

本研究の目的は、スピーキングとライティングの2技能における産出結果において、どのような違いがあるか、さまざまなテキストの特徴を示す指標から統計的に分析し、明らかにすることである。そこで以下のリサーチクエスチョン(RQ)を設定した。

**RQ:** 同一学習者によるスピーキングとライティングの産出結果におけるテキストの特徴を示す指標に統計的有意差はあるか。

#### 3.2 データ

本研究では、日本の高校に通う高校1年生 31 名からスピーキングとライティングの産出データを収集した。スピーキングに関しては1つのタスクのデータを収集し、ライティングに関しては3つのタスクのデータを収集した。それぞれのタスクのテーマと学習者が産出する際の条件については以下の通りである。

##### スピーキングタスク

テーマ: Choose one of the prefectures in Japan you want to visit and make a 1-minute speech. Please include why you want to visit the prefecture.

条件: 準備時間を 1 分間設け、メモを取ることを許可。

##### ライティングタスク

タスク 1 テーマ: 次の問いについて、あなたの意見とその理由を 2 つ挙げ、30 語～50 語程度の英文で書きなさい。 What is the best season to visit your hometown?

タスク 2 のテーマ: 次の問いについて、あなたの意見とその理由を 2 つ挙げ、30 語～50 語程度の英文で書きなさい。 What is your favorite school subject and why?

タスク 3 のテーマ: 次の問いについて、あなたの意見とその理由を 2 つ挙げ、30 語～50 語程度の英文で書きなさい。 Do you prefer to eat Japanese food or Western food?

条件: ネット情報、ネット辞書の使用可、翻訳機能の使用不可、時間制限無し。

表1 コーパスの概要

	n	総語数	平均語数	最大語数	最小語数
スピーキングタスク	31	1,398	45.1	129	14
ライティングタスク1	31	1,314	42.4	56	30
ライティングタスク2	31	1,068	34.5	59	30
ライティングタスク3	31	1,325	42.7	58	30
ライティング平均	N/A	1,236	39.9	N/A	N/A

### 3.3 手法

本分析には、CVLA (Uchida & Negishi, 2018) というウェブソフトを使用した。CVLA は入力されたテキストに対して、テキストの特徴に基づき、CEFR-J レベルを付与するシステムである (Uchida & Negishi, 2018)。本分析では、CVLA で得られる、ARI, VperSent, AvrDiff, そして BperA というテキストの特徴を示す 4 項目の数値を取得し、それらの項目ごとにスピーキングとライティング間の数値に統計的有意差があるか確認するため、Python の spacy を用いて統計的検定を行なった。各項目の簡易的な説明は以下の通りである。

ARI : 英語テキストの読みやすさを数値化したもの。

VperSent : 一文あたりの動詞の平均割合。

AvrDiff : 語彙難易度の平均値 (A1, A2, B1, B2 はそれぞれ 1, 2, 3, 4 としてカウントされる。内容語のみ、機能語は除外。)

BperA : A レベルの内容語に対する B レベルの内容語の割合

また CVLA で取得できる上記の 4 項目以外にも、多角的にスピーキングとライティングの産出結果の差を測るため、TTR を 50 語で調整した STTR (水本, 2008) と機能語に対する内容語の割合についても算出し、統計的検定を行なった。TTR はテキストが長くなればなるほど数値が小さくなる傾向がある。スピーキングでは、学習者間で語数にばらつきが見られたため、TTR ではなく 50 語で調整した STTR の数値を使用した。

ライティングについては、タスクのテーマによる影響を減らし、より一般的なライティングの産出結果としての数値として扱うため、3 つのタスクの平均値を検定に使用した。

#### 3.3.1 データの事前処理

スピーキングの産出データに関しては、一度、学習者が自身で録音データを基に書き起こしを行なった。その後、書き起こしデータを筆者が録音データを聞きながら確認し、必要に応じて加筆修正を行なった。「un」や「hmm」といったフィラーは取り除き、繰り返しや、自己訂正についてはそのままとした。文の区切りについては、学習者による書き起こしを元に筆者が音声データを確認する

過程の中で、ポーズの長さや文構造を参考にしながら、ピリオドの位置を設定した。

ライティングの産出データに関しては、各項目の数値の算出に影響が出ることを防ぐため、スペルミスのみ、Google スプレッドシートのスペルチェック機能を使用して訂正を行った。訂正が適切かどうかの判断については、筆者が提示されたスペルミスを全て確認した上で行なった。

### 3.3.2 分析の手順

本分析では、まず学習者から取得したスピーキングとライティングの産出データをエクセルファイルに集約し、簡易的なミニ学習者コーパスを構築した。その後、分析対象となる項目の数値データを取得するための Python コードと、その数値データに対して行う統計的分析のための Python コードを作成した。それらを用いてミニ学習者コーパスに格納されているスピーキングとライティングの産出データ(計 124 個分)を自動分析し、その結果得られた数値データに対し、統計的な分析を行なった。

## 4. 結果と考察

### 4.1 CVLA から得られた4項目の分析結果

CVLA を用いて得られたスピーキングとライティングの各項目の平均値と有意差検定の結果を以下の表に示す。

表2 CVLA から得られた各項目の平均値と検定結果

	スピーキング	ライティング	t 検定	マン・ホイットニーの U 検定
ARI	0.0594	3.7380	$t = -8.38, p < .05$	N/A
VperSent	1.8713	1.6756	N/A	$U = 505.5, p = .73$
AvrDiff	1.1694	1.4548	$t = -10.23, p < .05$	N/A
BperA	0.0384	0.1828	N/A	$U = 22.0, p < .05$

項目ごとにスピーキングとライティングの平均値に、統計的有意差があるか分析を行なった。まず項目ごとに得られた数値が正規分布しているか調べるため、シャピロ・ウィルク検定を行った。そこでどちらも正規分布に従っている可能性が高いと判断された場合、対応ありの t 検定を行なった。どちらか一方が正規分布に従っていない可能性が高いと判断された場合は、マン・ホイットニーの U 検定を行なった。統計的検定の結果から、VperSent 以外の項目において、スピーキングとライティング間に有意差が確認された。ARI に関しては、文長や単語の長さに影響を受けやすい指標であり、ライティングに比較して、スピーキングでは、発話語数が少ない学習者が多かったこと、学習者間で語数の差が極端にあったことで有意差が出たと考えられる。また、ARI はテキストのリーダビリティを測る指標であるため、スピーキングの産出結果に使う指標として適切ではなかった可能性も考えられる。テキストで使用されている語彙の難易度を CEFR-J Wordlist に基づいて算出する AvrDiff に関して統計的な有意差が確認されたことから、ライティングよりもスピーキングの方が

産出語彙レベルが低かったということが言える。ライティングでは、準備時間や産出の際の時間に制限がなかったこと、わからない語彙についてはインターネットで調べて書くことが許されていたことから、スピーキングよりも語彙レベルが高くなったことも要因の一つとして考えられる。BperAについても AvrDiff と同様、辞書の使用や時間制限等が影響を与えた可能性が考えられる。VperSent に有意差が見られなかったこと、また統計的有意差は認められなかったが、ライティングよりもスピーキングの方が数値が高かったことは仮説とは異なる結果であった。これにはピリオドの位置が大きく影響を与えるが、スピーキングの書き起こしの際、どこまでを一文とするかについては、主観的な判断となってしまったため、信頼性については疑問が残る。

#### 4.2 STTR や機能語に対する内容語の割合

表 3 タスクごとの記述統計量(カッコ内には割合を示した)

	n	STTR(50)	平均内容語数	平均機能語数	内容語/機能語
スピーキングタスク	31	0.54	23.6(51.8)	22.4(48.2)	1.1
ライティングタスク1	31	0.76	24.6	18.3	1.4
ライティングタスク2	31	0.70	18.8	15.8	1.2
ライティングタスク3	31	0.64	26.4	16.5	1.6
ライティング平均	N/A	0.70	23.3(57.9)	16.9(42.1)	1.4

表 4 STTR と機能語に対する内容語の割合と有意差検定の結果

	スピーキング	ライティング	t 検定
STTR	0.54	0.70	t = 9.81, p < .05
内容語/機能語	1.1	1.4	t = -4.38, p < .05

標準化された TTR である STTR については、スピーキングタスクでは 0.54 であったのに対し、ライティングタスクでは、平均で 0.70 という結果であった。これについて t 検定を実施したところ有意差が認められた。このことからライティングよりもスピーキングの方が、同じ語を繰り返し使用する傾向があると言える。また機能語に対する内容語の割合に関しても統計的有意さが認められた。しかしながら、スピーキングの産出データから、繰り返しや自己訂正を取り除かなかったため、このような結果が得られた可能性が考えられる。また、スピーキングの方がライティングよりも機能語の割合が高いという結果が得られたが、これは Vibar(2020)の研究結果と同じ結果である。より正確に実態を把握するには、繰り返しや自己訂正の除いたデータでの分析が必要と考える。

#### 5. まとめ

本研究は同一学習者によるスピーキングとライティングの産出結果にどのような差があるかを明ら



かにすることを目的に、テキストの特徴を示す様々な項目から分析を行なった。

分析の結果、スピーキングとライティングの産出結果において、ARI, AvrDiff, BperA, STTR, 機能語に対する内容語の割合の項目に、統計的有意差が認められた。これらの結果から、今回の研究対象となった学習者は、ライティングと比較してスピーキングでより難易度の低い語彙を使用すること、より限定的な語彙使用をすることが明らかになった。

しかし、分析対象とした産出データの量が限られていたこと、特にスピーキングに関しては1つのタスク分しか取れておらず、一般化するにはデータが非常に少なかったことは今後の課題である。またデータの取得方法についても、スピーキングとライティングとで、条件が異なっていた。今後、分析結果をより一般的な結果とするためには、データ量とデータ取得時の条件を揃えていく必要がある。

またスピーキングの産出結果の処理方法が分析結果に影響を与えるため、フィラーや繰り返し、自己訂正の扱いについて、研究の目的に合わせ、どのように扱うか、再度検討する必要がある。

## 謝辞

本研究は九州大学と新興出版社啓林館の共同研究として助成を受けたものです。

## 引用文献

- Abe, M. (2007). Grammatical errors across proficiency levels in L2 spoken and written English. *The Economic Journal of Takasaki City University of Economics*, 49(3), 4.
- 水本篤. (2008). 自由英作文における語彙の統計指標と評定者の総合的評価の関係. 統計数理研究所共同研究レポート 215 「学習者コーパスの解析に基づく客観的作文評価指標の検討」, 15-28.
- 野村真理子. (2009). 日本人英語学習者の産出モードの違いによる言語特徴の分析—中学・高校段階の学習者に焦点をあてて—. *四国英語教育学会紀要*, 29, 15-23.
- Uchida, S., & Negishi, M. (2018). Assigning CEFR-J levels to English texts based on textual features. In *Proceedings of Asia Pacific Corpus Linguistics Conference*. 4, 463-467.
- Vibar, J. D. (2022). Syntactic analysis between spoken and written language: A case study on a senior high student's lexical density. *Asian Journal on Perspectives in Education*, 3, 86-106.

# Than I/me を伴う比較構文における格のゆれ

久屋 愛実(立命館大学)

kuya@fc.ritsumeai.ac.jp

## The Choice of Case in the Comparative Construction with *Than I/Me*

KUYA Aimi (Ritsumeikan University)

### Abstract

This study describes the shift from the nominative case to the accusative case in the comparative construction with *than I/me*. Personal pronouns after *than*, despite the rules of prescriptive grammar, often appear in the accusative case even when they correspond with subjects in the main clause. Data from the 2000s and later show that the shift has made further progress to the extent that nominative and accusative forms are not anymore “fairly evenly divided” (Biber et al., 1999, p.336), suggesting a stronger tendency for *than* to be recognized as a preposition rather than a conjunction. It is also suggested that the shift toward the accusative case in the current construction is a little delayed compared with that in the typical *It is I/me* construction (Kuya, 2021) but is advanced compared with that in the *It is I/me who* construction (Kuya, 2024).

### Keywords

比較構文, 人称代名詞の格, *than*, 接続詞, 前置詞

### 1. はじめに

本研究の目的は, *than I/me* を伴う比較構文における, 主格に替わる目的格使用の拡大について記述することである。He is taller than I/me.のように, *than* 節が述部を持たない場合(以下, 「*Than I/me* 構文」), *than* 節内の人称代名詞の格はその文法機能に一致させることが規範文法の説くところである。しかし, 運用面では *than* 節内の人称代名詞が目的語機能を有する場合だけでなく主語機能を有する場合でもしばしば目的格として現れる。主格から目的格への変化は, 言い切り型の *It is I/me* 構文(Who's there?等への返答としての)や *who* 節が後続する *It is I/me who* 構文(*It was I/me who was wrong* 等)などにおいても拡大していることが報告されている。言い切り型構文については, コーパス調査の結果から, 規範文法の求めるところにもかかわらず, 主格から目的格への変化がほぼ完了したと見られている(Dekeyser, 1975; Erdmann, 1978; Biber et al., 1999; Nakayama, 2014; Kuya, 2021)。It is I/me who 構文においても, relative

attraction (Jespersen, 1894) と呼ばれる関係代名詞による文法的制約にもかかわらず、主格 I に替わり目的格 me の使用が拡大しているとの指摘がある(久屋, 2024)。そこで本研究では、It is I/me who 構文と同じく規範文法的な格の一致のルールに反して目的格使用が拡大していると見られる Than I/me 構文において、変化がどの程度進行しているかを 2000 年代以降のデータを含めて調査し、上記関連構文との比較を行う。

## 2. 先行研究: Than I/me 構文における目的格使用

比較構文に現れる than は接続詞の位置付けであるため、従属節を導く(1a-d)。ただし、この than 節において主節と重複する部分(1a-d の括弧内の要素)は通常省略され、than に続く人称代名詞は主節主語と対応する場合に主格として現れ(1a-c)、主節目的語と対応する場合には目的格として現れる(1d)。しかし、運用面(特にインフォーマルな場面)ではこのような区分が厳密にはなされず、(2a-c)のように目的格として現れることが多いという(Jespersen, 1949; Quirk et al., 1972)。ただし、主節動詞が目的語を伴う(2c)のような場合は、(1c)の意味にも(1d)の意味にもなりうるため曖昧性の問題が生じる(Fowler & Fowler, 1931; Curme, 1931)。

(2a-b)のように、than 節内の人称代名詞が明らかに主語役割を担う場合でも目的格の使用が拡大している背景としては、than が接続詞ではなく前置詞としての機能を獲得しつつあること(Jespersen, 1949; Quirk et al., 1972; OED 2nd ed.), さらに言い切り型構文や It is I/me who 構文などの関連構文においても主格に替わり目的格の使用が拡大していることが挙げられる。

(1)

- a. She works harder than I (work/do).
- b. He is taller than I (am).
- c. I love you more than he (loves/does you).
- d. I love you more than (I love/do) him.

(2)

- a. She works harder than me. (1a の意味で)
- b. He is taller than me. (1b の意味で)
- c. I love you more than him. (1c または 1d の意味で)

Than I/me 構文自体は会話体やフィクションで多く観察され、新聞や学術的文章ではほとんど見られない(Biber et al., 1999)。この点は、言い切り型構文(Kuya, 2021)や It is I/me who 構文(久屋, 2024)において、フィクションやテレビ・映画の字幕といった口語性の高いジャンルで用例が多く収集されたことと共通する。

関連構文との比較の視点からは、まず、Biber et al. (1999) が、「as/than + 人称代名詞」のコロケーションにおいて、会話体ではほぼ目的格のみ、フィクションでは「主格と目的格が半々」(p.336)

で出現する(Erdmann, 1978 も参照)と述べている。これに対して言い切り型構文「it + be + pronoun (final)」では、会話体とフィクションどちらにおいても目的格の比率が圧倒的に高い(Biber et al., 1999)。さらに、言い切り型構文よりも Than I/me 構文において主格の使用がわずかに多く見られるとの記述もある(Leech & Svartvik, 1975)。これらのことから、Than I/me 構文における主格から目的格への変化は、言い切り型構文におけるそれよりもやや遅れて進行していると予測される。

人称代名詞の種類については、Kuya (2021)が、言い切り型構文において目的格への移行が最も早く完了段階に近づいたのは 1 人称単数 I/me で、それにやや遅れる形でその他の人称代名詞 they/them, we/us, he/him, she/her も完了に向かっていることを示した。Than I/me 構文でも人称代名詞の種類によって変化の進行の度合いが異なる可能性がある。

### 3. リサーチデザイン

本研究では、than 省略節を伴う比較構文のうち、「than +主格」を既存形、「than +目的格」を革新形として扱い、主格から目的格への変化を記述する。前節の議論を踏まえ、(i)当該言語変化(主格に替わる目的格使用)がどの程度進行したか、(ii)変化の進行の度合いは関連構文である言い切り型構文や It is I/me who 構文と比較して異なるかという問いについて論じる。手順としては、まず、than 節が省略節として文末に現れる用例「than +人称代名詞+ピリオド/?/!」を文字列検索し、(i)について観察する。その後、収集された用例のうち than 節の人称代名詞が明らかに主格役割を担う用例に絞って、(ii)についての考察を行う。

上記関連構文との比較のために COHA (Corpus of Historical American English) (Davies, 2010)を利用する。COHA では 1820 年代から 2010 年代までの 200 年間のデータの採取が可能で、Fiction, Poplar Magazines, Newspapers (1860 年代以降), Non-fiction Books, TV/Movies (1930 年代以降, 2021 年更新時追加)の 5 つのジャンル(合計約 5 億語)からなる。言い切り型構文(Kuya, 2021)や It is I/me who 構文(久屋, 2024)の用例数が最も多かった Fiction が全体の約半数を占めるコーパスである。

COHA から採取された用例の代表的なものを(3a-i)に示す。than 節の人称代名詞が明らかに主格役割を担う典型的な例は前述した(1a-b, 2a-b)のように主節動詞が目的語を伴わない場合(Jespersen, 1949)である。COHA からの類例は(3a-c)で、比較級は形容詞(3a)と副詞(3b)を含み、be 動詞の省略等が起きている例(3c)もある。主節動詞が目的語を伴う場合は、than 節の人称代名詞が主節主語に対応すると判断できる(3d-e)のような例から、Huddleston & Pullum (2002, p.460)の指摘する、主節目的語と結びつけて解釈されやすい例(I've not met a nicer man than he/him) (COHA からの類例は 3f-g)までさまざまで、中には(2c)で見たように than 節内の人称代名詞が担う文法役割を特定することが容易ではないものもある。また、Jespersen (1949, p.231)の指摘する、構造的にどちらの格が「正しい」のかが判断しにくい例(I never saw anybody stronger than he/him)もある(COHA からの類例は 3h-i)。今回は、人称代名詞が主語役割を担うと見られる(3a-e)のタイプに限って分析を進め、(2c)や(3f-i)タイプの用例と、熟

語・慣用的な表現 (better you than me, different than, (would) rather than, other than 等) は除外することとした。

(3)

- a. You're even a little older **than I**. (1937, Fiction, I Can Get It You Wholesale)
- b. “Well you drive better **than me**.” (1975, Fiction, Ragtime)
- c. She had a brilliant mind, so much smarter **than I**. (1950, TV/Movies, The Underworld Story)
- d. You know that better **than me**. (2003, TV/Movies, 2003 ABC World Stunt Awards)
- e. Well, nobody's got a better dream **than me**. (1952, TV/Movies, Pat and Mike)
- f. as though it were a question that concerned her more **than me**. (1973, Fiction, After)
- g. and you will marry a better woman **than me!** (1865, Magazines, Harpers)
- h. But surely the Prophets would have chosen someone more worthy **than I**. (1999, TV/Movies, Star Trek: Deep Spac...)
- i. he took me and my brother, a little older **than me**. (1942, Magazines, Atlantic)

#### 4. 結果と考察

COHA にて「**than** + 人称代名詞 + ピリオド / ? / !」を文字列検索したところ、全期間で主格 1,452 例、目的格 1,220 例 (合計 2,672 例) が採取された。主格と目的格の合計を 100 とした時の目的格の比率は 46% で、全体としては主格が目的格をやや上回っている。人称代名詞ごとの用例数 (主格用例数 / 目的格用例数) は I/me (207/737), he/him (571/171), she/her (258/103), we/us (159/130), they/them (257/79) である。1 人称単数 I/me のみ目的格の出現数が主格のそれを上回っていることから、言い切り型構文 (Kuya, 2021) と同様、他の人称代名詞よりも I/me で変化が進んでいる可能性がある。

年代ごと (10 年区切り) の変遷を見ると、目的格の比率は 20 世紀後半に 50% を上回り、2010 年代には 90% に達している。Biber et al. (1999) と比較するためにフィクションにおける I/me と he/him に限定しても、目的格の比率が 20 世紀後半に 50% を上回り、2010 年代には 90% に達するというほぼ同様の増加傾向を示す。Biber et al. (1999) が主に 1980–90 年代のデータに依拠していることを踏まえると、目的格の使用は 2010 年代に至るまで拡大し続けており、「主格と目的格が半々」(Biber et al., 1999, p.336) の状況から目的格への移行がさらに進んだと見られる。

ここからは言い切り型構文 (Kuya, 2021) において目的格への変化の進行が最も早かった I/me について詳しい分析を進める。3 節での議論を踏まえて、**than** 節内の人称代名詞が明らかに主語機能を有する用例に絞った場合、**than I** が 175 例、**than me** が 537 例、合計で 712 例である。I と me の合計を 100 とした時の me の出現率 (以下、「me 率」) は全期間で 75% である。

まず、通時的変遷を 50 年区切り (I: 1820s-60s, II: 1870s-1910s, III: 1920s-60s, IV: 1970s-2010s) で見ると、me 率は III 期に 50% に達し、IV 期には 90% に達している。これをジャンルご

とに分けても、全てのジャンルで IV 期には *me* 率が 90%に達していることから、現在 *me* への移行が完了しつつある段階にあると見られる（ただし *Newspaper* や *Non-fiction Books* は用例が少ないことに注意）。また *Fiction* と *TV/Movies* から用例の多くが採取されている点は、言い切り構文や *It is I/me who* 構文と同じく口語的な文体との関連性を示唆する。

次に、主格から目的格への変化の進行度について、*Than I/me* 構文、言い切り型構文 (Kuya, 2021), *It is I/me who* 構文 (久屋, 2024) の 3 つで比較する。2021 年の COHA 更新前のデータを利用した言い切り型構文 (Kuya, 2021) のデータと対応させるために年代の区切りを調整し、I: 1810s-50s, II: 1860s-1900s, III: 1910s-50s, IV: 1980s-2000s, V: 2010s とした。全ての構文で用例が多く採取されたフィクションに注目すると、*me* 率が 90%程度に達したのが言い切り型構文では III 期、*Than I/me* 構文では IV 期であることから、後者の方で変化がやや遅れて進行していると見られる。一方 *It is I/me who* 構文は最新の V 期 (2010s) に *me* 率が 50%に達しつつある程度であった (久屋, 2024) ことから、*Than I/me* 構文よりも変化が遅れて進行していると見られる。*Than I/me* 構文については *than* が伝統的には接続詞と認識されていること、*It is I/me who* 構文については *relative attraction* と呼ばれる制約が、言い切り型構文に比べて変化の進行速度を遅らせる要因として働いている可能性がある。

## 5. まとめ

本研究は、規範文法の求める格の一致のルールにも関わらず、*Than I/me* 構文でも、主格 *I* から目的格 *me* への変化がいまだ進行中で、現在ではほぼ完了の段階に来ていることを示した。このことは、*than* が接続詞ではなく前置詞として認識される傾向がより一層強まっていることの表れである。ただし、*Than I/me* 構文における変化の度合いは、言い切り型構文よりやや遅れつつも *It is I/me who* 構文よりは進んでおり、また人称代名詞の種類によっても異なると見られる。*Than I/me* 構文の出現するジャンルに偏りが見られた (口語性の高い *Fiction* や *TV/Movies* に多い) ことや、1 人称単数 *I/me* で変化がいち早く進行している可能性のあることについては、Kuya (2021) の結果と通ずるところがある。

今後の課題としては、調査対象を *I/me* 以外の人称代名詞にも拡大し、人称代名詞の種類によって変化の度合いが異なる可能性について改めて検証することや、今回分析に含めなかった (3h-i) のような用例の追加分析 (今回分析した例よりも構造的に目的格使用が出現しやすいか) を行うことが挙げられる。さらに、アメリカ英語で (2c) のような用例の曖昧性を排除するために使われるという *than* 節の述部を省略しない主格形 (例えば *I love you better than him* に対する *I love you better than he does*, Mencken, 1963, p.556) を含めた変異の全体像を把握する必要がある。

## 謝辞

本研究は JSPS 科研費 23K12179 の助成を受けたものである。

## 引用文献

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, F. (1999). *Longman grammar of spoken and written English*. Longman.
- Curme, G. O. (1931). *Syntax*. Heath.
- Davies, M. (2010). *The Corpus of Historical American English (COHA)*. <https://www.english-corpora.org/coha/>. (2024年8-9月アクセス)
- Dekeyser, X. (1975). *Number and case relations in 19th century British English: A comparative study of grammar and usage*. Uitgeverij de Nederlandsche Boekhandel.
- Erdmann, P. (1978). It's I, it's me: A case for syntax. *Studia Anglica Posnaniensia*, 10, 67-80.
- Fowler H. W., & Fowler F. G. (1931). *The king's English* (3rd ed.). Oxford University Press.
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge University Press.
- Jespersen, O. (1894). *Progress in language with special reference to English*. Swan Sonnenschein & Co.
- Jespersen, O. (1949). *A modern English grammar on historical principles, Part VII, Syntax* (Completed and edited by Haislund, N.). George Allen & Unwin.
- Kuya, A. (2021). A corpus-based variationist approach to the use of *it is I* and *it is me*: A real-time observation of a syntactic change nearing completion in COHA. *Gengo Kenkyu*, 159, 7-35.
- 久屋 (2024)「It is I/me who 構文における relative attraction の効果を検証する:最近の動向」『社会言語科学会第48回大会発表論文集』, 95-98.
- Leech, G. & Svartvik, J. (1975). *A communicative grammar of English*. Longman.
- Mencken, H. L. (1963). *The American language: A preliminary inquiry into the development of English in the United States* (Abridged ed.). Alfred A. Knopf.
- Nakayama, M. (2014). *It is I* vs. *It is me* in 19th-century English novels. *Studies in Modern English*, 30, 1-23.
- Oxford University Press. (1989). Than. In *Oxford English dictionary* (2nd ed., p.245).
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A grammar of contemporary English*. Longman.

否定接頭辞 *un-*付加派生語の歴史的考察  
— *Unkempt* の事例を中心に —

岡田 晃 (小山工業高等専門学校)  
akiraokada777@gmail.com

A Historical Study of the Negative Prefix *un-* Derivatives  
— Focusing on the Case of *unkempt* —

OKADA Akira (National University of Technology, Oyama  
College)

Abstract

This presentation explores the historical development of derivational words formed with the negative prefix *un-*, with a particular focus on the word *unkempt*. Although negative prefixes with extensive historical roots have been widely studied, this research shifts the focus to the Old English-derived *unkempt*, examining its emergence, usage patterns, and factors contributing to its decline. Using data from the British National Corpus (BNC), the study analyzes the word's frequency and distribution in contemporary English. Additionally, the Oxford English Dictionary (*OED*) is employed to trace its historical occurrences and shifts in meaning over time.

Keywords

Negative prefix, Historical research, Frequency, OED, BNC

1. はじめに

本発表では、否定接頭辞 *un-*によって派生した語彙 *unkempt* を、主に歴史的視点から考察する。これまで、多くの言語学者が派生語に関する研究を行っており、特に否定接頭辞については長い研究の歴史がある。発表者はこれまでの研究において、形態論における「二重語」を主題に取り上げ、British National Corpus (BNC)を用いて現代英語における使用頻度と使用域を、また Oxford English Dictionary (OED)を用いて過去の使用事例を調査してきた。今回の研究では、特に古英語由来の単語 *unkempt* に焦点を当て、その語がどの時代に使用され、いつから使用頻度が減少していったのかを探る。

言語使用がコミュニケーションの主要な手段であることは広く知られている、言語や語形成に関しては、依然として解明されていない部分が多く残されている。言語使用に関するすべての問題を解決することは容易ではないが、本研究が人々の言語使用をより深く理解し、効果的な言語運用



に寄与する一助となることを期待している。

## 2. BNC で見られる unkempt の頻度及び使用域について

ここでは BNC を使用し、現代英語における unkempt の使用頻度と使用域を調査する。BNC を使用する意義は、現代英語における当該語彙の使用実態の把握と言語変化の追跡が可能となる点にある。下の表 1. は BNC における unkempt の使用頻度と使用域別頻度をまとめたものである。

表 1. Unkempt の使用頻度と使用域

Spoken or Written:				
Category	No. of words	No. of hits	Dispersion (over Frequency per files)	million words
Written	87,903,571	109	97/3,140	1.24
Spoken	10,409,858	0	0/908	0
<b>total</b>	<b>98,313,429</b>	<b>109</b>	<b>97/4,048</b>	<b>1.11</b>
Derived text type:				
Category	No. of words	No. of hits	Dispersion (over Frequency per files)	million words
Fiction and verse	16,143,913	57	48/452	3.53
Other published written material	17,924,109	19	19/710	1.06
Non-academic prose and biography	24,178,674	24	21/744	0.99
Newspapers	9,412,174	5	5/486	0.53
Unpublished written material	4,466,673	2	2/251	0.45
Academic prose	15,778,028	2	2/497	0.13
Other spoken material	6,175,896	0	0/755	0
Spoken conversation	4,233,962	0	0/153	0
<b>total</b>	<b>98,313,429</b>	<b>109</b>	<b>97/4,048</b>	<b>1.11</b>

BNC のデータによれば, *unkempt* の総使用頻度は非常に低いことが確認された。BNC 全体の語彙数約 1 億語のうち, *unkempt* の出現回数はわずか 109 回であり, これは非常に稀な語彙であることを示している。頻度としては, 100 万語あたり 1.11 回という極めて低い数値であり, 日常の会話や一般的なテキストにおいては, この語がほとんど使用されていないことが明らかになった。

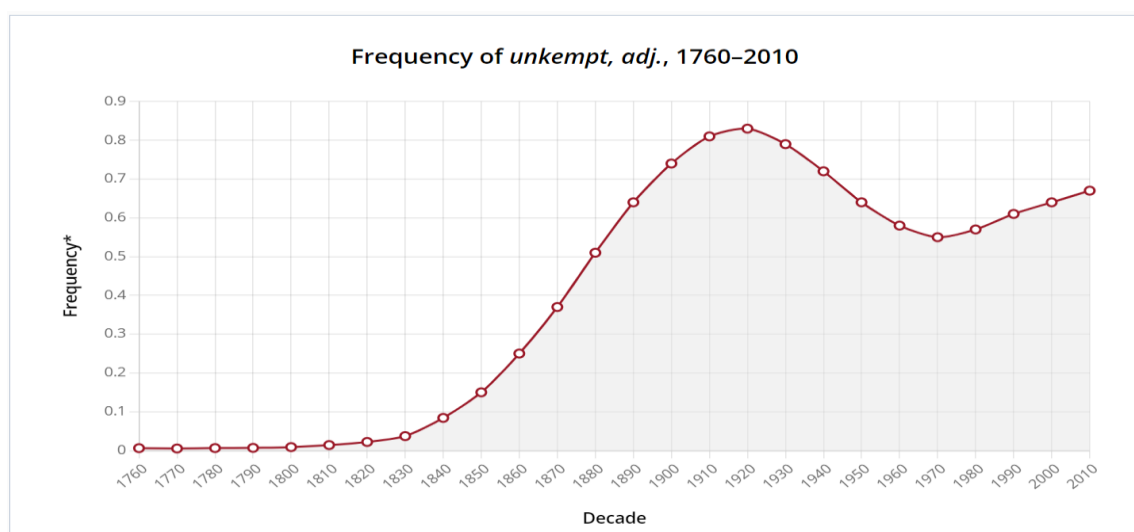
また, 使用域に関しては, 主に「書き言葉」に限定されており, 書き言葉では 109 回の使用が確認されたが, 話し言葉での使用は一度も見られなかった。これは, *unkempt* が一般的な日常会話で使用されることは少なく, むしろ文学作品や学術論文といった書き言葉において主に用いられる語であることを示している。さらに, テキストの種類ごとの使用状況を分析した結果, 特にフィクションや詩のジャンルで最も多く使用されており, 文学的な表現としての *unkempt* の使用が一般的であることが分かった。

以上の分析から, *unkempt* は BNC において極めて稀な語彙であり, 主に書き言葉, 特に文学的な文脈で使用されることが明らかとなった。一方で, 日常会話やその他の話し言葉のテキストではほとんど使用されていない。この結果は, 英単語の使用頻度とその使用文脈を理解する上で重要な示唆を与えるものであり, 今後の言語教育や辞書編纂における基礎的なデータとして活用されることが期待される。

### 3. OED で見られる *unkempt* の年代別使用頻度と使用例について

OED を用いて *unkempt* を通時的に考察する目的は, 単なる語彙の変遷を追跡するだけでなく, 言語の進化, 社会的・文化的背景, 標準化のプロセス, さらに形態論的規則性など, 多角的な視点から英語の歴史を理解することにある。このような研究は, 現代英語の語彙や表現の背後にある複雑な歴史的プロセスを明らかにし, 言語がどのように変化し続けるかを示す重要な手がかりを提供してくれる。以下の図は, OED に見られる *unkempt* の年代別使用頻度の推移を示している。

図 1 OED に見られる *unkempt* の年代別使用頻度の推移



上の図によると、1760年から2010年までの **unkempt** の使用頻度は、いくつかの特徴的な変化を示していることがわかる。以下、その詳細を時代ごとに説明する。

(1) 1760年～1830年

1760年代から1830年代にかけて、**unkempt** の使用頻度は非常に低いまま推移している。この時期にはほとんど使用されておらず、言語的な一般性や普及度は限定的であったと考えられる。

(2) 1830年～1880年

1830年代以降、使用頻度が急速に増加し始め、1870年代から1880年代にかけてピークに達する。この増加は、ビクトリア朝時代における文学や文章表現の変化に影響を受けた可能性がある。特に、19世紀中盤から後半にかけて、詳細な描写や社会的規範に対する言及が増えたことが、**unkempt** の使用頻度の上昇に寄与していると考えられる。

(3) 1880年～1950年

1880年代をピークに、**unkempt** の使用頻度は減少に転じ、1950年代には再び低下する。この時期の減少は、言語使用の変化や、社会的な規範の変動と関連している可能性がある。20世紀前半には、新たな語彙や表現が普及し、**unkempt** のような伝統的な単語の使用が抑えられたと推測される。

(4) 1950年～2010年

1950年代以降、**unkempt** の使用頻度は再び増加し、2000年代に至るまで緩やかな上昇傾向を示している。この再増加の背景には、現代文学やメディアにおける多様な表現の普及、そしてファッションや外見に関する議論が再び注目を集めたことが関係している可能性がある。

**Unkempt** の使用頻度は、1760年から2010年までの期間で大きな変動を経験しており、特に19世紀後半にピークを迎え、その後一時的に減少しつつも、現代に至るまで再び増加傾向にあります。このような使用頻度の推移は、社会的背景や言語文化の変遷を反映しており、語彙使用の歴史的な変化を理解する上で重要な示唆を与えるものである。

#### 4. 結論

**Unkempt** の歴史的な使用の変遷を通じて、言語がどのように時代の文化的背景や社会的価値観に応じて進化してきたかが明らかになる。特定の単語が時代を超えてその意味を変容させる過程は、単に語彙の変化に留まらず、文化の変遷や社会の変動も反映している。現代においても **unkempt** は理解可能な単語であるものの、その使用頻度は減少している。この現象は、言語が常に変化し続ける「生きたシステム」であることを示唆している。

以上のように、**unkempt** という単語の歴史的な使用を探ることによって、英語の語彙とその文化的背景との相互作用を理解する手がかりが得られる。本研究が、英語の言語史研究において一助となることを期待する。

オンライン資料

The British National Corpus URL <http://es-corp.unizh.ch>

*The Oxford English Dictionary online*. URL <http://www.oed.com/>

引用文献

Lieber, Rochelle (2005) *Morphology and Lexical Semantics*. Cambridge: Cambridge University Press.

Lieber, Rochelle (2010) *Introducing Morphology*. Cambridge: Cambridge University Press.

大石 強 (1994) 『形態論』 開拓社.

片見 彰夫, 川端 朋広, 山本史歩子 (2023) (編者) 『英語教師のための英語史』 開拓社.

高橋 英光 (2020) 『英語史を学び英語を学ぶ』 開拓社.

竝木 崇康 (1994) 『語形成』 大修館書店.

西川 盛雄 (2013) 『英語接辞の魅力』 開拓社.

米倉 綽 (2015) 『歴史的にみた英語の語形成』 開拓社.

現代英語の態度マーカー  
—CORE コーパスを用いた 33 種のレジスター別比較に基づいて—

飯島 真之(神戸大学 大学院生)  
christmasayuki98@gmail.com

Attitude Markers in Contemporary English  
—A Comparison between 33 Registers in the CORE Corpus—

IIJIMA Masayuki (Kobe University, Graduate Student)

Abstract

This study aims to clarify the use of attitude markers (AMs) (e.g., *agree*, *important*, and *unfortunately*) across various registers of contemporary English using the CORE corpus. The research has three main objectives: (1) clarifying registers in which AMs are frequently used, (2) determining individual AMs that are both frequently and widely used, and (3) exploring the relationships between registers and AMs. The findings reveal that: (1) AMs are frequently used in registers where the writer's or speaker's stance is likely to be expressed, (2) some verbs (e.g., *agree* and *expected*), adjectives (e.g., *important*), and adverbs (e.g., *even*) are both frequently and extensively used as AMs, and (3) registers and AMs can be categorized based on the generality of the content and the degree to which they are oriented towards speakers/writers or listeners/readers.

Keywords

態度マーカー (AM), レジスター, 頻度, 汎用度

1. はじめに

言語使用の際、話し手や書き手は陳述内容のみならず、心的な態度についても受け手である聞き手や読み手に伝達する。現代英語において、これらはしばしばメタディスコース (metadiscourse) やスタンス (stance) といった枠組みで議論されており (Hyland, 2005; Hyland, 2019 他), 本研究は Hyland がそれらの下位概念の 1 つとして提示する, 態度マーカー (attitude markers) (以下, AM) (*agree*, *important*, *unfortunately* 等) に焦点を当てる。AM は, 受け手への驚き, 同意, 重要性等の伝達に関連し, 言語使用において重要な役割を果たしているが (Hyland, 2019), Mur Dueñas (2010) や Azar and Hashim (2019) を含む先行研究の多くは, 考察対象を研究論文等のかたい書き言葉に限定しており, 異なる言語使用場面, すなわちレジスタータイプによって

AM 使用がどのように変化するかについては必ずしも明らかにしておらず、考察の余地がある。

## 2. 先行研究

以下、コーパスを用いた現代英語における AM 使用に関する複数の研究事例を、研究ごとに知見を整理していく。

表 1 AM に関する先行研究

研究事例	データ	知見
Mur Dueñas (2010)	経営分野の国際学術誌(英語)と国内学術誌(スペイン語)	<ul style="list-style-type: none"> <li>・2 言語ともに、形容詞が高頻度で副詞が低頻度</li> <li>・機能別頻度では、「評価性&gt;重要性&gt;感情性」</li> </ul>
Azar & Hashim (2019)	応用言語学分野のレビュー論文	<ul style="list-style-type: none"> <li>・論文の結論部において AM が多用</li> <li>・AM を, <b>significance, limitations and gaps, emotion, assessment</b> の 4 機能に分類</li> </ul>
Soylu et al. (2023)	教師教育に関する論文(英語母語話者とトルコ語母語話者)	<ul style="list-style-type: none"> <li>・AM 全体使用量は、トルコ語母語話者が英語母語話者を上回る</li> <li>・2 言語共通: 評価性, 重要性 AM が高頻度で使用され, 形容詞を多用</li> <li>・トルコ語母語話者は感情性 AM を過少使用</li> <li>・評価性・重要性 AM については, 英語母語話者よりもトルコ語母語話者のほうが高頻度</li> </ul>

上記の先行研究は、主に特定分野の学術論文における AM 使用実態をコーパス言語学の知見から調査し、著者が読み手に対してどのように心的態度を示しているかを明らかにした点で示唆に富む。しかしながら、学術論文以外の幅広いレジスターにおける AM 使用実態についてもいまだ考察の余地がある。

## 3. リサーチデザイン

### 3.1 研究目的と研究設問

本研究では、現代英語の幅広いレジスターにおける AM 使用実態を明らかにし、L2 スタンス指導のための示唆を得ることを目的とする。本研究における研究設問(RQ)は以下の通りである。

RQ1. AM はどのようなレジスターで多用されるか。

RQ2. 高頻度および高汎用度な AM は何か。

RQ3. AM とレジスターの相互関係性はどのようなようであるか。

### 3.2 コーパス

本研究ではウェブから収集した英語資料を 33 種のレジスター(単一レジスター28 種, 複合型レジスター5 種)に分類した, **Corpus of Online Registers of English (CORE)** を使用する。当該コーパスは総語数 52,933,543 語のコーパスであり, 現代英語コーパスとして広く知られる BNC や COCA ほど大規模なコーパスではないが, レジスター数において, こ

これらのコーパスを上回っており、レジスター研究に適したコーパスといえる。なお、当該コーパスは、コーパス運営サイトである、[English-Corpora.org](http://English-Corpora.org) 上で使用可能である。加えて、本研究では頻度調査を実施する際、同ウェブサイトで使用可能なオンラインインターフェースを使用し、レジスターごとに対象表現の頻度情報を取り出す。

### 3.3 対象とする AM

本研究は、Hyland (2019) が Appendix 中で AM として提示している以下の 65 語（動詞の場合活用形別）をサンプルとして頻度調査を実施する。

admittedly, agree, agreed, agrees, amazed, amazing, amazingly, appropriate, appropriately, astonished, astonishing, astonishingly, correctly, curious, curiously, desirable, desirably, disagree, disagreed, disagrees, disappointed, disappointing, disappointingly, dramatic, dramatically, essential, essentially, even, expected, expectedly, fortunate, fortunately, hopeful, hopefully, important, importantly, inappropriate, inappropriately, interesting, interestingly, prefer, preferable, preferably, preferred, remarkable, remarkably, shocked, shocking, shockingly, striking, strikingly, surprised, surprising, surprisingly, unbelievable, unbelievably, understandable, understandably, unexpected, unexpectedly, unfortunate, unfortunately, unusual, unusually, usual

### 3.4 手法

以下、本研究の手法を説明する。RQ1 では、33 レジスター別に AM の全体使用量を調査する。しかしながら、CORE はレジスターごとに総語数が異なるため、100 万語当たり調整された頻度 (PMW) を用いる。その後 PMW の値の降順にレジスターを並び替える。

RQ2 では、個々の AM ごとに (1) コーパス全体頻度と (2) 33 レジスター別頻度 (いずれも PMW) を調査する。まず、コーパス全体頻度を頻度指標と捉え、値の降順に個々の AM を並び替え、順位付けを行う。次に、33 レジスター別頻度に基づいて、個別 AM ごとに、1 レジスターあたりの平均頻度、及び、標準偏差 (SD) を算出する。SD はレジスター間の頻度のばらつきを示すが、頻度の影響を受けやすいため、その値をさらに平均頻度で割ることで、変動係数 (CV) を算出し、個々の AM の汎用度指標とする。しかしながら、CV の場合、その値が低いほどレジスター間のばらつきが小さい、すなわち、高汎用度であることを示すため、頻度指標と汎用度指標の大小関係が逆になってしまう。そこで、CV の逆数 ( $1/CV$ ) をとることで頻度指標と大小関係を統一する。その後、 $1/CV$  の値の降順に個々の AM を並び替え、順位付けを行う。

上記の手法により、頻度指標、汎用度指標それぞれに基づく順位付けがなされるわけだが、個別 AM ごとに、2 指標それぞれの順位を平均化し、さらにその値を昇順に並び替えて順位付けすることで、2 指標を統合した順位付けが可能となる。

RQ3 では、個別 AM の 33 レジスター別頻度表 (PMW に基づく) を作成し、多変量解析手法の 1 つである、対応分析を実施する。これにより、表に含まれる大量の行列情報を少数の次元 (2 次元) に基づく散布図に整理することが可能となり (石川, 前田, 山崎, 2010),

個別 AM と個別レジスターの関係性を 2 観点で整理することを試みる。ただし、頻度表を作成する際、コーパス全体の粗頻度が 100 未満の *disappointingly*, *expectedly*, *desirably* は除外する。

#### 4. 結果と考察

##### 4.1 RQ1 AM が多用されるレジスター

はじめに、サンプルとした AM (65 語) の全体使用量をレジスター別に比較した結果、高頻度レジスターおよび低頻度レジスターとして以下のものが確認された。下表は頻度上位 5 位と下位 5 位のレジスターと全体使用量 (PMW) の値を整理したものである。

表 2 高頻度上位 5 位と低頻度下位 5 位のレジスター

順位	高頻度レジスター	PMW	順位	低頻度レジスター	PMW
1	Advice	4,600.48	33	Lyrical	2,061.95
2	Review	4,579.86	32	Encyclopedia Article	2,271.43
3	Opinion Blog	4,503.28	31	TV/Movie Script	2,615.22
4	Informational Description+ Opinion	4,316.56	30	Historical Article	2,790.33
5	Interview	4,298.12	29	Informational Persuasion	2,821.47

上表を概観すると、AM は Advice, Review, Opinion Blog, Interview などの受け手を意識した直接的主張が強く反映されるレジスターで多用されている。その一方で、間接的なスタンスを多く含むと考えられる Lyrical や、事実的な内容提示を優先すると考えられる Encyclopedia Article, Historical Article などのレジスターでは低頻度であることがわかる。

##### 4.2 RQ2 高頻度で高汎用度な AM

次に、3.4 節で述べた手法に従い、頻度指標と汎用度指標の両観点を統合し、高頻度で高汎用度な AM を取り出した結果、下表のような結果を得た。下表は、サンプルとして扱った全 65 語の AM のうち、上位 10 語を扱っている。



表3 頻度指標・汎用度指標統合型 AM リスト

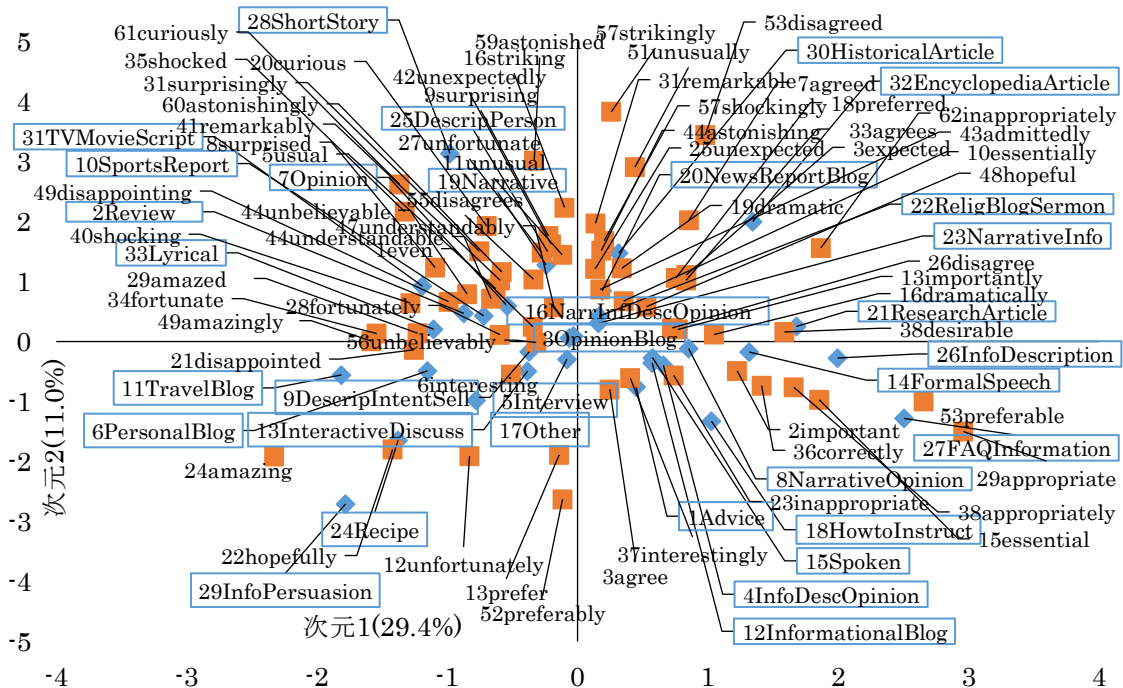
統合順位	頻度順位	汎用度順位	AM	頻度指標 (PMW)	汎用度指標 (1/CV)
1	1	1	even	1374.34	4.72
2	2	3	important	459.91	2.20
3	3	6	agree	237.11	2.03
3	5	4	expected	130.6	2.14
5	13	2	usual	62.74	2.23
6	4	15	interesting	202.26	1.74
7	9	11	agreed	85.45	1.91
8	11	10	surprised	64.4	1.92
9	19	5	surprising	31.45	2.04
10	15	12	essentially	50.7	1.77

上表を概観すると、動詞や形容詞(もしくは形容詞的に機能するもの)が多くを占めていることがわかる。なお、1位の even は頻度指標と汎用度指標の2指標において、他のAMを大きく上回っている。この要因としては、even が有する用法の幅広さが考えられる。

#### 4.3 RQ3 AM とレジスターの相互関係性

最後に、個別AMと個別レジスターの相互関係性を考察するため、3.4節で述べた手法に従い、対応分析を実施した結果、図1の散布図を得た。なお、散布図中において、レジスターにはRQ1の高頻度順位を、AMにはRQ2の2指標(頻度・汎用度)統合型順位を付している。

図1 個別AMと個別レジスターの相互関係性(対応分析)



上図を概観すると、各次元の寄与率の合計は約 40%となっており、作成した頻度表に含まれるデータ全体の 40%が上図において説明されている。レジスターの分布に基づき、各次元（軸）を評価すると、寄与率の高い次元 1（横軸）は、右方に Historical Article, News Report/Blog, Formal Speech などの内容の一般性が高いレジスターが分布し、左方に Short Story, Lyrical, Personal Blog などの内容の一般性が低いレジスターが分布している。一方で、次元 2（縦軸）は、上方に Review, Short Story, News Report/Blog などの発信者志向のレジスターが分布し、下方に Recipe, Informational Persuasion, How-to/Instructional などの受信者志向のレジスターが分布している。

## 5. まとめ

本研究は、現代英語の幅広いレジスターにおける AM 使用実態を明らかにし、L2 スタンス指導のための示唆を得ることを目的とした。

RQ1 では、CORE のレジスターごとに、サンプルとした AM(65 語)の全体使用量を調査した。その結果、AM は受け手を意識した直接的主張が反映されやすいレジスターで高頻度である一方で、直接的な心的態度が反映されにくいと考えられるレジスターや、事実的な内容提示が優先されやすいと考えられるレジスターでは低頻度であることが示唆された。

RQ2 では、頻度指標と汎用度指標を統合し、個別の AM に対し順位付けを行った。その結果、上位 10 語の多くが動詞や形容詞として(あるいは形容詞的に)機能することが示された。なお even が 1 位の理由としては、副詞としての用法の幅広さによるものと考えられる。

RQ3 では、個別 AM のレジスター別頻度表に対して対応分析を行った結果、AM とレジスターの相互関係性が、(1)内容の一般性の程度、(2)発信者・受信者志向の程度の 2 観点によって説明できる可能性が示された。

RQ1 から RQ3 を通じて明らかになった知見を組み合わせることで、レジスタータイプ別にどのような AM が重要となるかについても考察が可能となる。これらの知見は L2 英語学習者に向けた、スタンス表現に関する教材開発に一定の示唆を与えるのではないだろうか。

一方で、本研究には課題も存在し、(1)対象とした AM はあくまでもサンプルであり、すべての AM を扱っているとは言えないこと、(2)レジスタータイプの 4 分類は対応分析の手法的影響が大きいことなどがあげられる。そのため、これらについては再度検証の余地があると考えられる。

## 引用文献

- Azar, A. S., & Hashim, A. (2019). The impact of attitude markers on enhancing evaluation in the review article genre. *GEMA Online Journal of Language Studies*, 19(1), 153-173.
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7, 173-192.
- Hyland, K. (2019). *Metadiscourse: Exploring interaction in writing*. Bloomsbury.

石川慎一郎・前田忠彦・山崎誠 (2010). 『言語研究のための統計入門』. くろしお出版.

Mur Dueñas, P. (2010). Attitude markers in business management research articles: A cross-cultural corpus-driven approach. *International Journal of Applied Linguistics*, 20(1), 50-72.

Soylu, M., Soylu, A., & Das, R. (2023). A new approach to recognizing the use of attitude markers by authors of academic articles. *Expert Systems with Applications*, 230(19).

コーパス

CORE. English-Corpora.org. <https://www.english-corpora.org/core/>

# 日本人英語学習者の発話における フィラーの生起位置と習熟度の関係性

田中 悠介(福岡大学)

瀬戸口 彩花(京都大学 大学院生)

近 大志(関西大学)

神澤 克徳(京都工芸繊維大学)

{yusuke.tanaka.07, guchi.a.chan7, nanou7614, rnkp43470}@gmail.com

## Relationship Between the Occurrence Positions of Fillers and Proficiency Levels in the Speech of Japanese English Learners

TANAKA Yusuke (Fukuoka University)

SETOGUCHI Ayaka (Kyoto University, Graduate Student)

CHIKA Taishi (Kansai University)

KANZAWA Katsunori (Kyoto Institute of Technology)

### Abstract

This study utilized the KIT Speaking Test Corpus to investigate how the placement of fillers in different sentence positions—sentence initial, clause boundary, phrase boundary, and within a phrase—affects the fluency ratings of utterances by Japanese learners of English. The findings indicate a negative correlation between the frequency of fillers at the beginning of sentences and the learners' test scores. Sentence initiation requires critical decision-making about message content, vocabulary, and syntactic structure. Learners with higher proficiency levels are more adept at making these decisions, resulting in fewer fillers at sentence commencements. This suggests that fillers located at sentence starts, unlike those in other positions, might be perceived more strongly as indicators of disfluency.

### Keywords

KIT Speaking Test Corpus, Fillers, Disfluency phenomenon, Proficiency levels

### 1. はじめに

学習者のスピーキング能力を測る指標の1つに、「流暢さ」がある(Ellis 2003)。したがって、これを阻害する要因である「非流暢性現象」(Biber et al. 1999)は、英語学習者のスピーキング能力を評価する上で重要な役割を果たす。本研究では、非流暢性現象の1つである「フィラー」に焦点を当てる。フィラーとは、「発話の隙間を埋めるのに用いられる音声要素で、それ自体は発話の意味

内容・メッセージ内容に関わらないもの」(横森他 2014)であり、日本語の「えっと」や英語の *you know* などが例として挙げられる。

非流暢性現象の 1 つとされるフィラーであるが、これが実際に非流暢性の指標となり得るかについては立場が別れている。実際、フィラーは必ずしも流暢性を阻害するわけではない。流暢性を阻害するものとして知覚されないフィラーは「平常的非流暢性」と呼ばれる(Biber et al. 1999)。フィラーは母語話者でも一定の割合で産出するものであり、フィラーの適切な使用はむしろ流暢さの評価を高める可能性がある。このことから、フィラーを流暢さの指標とみなすべきという主張もある(Tottie 2014)。実際、フィラーが流暢さの指標である可能性を示す調査結果がある。Rieger (2003)は英語を母語とするドイツ語学習者の発話を分析し、習熟度の高い学習者ほどフィラーの使用頻度が高いことを明らかにしている。Cenoz (1998)もスペイン語を母語とする英語学習者の分析から、同様の傾向を報告している。一方で、フィラーが非流暢性の指標である可能性を示す調査結果もある。Lennon (1990)はイギリス留学中のドイツ語を母語とする中級英語話者の 2 週目と 23 週目を比較し、フィラーの減少を観察している。また、田中他 (2023)は KIT Speaking Test Corpus (KISTEC)を利用したコーパス調査から、スピーキングテストの得点とフィラーの使用頻度に負の相関があることを明らかにした。これは、習熟度が低い学習者の方がフィラーを多用する傾向があることを示している。

以上のとおり、フィラーが流暢性と非流暢性のどちらの指標とみなされるべきかについては議論が続いているが、流暢さに対するフィラーの影響はより複雑な可能性がある。Kahng (2018)は流暢さの評価に対するポーズの影響が生起位置によって異なる可能性を検証するために、英語母語話者と韓国語を母語とする英語学習者の発話における節の境界または内部に人為的にポーズを挿入した。その結果、母語話者と学習者いずれの発話においても、節の内部にあるポーズは境界にあるポーズよりも大きく流暢さの評価を下げるということがわかった。

Kahng (2018)が分析対象としているのはフィラーではなくポーズであるが、フィラーについても同様に、生起位置によって流暢さに与える影響の方向性や大きさが異なる可能性がある。本研究ではこの可能性を検証するべく、以下の 3 つの仮説を立てる。

- 仮説 1 文頭は伝達する内容や使用する語彙、構文など、さまざまな選択をする必要がある位置であり、母語話者でもフィラーが生じやすい位置であるため、文頭で生じるフィラーは非流暢性現象として知覚されにくい
- 仮説 2 節境界や句境界に生起するフィラーは聞き手に統語構造を理解するための手がかりと時間を与えるため、非流暢性現象として知覚されにくい
- 仮説 3 句内部では統語的なまとまりの中にフィラーが生じることになるため、他の生起位置に比べてより強く非流暢性現象として知覚される

以上の仮説が妥当であれば、文頭や節境界、句境界で生じるフィラーは流暢さの評価に影響を与えない一方で、句内部におけるフィラーは負の影響を与えると予測される。次節では、これらの

仮説を検証するために実施したコーパス調査を報告する。

## 2. コーパス調査

### 2.1 使用したコーパス

本研究では、フィラーの収集に KISTEC を利用した。理由としては、各受験者にスピーキングテストの得点が付与されているため習熟度に基づく分析が可能であることと、フィラータグが付与されているためフィラーの収集が容易であることの 2 点が挙げられる。KISTEC は 2018 年に京都工芸繊維大学の 1 年生が受験した CBT 方式の英語スピーキングテストの回答音声を書き起こしたものであるが、詳細については <https://kitstcorpus.jp/> を参照されたい。

### 2.2 手続き

ランダムサンプリングした 100 名のデータを調査対象とし、フィラータグが付与されている 4463 件を、生起位置に基づき 8 つに分類した(表 1)。4 人の著者が分担して分類した後に、相互チェックを実施した。

表 1 生起位置とそれぞれの例

生起位置	例
文頭	<F>Mm</F> here is <F>uh</F> my house.
節境界	<R>If</R> if I am Mariam, <F>eh</F> I try to work more faster than before.
句境界	At first I <SC>call a friends</SC> call friends <F>uh</F> to join a practice cycling.
句内部	<F>Eh</F>, this is <SC>why I</SC> <F>eh</F> one reasons why I want to go there with <F>eh</F> some friends.
自己訂正の前	<F>Eh</F>, this is <SC>why I</SC> <F>eh</F> one reasons why I want to go there with <F>eh</F> some friends.
繰り返しの前	But, <F>eh</F> and the <R>dream</R> <F>eh</F> dream <F>eh</F> is now <F>eh</F> I have.
独立	<F>Eh</F>.
その他	<CO>So he <F>ah</F> so</CO>

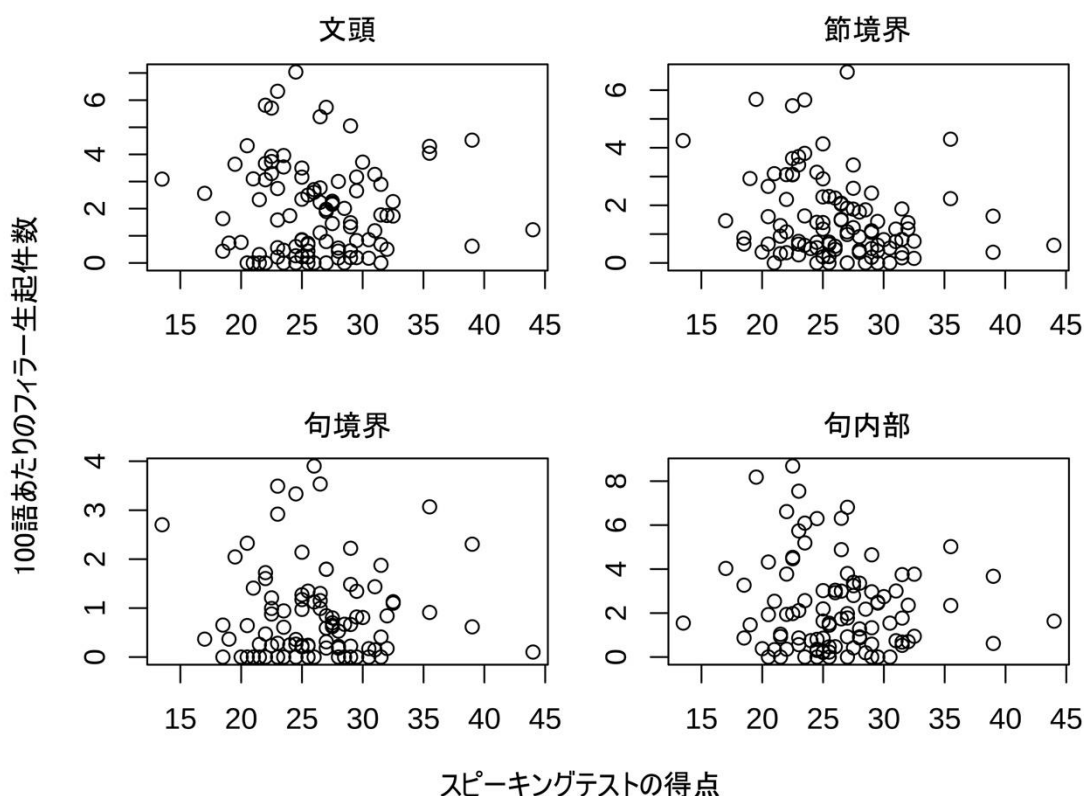
本研究では、フィラーを統語的な観点から文頭、節境界、句境界、句内部の 4 つに分類した。また、分析対象外のフィラーとして、自己訂正の前、繰り返しの前、独立、その他の 4 つを設けた。自己訂正とは、受験者自身が発話を訂正するものである。表 1 の例では、“why I” と発話した後に “one reasons” を追加するという訂正を行っている。このような自己訂正の前に生起しているフィラーは、どのように訂正するかを考える時間を稼ぐために使用されているという意味で、その他のフィ

ラーと同等に扱うべきではないため、分析から除外した。また、フィラーが連続している場合は、2つ目以降のフィラーは分類対象外とした。つまり、フィラーが連続していても、その位置での生起回数は1回とカウントするということである。繰り返しはフィラーと同様の役割を担うことがあるため、繰り返しに後続するフィラーも分析対象外とした。また、フィラーのみで発話が完結しているなど、4つの生起位置のいずれにも当てはまらないものは「独立」として分析から除外した。最後に、主に統語構造の崩壊が原因で生起位置が特定不可能なものを「その他」として分析から除外した。

### 3. 結果と考察

分析対象とした100名のスピーキングテストの得点と各生起位置における100語あたりのフィラーの生起件数は図1のとおりである。KISTECには「タスク達成度」と「タスク伝達度」という2つの評価項目があるが、後者が主に流暢性の評価をしているため、分析には後者の得点を用いる。

図1 スピーキングテストの得点と100語あたりのフィラー生起件数の関係性



スピーキングテストの得点を目的変数、4つの生起位置における100語あたりのフィラー生起件数を説明変数とする線形モデル分析を行った。その結果、文頭の主効果のみが統計的に有意であった(表2)。この結果は、文頭のフィラーはスピーキングテストの得点に負の影響を与えるが、その他の生起位置におけるフィラーは影響を与えないことを示唆している。

表 2 線形モデル分析の結果

生起位置	推定値	標準誤差	t 値	p 値
文頭	-1.518	0.511	-2.973	0.004
節境界	1.104	0.577	1.913	0.059
句境界	-0.109	0.429	-0.253	0.801
句内部	-0.449	0.775	-0.579	0.564

以上の結果から、仮説 2 は支持された一方で、仮説 1 と仮説 3 は支持されなかった。まずは仮説 1 について考察する。文頭は伝えたい内容や使用する語彙、構文など、さまざまな選択をする必要がある位置であるため、母語話者であってもフィラーが生じやすい位置である。このことから、当初は文頭でのフィラーは非流暢性現象として知覚されにくく、得点に影響を与えないと予測していたが、実際には負の相関が観察された。これは、内容等の選択を素早く行い文をスムーズに始められる学習者は流暢な印象を与えることができたのだと考えられる。対照的に、文頭で詰まりフィラーを使用することは、非流暢性をより強く印象付けるのだと思われる。

次に仮説 2 について、節境界と句境界のフィラーは統語構造を理解するための手がかりと時間を与えるため非流暢性現象として知覚されにくく、得点に対する影響はないと予測していた。結果はこの予測のとおり、得点に対する節境界と句境界のフィラーの影響は認められなかったが、これには別の理由が考えられる。すなわち、KISTEC におけるスピーキングテストの受験者の習熟度が全体としてあまり高くないということである。母語話者と比べて発話語数も少なく、統語構造も複雑ではないため、節境界と句境界にフィラーがあっても、それが理解の助けとなることはなかったのだと推測される。

最後に仮説 3 について、句内部でのフィラーは予測に反し、得点に影響を与えなかった。これについては、句内部におけるフィラーの生起頻度の低さが関係している可能性がある。文頭(751 件)、節境界(994 件)、句境界(1126 件)と比べて、句内部は 436 件と生起件数が少ない。Kahng (2018) の実験では人為的に(本研究でいうところの)句内部にポーズを挿入していたが、実際の発話では句内部のフィラーやポーズはあまり見られないため、流暢さの評価に対する有意な影響が観察されなかった可能性がある。また、ポーズとフィラーの違いが関係している可能性もあり、この点については更なる調査が必要であると考えられる。

#### 4. おわりに

本研究は、フィラーの生起位置と習熟度の関係性を明らかにするために KISTEC を使用したコーパス調査を実施した。その結果、文頭におけるフィラーはスピーキングテストの得点に対して負の影響がある一方で、節境界、句境界、句内部におけるフィラーの統計的に有意な影響は観察されなかった。文頭は伝達する内容から使用する語彙、構文など、さまざまな選択を行う必要がある位置であるため、文頭でのフィラーの使用は他の位置での使用と比較して流暢性が低いことをより強く印象付けるのだと考えられる。



最後に、本研究の限界について述べる。本研究は文や節、句の生起数をカウントしていない。したがって、例えばスピーキングテストの得点と文頭におけるフィラーの生起件数に負の相関があるというも、得点の低い学習者は一文が短いが故に文の数が多く、それに応じて文頭のフィラーも多くなっているだけという可能性も否定できない。より厳密な分析を行うためには文や節の生起件数をカウントする必要があるため、今後の課題としたい。

## 謝辞

本研究は、文部科学省科学研究費補助金「英語スピーキングテストの採点効率化に向けた Task Achievement の解明」(課題番号:22K00736)を受けたものです。

## 引用文献

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. London: Longman.
- Cenoz, J. (1998). Pauses and communication strategies in second language speech.
- Ellis, R. (2003). *Task-Based Language Learning and Teaching*. Oxford: Oxford University Press.
- Kahng, J. (2018). The effect of pause location on perceived fluency. *Applied Psycholinguistics*, 39, 569–591.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40, 387–417.
- Rieger, C. L. (2003). Disfluencies and hesitation strategies in oral L2 tests. In Eklund, R. (ed.), *Gothenburg Papers in Theoretical Linguistics 90*, pp. 41–44.
- 田中悠介・瀬戸口彩花・近大志・神澤克徳 (2023).「日本語を母語とする英語学習者が使用するフィラーの分析:習熟度との関連性および英語母語話者との比較から」『英語コーパス学会大会予稿集 2023』13–18.
- Temple, L. (1992). Disfluencies in learner speech. *Australian Review of Applied Linguistics*, 15, 29–44.
- Tottie, G. (2014). On the use of *uh* and *um* in American English. *Functions of Language*, 21, 6–29.
- 横森大輔・遠藤智子・河村まゆみ・鈴木正紀・原田康也 (2014).「日本語を第一言語とする英語学習者の比較的自発的な発話におけるフィラーに見られるいくつかの特徴」『日本英語教育学会第 43 回年次研究集会論文集』89–96.

---

『英語コーパス学会大会予稿集 2024』 (ISSN 2436-6647)

*Proceedings of the JAECs Conference 2024*

刊行日 2024年 10月 5日

発行所 英語コーパス学会

事務局 〒464-8601 名古屋市千種区不老町 名古屋大学言語教育センター 小島ますみ研究室気付

e-mail: jaecs.hq@gmail.com URL: <https://jaecs.com/>