



日本人学習者の英作文を特徴づける語彙文法指標と スタンス表出

飯島, 真之

(Citation)

統計数理研究所共同研究リポート, 481:39-56

(Issue Date)

2025-03-25

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/0100493758>

(URL)

<https://hdl.handle.net/20.500.14094/0100493758>



日本人学習者の英作文を特徴づける語彙文法指標とスタンス表出

飯島 真之(神戸大学・大学院生)

christmasayuki98@gmail.com

Linguistic Features and Stance Expressions in English Essays Written by Japanese Learners of English

Iijima, Masayuki (Kobe University, Graduate Student)

概要

本研究は、日本人学習者の英作文に見られるスタンス表出の全体像を、現代英語を構成する様々な語彙文法指標の評価を踏まえて明らかにすることを試みる。データとして、アジア圏国際英語学習者コーパスである ICNALE の英作文モジュールである **Written Essays**(2.6)を使用する。本研究では、英作文データのトピックを「大学生アルバイトの是非」に限定し、個々の語彙文法指標の評価には、**Multidimensional Analysis Tagger (MAT)**を用いる。結果として、日本人学習者のスタンス表出が、(1) 書き手自身の観点を直接的に反映すること、(2) 読み手に対して直接的な内容提示を行う一方で、習熟度の上昇に伴い中立的視点を反映するようになること、(3) 読み手に直接的である点が、アジア圏内で一定の地域的共通性を有するが、議論が一般化されにくいこと等の特徴を有する可能性が示された。

キーワード

日本人学習者、語彙文法指標、スタンス

1. はじめに

我々は言語使用の際、スタンス(stance)、すなわち自身の心的態度を自然に表出するが、こうした現象は英語学や英語教育学においても広く議論されてきた。言語使用時のスタンス表出に関する研究の多くは、事前にスタンスの枠組みとその下位項目を設定し、その枠組みの中で、各種のスタンス標識に関して評価や考察を行っている。とくに、Hyland が示すスタンスの枠組みは、広く取り入れられており、驚き・同意・重要性等の態度表出に関与する「態度マーカー」(attitude markers)、命題への関与を保留する「ヘッジ」(hedges)、命題に強意を付与する「ブースター」(boosters)、書き手や話し手を明示的に表出する「自己言及」(self-mentions)という 4 種の下位概念は(Hyland, 2005; Hyland, 2019 他)、スタンス表出の中心的な機能として、多くの研究で調査されている。

ただ、スタンスという概念の明確な定義は難しく、どのような枠組みを立てたとしても、そこからこ

ばれ落ちる要素が存在することは否定できない。この点を踏まえれば、既存の枠組みを導入し、その枠内でスタンスを議論するだけではなく、あえて事前の枠組みを持たず、スタンスに関連する可能性のある要素を網羅的かつ悉皆的に調査するというアプローチが想定される。

本研究では、アジア圏国際英語学習者コーパスである ICNALE に含まれる英作文データをサンプルとして、過去の言語学研究においてテキストの言語特性を表す指標とみなされてきた 65 種の語彙文法指標 (Biber, 1988 他) の出現状況を網羅的に調査する新しいアプローチを取ることで、L2 英語学習者のスタンス表出を考察することとしたい。

2. 先行研究

L2 英語学習者の言語特性を調査するにあたり、アジア圏の学習者が使用する語彙文法指標の特徴について、先行研究の知見を概観する。

Ishikawa (2024) は、アジア圏国際英語学習者コーパス ICNALE の英作文データを使用し、日本人を含むアジア圏地域の学習者が使用する個々の語彙文法指標に関して、英語母語話者と比較しながら考察した。なお、当該研究では Multidimensional Analysis Tagger (MAT) を使用した調査が行われた。結果として、アジア圏学習者が使用する語彙文法指標の全般的な特徴として、内容の対話性 (conversationality)、様々なモダリティ使用 (mixed modality)、内容提示の断片性 (fragmentality)、単純な構造 (structural simplification) といったものが示された。また、対応分析を通じ、地域別学習者と使用される個々の語彙文法指標の関係性を調査し、日本人学習者によって使用される語彙文法指標が口語的 (colloquial) で個人的 (personal) であることを示した。

Li (2023) は、学習者コーパスの TECCL、WECCL2.0 から得た中国人学習者の英作文データと ICNALE から得た比較用の母語話者データを使用し、Biber (1988) の MDA (Multidimensional Analysis) に基づく 6 次元のうち 5 次元の程度に関して、個々の語彙文法指標の調査を踏まえ考察を行った。なお、当該研究においても MAT を用いて調査が行われた。結果として、中国人学習者は英作文中で、(1) 関与性 (involvement) と情報説明性 (informativeness) が高いこと、(2) 内容を語らない非物語的関心 (non-narrative concern) が高いこと、(3) 明示的言及 (explicit reference) が多いこと、(4) 説得性 (persuasiveness) が低いこと、(5) 抽象的情報 (abstract information) を提示することといった特徴を有することが示された。なお、(3) の明示的言及に関しては、学習者も母語話者も頻度が高いが、句の並列 (phrasal coordination) や名詞化 (nominalization) の傾向は母語話者以上に学習者に目立つことが示された。また、(5) の抽象的な情報提示に関しても、学習者と母語話者は共に同程度であることが指摘されているが、学習者は接続詞 (conjunct) を多用する一方で、母語話者は受動態等を多く使用することが示唆された。

上記の先行研究は、学習者コーパスのデータに含まれる語彙文法指標の特徴を、計量的手法を踏まえて考察している点において、多くの示唆に富む。Ishikawa (2024) は、アジアの地域別学習者を広く国際比較を通じて考察しており、Li (2023) は中国人学習者を主たる関心の対象とし、

特定地域の学習者に関する詳細な語彙文法指標の特徴を考察している。

本研究は、ICNALE の英作文データを使用し、日本人学習者が使用する語彙文法指標の特徴、及びスタンス表出傾向を、アジアの EFL 圏内の国際比較を通じて行うが、Ishikawa (2024) において、本研究と共通するデータを使用した研究がすでに実施されている。しかしながら、当該研究では、対象とした学習者・母語話者グループすべてに頻度が確認される 49 種の語彙文法指標に対象項目が限定されている。また、当該研究は日本人学習者を含め、アジア圏の地域別学習者を総合的に考察したものであり、必ずしも日本人学習者を議論の中心として位置付けているわけではない。そこで、本研究では、Ishikawa (2024) と同様のデータを使用しつつも、(1) 一部の地域別・習熟度別学習者データに頻度が確認されない指標も含め、より幅広い指標を対象とすること、(2) 国際比較の観点を取り入れつつも、日本人学習者を議論の中心としてリサーチデザインを構築することの 2 点に考慮する。

3. リサーチデザインと手法

3.1 研究目的と RQ

本研究では、現代英語を成り立たせる多様な語彙文法指標 (計 65 種) を網羅的に調査するというアプローチを取り、日本人学習者の英作文に見られる語彙文法指標、及びスタンス表出傾向を分析する。

なお、Granger (2015) が提唱する改訂版中間言語対照分析 (Contrastive Interlanguage Analysis: CIA²) では、従来のモデルで NL (native language) と IL (interlanguage) と呼ばれていたものが、新たに RLV (reference language varieties) と ILV (interlanguage varieties) と置き換えられ、参照データ (母語話者データを含む) と学習者データの両面において、内部的な多様性に配慮した比較を行うことが推奨されている。この理念を取り入れ、本研究において、母語話者の産出については職業属性別に分析するとともに、学習者については対象学習者内の習熟度の多様性に加え、母語を異にする同じアジア圏学習者との比較を取り入れる。

以上の目的に沿い、以下の 3 種の研究設問 (RQ) を設定する。

母語話者と比較した場合：

RQ1 日本人学習者と母語話者が全体として多用する語彙文法指標には違いがあるか。

(日本人と母語話者の全体比較)

RQ2 4 段階の習熟度段階別日本人学習者と、3 種の職業属性別母語話者が多用する語彙文法指標にはどのような違いがあるか。

(習熟度・職業属性別比較)

RQ3 日本人学習者と習熟度を等しくする 5 か国の EFL 圏学習者が多用する語彙文法指標にはどのような違いがあるか。

(地域別比較)

本研究では以上の 3 種の RQ に基づき、計量的手法を踏まえて調査を実施する。各 RQ の調査を通じて、日本人学習者と母語話者、及び他の EFL 圏地域の学習者の間には、多用される語彙文法指標、及びスタンス表出傾向にどのような差があるかを考察する。

3.2 データ

本研究では、神戸大学の石川慎一郎研究室において開発された、アジア圏国際英語学習者コーパス(International Corpus Network of Asian Learners of English: ICNALE)の英作文モジュールである Written Essays(2.6) (以下、WE)を使用する。当該モジュールは日本人を含むアジアの地域別英語学習者の意見英作文のデータを収集したものである。WE に収録されている作文データは、(1)作文のトピックの統制(「大学生アルバイトの是非」と「レストランにおける全面禁煙の是非」の 2 種)、(2)1 作文当たりの執筆時間(20 分から 40 分)や必要語数(200 語から 300 語)の事前設定、(3)辞書使用の禁止等、徹底的な条件統制の下で収集され、さらに、各参加者は TOEIC 等の語学試験や語彙テスト等を踏まえ、4 段階の習熟度(CEFR に基づく A2、B1_1、B1_2、B2+)に区分されている(Ishikawa, 2023)。

本研究では、日本、インドネシア、中国、台湾、韓国、タイのアジアの EFL 圏 6 地域の学習者による作文データ、及び母語話者の作文データ(大学生、教師、その他職業の 3 職業属性)を使用する。加えて、本研究はスタンスに焦点を当てることから、個人的な体験や主張がより反映されやすいと推測される「大学生アルバイトの是非」のトピックの作文に分析の対象を絞る。

以下は WE のうち、本研究が使用する地域別学習者の作文数を整理した表である。なお、下表では、本研究が使用しないデータ項目については作文数を提示していない。

表 1. 地域別・習熟度別学習者データに含まれる作文数

地域別学習者	A2	B1_1	B1_2	B2+
日本(JPN)	154	179	49	18
インドネシア(IDN)		82	83	
中国(CHN)		232	105	
台湾(TWN)		87	61	
韓国(KOR)		61	88	
タイ(THA)		179	100	
母語話者(ENS)	学生:100 教師:44 その他:56			

本研究において、日本人以外の学習者データについては、地域別学習者間の比較(RQ3)で使用する。その際、より妥当な地域間比較を実施するため、学習者の習熟度を B1 レベル(B1_1 と B1_2 を統合)に限定する。

3.3 手法

3.3.1 データの事前処理: MAT による語彙文法指標の評価

本研究では、Multidimensional Analysis Tagger (MAT) (v. 1.3.3) を使用し、地域別学習者や母語話者による作文に含まれる語彙文法指標を評価する。MAT は、Biber (1988) の知見に基づき、入力した英語のテキストデータに含まれる、多様な語彙文法指標 (計 67 種) の使用量を計量的に評価するソフトである。なお、当該ソフトはウェブサイトからダウンロードして使用することが可能である。

下表は MAT が評価対象とする 67 種の語彙文法指標を整理したものである。なお、下表の内容は、当該ソフトウェアの公開マニュアルである Nini (2019b)、及び Santorini (1990) の内容を参照しつつ整理したものである。

表 2. MAT が評価対象とする個別語彙文法指標

1	AMP	amplifiers
2	ANDC	independent clause coordination
3	AWL	word length
4	BEMA	<i>be</i> as main verb
5	BYPA	<i>by</i> passives
6	CAUS	causative adverbial subordinators
7	CONC	concessive adverbial subordinators
8	COND	conditional adverbial subordinators
9	CONJ	conjuncts
10	CONT	contractions
11	DEMO	demonstratives
12	DEMP	demonstrative pronouns
13	DPAR	discourse particles
14	DWNT	downtoners
15	EMPH	emphatics
16	EX	existential <i>there</i>
17	FPP1	first person pronouns
18	GER	gerunds
19	HDG	hedges
20	INPR	indefinite pronouns
21	JJ	attributive adjectives
22	NEMD	necessity modals
23	NN	total other nouns

24	NOMZ	nominalizations
25	OSUB	other adverbial subordinators
26	PASS	agentless passives
27	PASTP	past participial clauses
28	PEAS	perfect aspect
29	PHC	phrasal coordination
30	PIN	total prepositional phrases
31	PIRE	pied-piping relative clauses
32	PIT	pronoun <i>it</i>
33	PLACE	place adverbials
34	POMD	possibility modals
35	PRED	predicative adjectives
36	PRESP	present participial clauses
37	PRIV	private verbs
38	PRMD	predictive modals
39	PROD	pro-verb <i>do</i>
40	PUBV	public verbs
41	RB	total adverbs
42	SERE	sentence relatives
43	SMP	<i>seem, appear</i>
44	SPAU	split auxiliaries
45	SPIN	split infinitives
46	SPP2	second person pronouns
47	STPR	stranded preposition
48	SUAV	suasive verbs
49	SYNE	synthetic negation
50	THAC	<i>that</i> adjective complements
51	THATD	subordinator <i>that</i> deletion
52	THVC	<i>that</i> verb complements
53	TIME	time adverbials
54	TO	infinitives
55	TOBJ	<i>that</i> relative clauses on object position
56	TPP3	third person pronouns
57	TSUB	<i>that</i> relative clauses on subject position
58	TTR	type-token ratio

59	VBD	verb, past tense
60	VPRT	verb, present tense
61	WHCL	<i>WH</i> -clauses
62	WHOBJ	<i>WH</i> relative clauses on object position
63	WHQU	direct <i>WH</i> -questions
64	WHSUB	<i>WH</i> relative clauses on subject position
65	WZPAST	past participial WHIZ deletion relatives
66	WZPRES	present participial WHIZ deletion relatives
67	XX0	analytic negation

※Nini (2019b)、Santorini (1990) 参照

MAT は上表の語彙文法指標を対象とするが、これらの指標を概観すると、AWL(ワード長)やTTR(タイプ・トークン比率)等の指標も含まれている。これら 2 種の指標は、英語における個々の表現形式を直接的に評価した指標ではないため、研究対象から除外する。これにより、最終的に 65 種の語彙文法指標が対象となった。

本研究では、事前にダウンロードした WE の地域別学習者及び母語話者の作文データを MAT に入力することで、個々の作文にみられる各語彙文法指標の使用量を計量的に調査する。なお、MAT を使用することで、(1) 指標別相対頻度表(100 語当たり)と(2) Biber (1988) のコーパスデータを基準とした指標別標準化得点 (z score) 表の 2 種の表がエクセルで自動作成される。なお、(2) の標準化得点については、指標別に以下の式によって自動計算される。式中の X は入力したテキストデータ中の各語彙文法指標の相対頻度を、 μ は Biber (1988) のコーパスにおける各指標のテキスト間平均頻度を、 σ は Biber (1988) のコーパスにおける各指標の標準偏差を示している (Nini, 2019a)。

$$Z_x = \frac{X - \mu}{\sigma}$$

上記の標準化得点を使用することで、本来頻度比較が妥当ではない異なる語彙文法指標間の比較が可能となる。なお、本研究では、RQ1 では(2)の標準化得点表のデータを使用し、RQ2・3 では(1)の相対頻度表のデータを使用する。

3.3.2 分析の手順

以下、RQ 別に手法を概観する。RQ1 では、日本人学習者(すべての習熟度を含む)、母語話者(すべての職業属性を含む)別に、作文別の語彙文法指標別標準化得点 (z score) を算出する。その後、指標別に、1 作文当たりの得点の平均値(平均スコア)を算出し、その値を降順に並べ替えることで、日本人学習者と母語話者それぞれに特徴的な指標の特定を試みる。

RQ2 では、日本人学習者の作文データを 4 種の習熟度別 (A2、B1_1、B1_2、B2+) に、母語話者の作文データを 3 種の職業属性別 (学生、教師、その他) に区分し、7 群のデータセットを作成する。その後、7 群別に、MAT を使用し、各作文における語彙文法指標別相対頻度 (100 語当たり) を算出し、データセットごとに指標別 1 作文当たりの平均頻度を算出する。その後、7 群それぞれに関する指標別平均頻度を整理した頻度表を作成し、多変量解析手法の 1 種である対応分析 (コレスポンデンス分析) を実施する。

RQ3 では、地域別学習者のデータセット (6 群)、職業属性別母語話者のデータセット (3 群) の計 9 群のデータセットを作成する。この際、各地域ともに、学習者の習熟度を B1 レベル (B1_1 と B1_2) に限定する。その後、RQ2 と同様に、9 群別に、1 作文当たりの語彙文法指標別平均頻度を算出する。その後、各群の指標別平均頻度を整理した表を作成し、対応分析を実施する。

以上が、本研究の各 RQ における手法である。なお、統計手法については無償で使用可能な統計プログラムである HAD を使用する。

4. 結果と考察

4.1 RQ1 日本人学習者と母語話者が多用する語彙文法指標

はじめに、日本人学習者全体と母語話者全体を比較し、各々が多用する語彙文法指標を比較する。手法に従い、1 作文当たりの平均スコアの高い上位 15 位の語彙文法指標を取り出したところ以下ようになった。なお、当初研究対象とした 65 種の指標のうち、SPIN については、一部の作文に対して極端に高いスコアが示され、平均値比較の妥当性が低いと判断し、当該指標を除いたうえで結果を提示する。また、表中では、上位 15 位の範囲内において、日本人学習者のみ、及び母語話者のみに見られる指標をボールドで示している。

表 3. 日本人と母語話者が好む語彙文法指標 (上位 15 位)

順位	日本人学習者全体		母語話者全体	
	指標	平均スコア	指標	平均スコア
1	PRED	4.87	PRED	3.51
2	CONJ	3.21	TO	3.33
3	TO	3.21	TSUB	2.67
4	POMD	2.96	OSUB	2.67
5	NEMD	2.76	TOBJ	2.28
6	WHCL	2.09	CONJ	1.94
7	CAUS	1.74	POMD	1.85
8	FPP1	1.48	COND	1.75
9	COND	1.44	EMPH	1.71
10	TOBJ	1.19	CAUS	1.55

11	THVC	0.97	THVC	1.39
12	PRIV	0.83	PRMD	1.37
13	EMPH	0.79	NEMD	0.98
14	DWNT	0.74	PHC	0.81
15	XX0	0.73	WHSUB	0.68

上表より、日本人学習者と母語話者では同一トピックの作文であっても、上位 15 位に含まれる語彙文法指標は必ずしも一致しないという点が示された。以下、両者を比較し、上位 15 位以内において一致しない指標を取り出し、考察を試みる。

はじめに、日本人学習者に関しては、WHCL (*WH*-clauses)、FPP1 (first person pronouns)、PRIV (private verbs)、DWNT (downtoners)、XX0 (analytic negation) の指標が該当する。これらをスタンス表出と関連づけた場合、一般的な日本人学習者の英作文では、直接的な主張が書き手自身の観点 (FPP1) に偏りやすく、主観性 (PRIV) や限定的 (DWNT)・否定的 (XX0) 要素が母語話者以上に反映されやすい傾向にあることがいえるだろう。また、日本人学習者は、母語話者以上に、受け手に対して明示的意味合いの強い語や句を用いることで、より簡潔な主張を行っている可能性があるといえる。

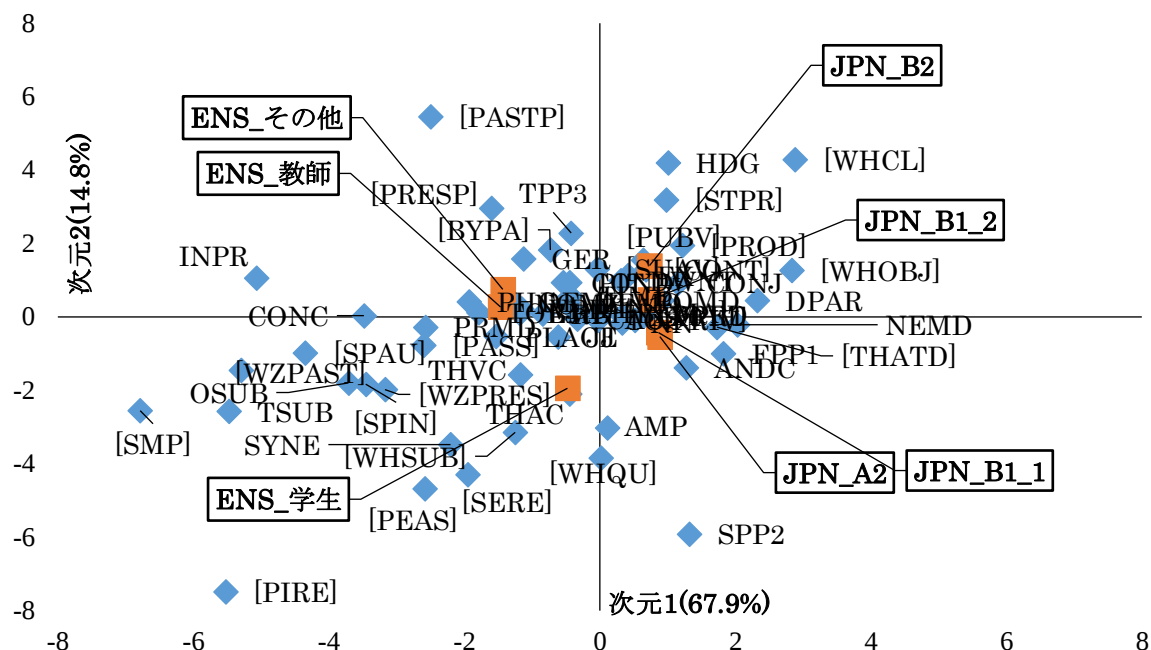
次に、母語話者に関しては、TSUB (*that* relative clauses on subject position)、OSUB (other adverbial subordinators)、PRMD (predictive modals)、PHC (phrasal coordination)、WHSUB (*WH* relative clauses on subject position) の指標が挙げられる。これらから、広く母語話者の作文では、関係代名詞の使用等、1 文あたりに示される情報が多様であり、従属節 (OSUB) や法助動詞 (PRMD) を適切に使用することで、状況を考慮しつつ、適度な推量が行われると考えられる。なお、日本人学習者のように、上位 15 位以内に FPP1 や PRIV は確認されず、母語話者は日本人学習者ほど、自身の主観を全面には出さずにスタンス表出を行う可能性が高い。

以上より、日本人学習者と母語話者の語彙文法使用の傾向、及びそれらから推測されるスタンス表出傾向の違いが示唆された。

4.2 RQ2 習熟度別日本人学習者と職業属性別母語話者が使用する語彙文法指標

RQ1 では、日本人学習者全体と母語話者全体を比較したが、日本人学習者を習熟度別に、母語話者を職業属性別に考察した場合、多用される語彙文法指標にどのような違いが生じるのかに関しても、調査と考察を行う必要があるだろう。そこで、すでに手法で言及した方法に従い、各群における指標別平均頻度表を作成し、多変量解析手法の一種である対応分析を実施した。その結果、以下の散布図を得ることができた。

図 1. 日本人の習熟度、母語話者の職業属性、個別語彙文法指標の関係性(対応分析)



上図において、右上の象限から反時計周りに象限 1・2・3・4 とする。散布図の次元 1(横軸)と次元 2(縦軸)の合計は、82.7%となっており、データ全体の 8 割以上が説明されていることを示す。以下、各次元に基づくデータの分布を評価する。

次元 1 を概観すると、右方には、WHCL(*WH*-clauses)、FPP1(first person pronouns)、HDG(hedges)、DPAR(discourse particles)、SPP2(second person pronouns)等、書き手が読み手に対して、直接的に内容を提示する際に使用すると推測される指標が分布している。また、左方には、SMP(*seem, appear*)、INPR(indefinite pronouns)、OSUB(other adverbial subordinators)、CONC(concessive adverbial subordinators)、TSUB(*that* relative clauses on subject position)、WHSUB(*WH* relative clauses on subject position)等、書き手が状況や物事を描写し、それらを踏まえて内容を提示するような指標が分布しており、右方に分布する指標ほど読み手への直接性が反映されるものではない。したがって、次元 1 は「読み手への直接性・間接性の程度」を示す軸であると評価することができるだろう。

次元 2(縦軸)を見ると、上方に PASTP(past participial clauses)、HDG(hedges)、TPP3(third person pronouns)、INPR(indefinite pronouns)、BYPA(*by*-passives)等が分布しており、これらは 3 人称的視点の反映や、事実への言及、控えめな主張等に関連し、自身の立場を示しつつも一定の中立的視点を増加させる指標であると推測される。また、下方には、PIRE(pied-piping relative clauses)、SPP2(second person pronouns)、WHQU(direct *WH*-questions)、AMP(amplifiers)、FPP1(first person pronouns)、SERE(sentence relatives)、SMP(*seem, appear*)等が分布しており、これらは書き手による読み手の直接的言及や、特定の物事への言及、主観的評価等に関連し、中立的視点を減少させる指標であると推測される。以上

により、次元 2 は「中立的視点の反映度」を指す軸であると評価することができるだろう。

各次元の評価を行った上で、習熟度別日本人学習者と職業属性別母語話者の分布傾向を見ると、日本人学習者と母語話者は次元 1 において区分されており、日本人学習者は、母語話者以上に直接的に内容を提示する傾向があるといえる。この点については、RQ1 と類似した傾向が確認され、日本人学習者と母語話者のスタンス表出方法の差を説明する 1 観点といえるだろう。その一方で、次元 2 より、日本人学習者は習熟度が高くなるにつれて、母語話者は学生から職業従事者（教師・その他）になるにつれて、中立的視点が増加する可能性が示唆された。このことは、日本人学習者も母語話者も、強意性の高い限定的な主張を行うスタンス表出傾向から、控えめで第三者的な立場から主張を行うスタンス表出傾向に変化していくことを示唆するのではないだろうか。次元 2 は次元 1 ほど寄与率が高くはないものの、L1 英語と日本人による L2 英語の間には、類似した変化があるといえるだろう。

以上を踏まえ、散布図の各象限における分布傾向を整理した結果、下表ようになった。

表 4. 散布図の各象限における分布特徴

象限番号	1	2	3	4
日本人の習熟度	B1_2 B2+			A2 B1_1
母語話者属性		教師 その他	学生	
語彙文法指標 (例)	WHCL HDG PUBV DPAR	PASTP TPP3 BYP PRES	PIRE SERE TSUB WZPRES	FPP1 SPP2 AMP
直接性・間接性	直接的	間接的	間接的	直接的
中立的視点の 反映度	高	高	低	低
特徴	主観的に陳述するが、絶対的なものではなく、強意性が低い。	3 人称的視点を踏まえ、状況的に、主張や内容の提示を行う。	主観性は低いものの、言及される物事は限定的である。	書き手が読み手の存在を意識し、強い主張を行う。

以上により、習熟度別学習者と職業背景別母語話者が使用する言語特性の傾向が示された。加えて、それらから推測されるスタンス表出方法における相違点や類似点についても示された。これにより、RQ1 では必ずしも示されなかった、日本人学習者内、及び母語話者内における言語使用実態の内部的变化がより明確化したといえるだろう。

4.3 RQ3 地域別学習者(B1レベル)と職業属性別母語話者が使用する語彙文法指標

図 2. 地域別学習者(B1)、職業属性別母語話者、個別語彙文法指標の関係性(対応分析)

上図において、RQ1と同様に、右上の象限から反時計周りに象限1・2・3・4とする。散布図の次元1(横軸)と次元2(縦軸)の合計は、63.8%となっており、データ全体の6割以上が説明されている。以下、各次元別に評価を行う。

次元 2 を概観すると、上方には、CAUS (causative adverbial subordinators)、SPP2 (second person pronouns)、TPP3 (third person pronouns)、TSUB (*that* relative clauses

on subject position)、COND (conditional adverbial subordinators)、HDG (hedges)などの指標が分布している。これらは、根拠の提示や物事の描写、また、絶対性の高い議論を避ける等、議論の一般化に関連する指標であるといえるだろう。その一方で、下方には、WHQU (direct WH-questions)、FPP1 (first person pronouns)、PRIV (private verbs)、SMP (*seem, appear*)、PUBV (public verbs)などの指標が分布している。これらは、1 人称的な視点に偏った内容提示に関連する指標と推測され、議論における一般化の程度は上方向高くはないと考えられる。以上より、次元 2 は、「内容における一般性の程度」と評価することができるだろう。

各次元の評価を行った上で、地域別学習者と職業背景別母語話者の分布傾向を見ると、学習者と母語話者は次元 1 において 2 区分されており、アジアの EFL 圏地域の学習者は、共通して、母語話者以上に内容提示が直接的である可能性が示された。RQ2 において、日本人学習者は、習熟度に関係なく、読み手に対して直接的であることが示されたが、当該傾向は、アジアの EFL 圏地域の学習者において共通の傾向であることがいえるだろう。また、次元 2 より、学習者の地域によって、内容的な一般性の程度が異なることが示されたが、日本人学習者の場合、韓国・中国・台湾人学習者と類似して、書き手の視点に偏った議論を行う可能性が高く、必ずしも議論が一般化されるわけではないと考えられる。この点は、当該地域の学習者によるスタンス表出が、アジア圏地域の学習者の中で、著者個人の視点に偏った限定的なものであることを示唆する。

なお、上記の散布図で示された地域別学習者と母語話者の分布傾向は、Ishikawa (2024) における分布傾向と必ずしも一致するものではない。本研究は、一部の習熟度別・地域別学習者データ、及び母語話者データにおける頻度の有無に関係なく、幅広い語彙文法指標を対象とした。それにより、RQ2 と RQ3 において、学習者と母語話者の差が大きく開いたと考えられる。

以上を踏まえ、散布図の各象限におけるデータの分布傾向を下表に整理した。

表 5. 散布図の各象限における分布特徴

象限番号	1	2	3	4
地域別学習者		IDN・THA	KOR・TWN・ CHN・JPN	
母語話者属性	その他・教師			学生
語彙文法指標 (例)	TPP3 PRMD COND NOMZ	CAUS POMD SPP2 HDG	WHQU WHOBJ FPP1 CONJ	SMP SYNE DWNT PUBV
直接性・間接性	間接性	直接性	直接性	間接性
一般性	高い	高い	低い	低い
特徴	第三者的視点から、条件を提示しつつ、一定の可能性を伴う内容提示を行う。	読み手に言及し、根拠等を含めつつ論じながらも、控えめな主張を行い、限定的な議論を避ける。	1 人称視点に偏った主張を行い、読み手に対して限定的な議論を行う。	状況的に論じているが、否定的・限定的主張を行い、内容を一般論化させるような議論は行わない。

以上により、地域別学習者(B1)と職業属性別母語話者が使用する言語特性の傾向が示された。これにより、学習者の地域差と個々の語彙文法指標の関係性が示され、アジアの EFL 圏地域における日本人学習者の使用言語特性の実態がより明確化したといえるだろう。

4.4 総合考察

以上により、日本人学習者の語彙文法指標、及びスタンス表出の傾向が示された。本節では、それらの知見がどの程度個々の英作文に当てはまるのかを質的に評価する。以下は、日本人学習者(B1_1)と母語話者(学生)の順に各々の作文を 1 例ずつ WE から引用したものである。はじめに、B1_1 レベルにおける日本人学習者の作文を引用する。なお、作文中では、本研究の知見と関連する一部の表現についてはボールドで示している。

I agree with this opinion. This is because that **I think** people can learn a lot of things that they cannot learn just at the university if they have a part-time job. There are many **kinds of** things **that** they can learn at the university, but **I think** there are also many another kinds of things **that** they cannot learn there. **Now I don't** have a part-time job, but **I want to** do some. **The reason why** is not only to earn some money by myself, but also to be helpful for someone. **I believe** that the experience of having a part-time job will enrich the people's life later. To have some part-time job, they meet many people in various situations, so they can have a chance to communicate with people in another generation and touch various way of thinking and living of the people. **I think** they also can get some information of how to live well in the world through working some part-time job. In other words, to have a part-time job is a kind of preparation for their future, **I think**. Therefore, **I agree** with the opinion that it is important for college students to have a part-time job.

(ICNALE_WE_JPN_PTJ0_227_B1_1)

上記の作文では、主格の 1 人称代名詞を主語とした文が多用されており、特に **I agree/think/believe/want to** といった表現が目立つ。また、**The reason why** のような、読み手に直接的で限定的な議論を行う印象を与える表現も使用されている。RQ1 において FPP1 や PRIV が日本人の特徴的な語彙文法指標であること、RQ2 において、B1_1 レベルの日本人学習者が直接的で強い言及を行うこと、RQ3 において、書き手視点で限定的な主張をおこなうことといった知見を示したが、上記の作文はこれらの傾向に一定の範囲内で当てはまるものといえるだろう。なお、日本人学習者による 1 人称代名詞の多用については、これまでも多くの先行研究で議論されており、小林(2010)、成田(2017)、飯島(2025)等は当該事項に焦点を当て、学習者コーパスを用いた頻度調査を実施した上で考察を行っている。

次に、大学生母語話者による作文を以下に示す。

I think that college students **who** work while they attend school have a very good idea. I recently read a study **in which** it was shown that students **who** work approximately 10 hours per week actually have increased academic performance over their counterparts who don't work. At this level of work, the **only** thing **that** the student is replacing his own schedule is usually time that he or she has wastes by doing things such as watching television or goofing around. Thus, as long as students can find a balance like this that works for them, there are **no** negative effects of having a part-time job. However, some students have to work fulltime all they attend college, and all of these people report that having a fulltime job while trying to attend classes is simply too much to deal with. Some of these people have no choice as they will have **no** other way to pay for their college tuition, and many of them drop out because they find that work is more important than school. But the money they might turn at the fulltime job over a few years pales in comparison if the money they can earn **if** they have a real career. Therefore, students should stick to part-time work during college.

(ICNALE_WE_ENS_PTJ0_056_XX_1)

上記の作文を見ると、内容における研究事例の提示や、**if** 節の使用が見られ、事例や条件等を踏まえながら読み手に対して間接的な主張を行っている印象を得る。このように間接的な内容提示を行う点は、前者における日本人学習者の作文事例とは異なっているといえる。また、母語話者

の作文では、関係代名詞(TSUB、PIRE)の多用や only (DWNT)の使用等、個別の限定的な議論がなされている印象を得る。加えて、内容中において否定語 no (SYNE)を用いた主張がなされている。RQ1 において、日本人学習者と母語話者を比較した際の特徴的指標の一つが TSUB であること、RQ2・3 において、母語話者が日本人学習者以上に読み手に対して間接的な内容提示をすること、及び大学生母語話者の主張は中立的視点の反映度が低い可能性があることが示唆されたが、上記で引用した大学生母語話者による作文は、当該傾向に一定の範囲内で当てはまるものといえるだろう。

以上より、本研究の各 RQ における計量的手法に基づく知見が、一定の範囲内で個々の作文にも当てはまるといえるのではないだろうか。

5. まとめ

本研究は、ICNALE を使用し、日本人英語学習者、及び母語話者による意見英作文で使われる多様な語彙文法指標を調査し、それらからスタンス表出傾向を考察した。以下、RQ ごとに本研究から得られた知見を整理する。

RQ1 では、日本人学習者全体と母語話者全体を比較し、同一トピックの作文中で多用される語彙文法指標の違い、及びそれらを踏まえたスタンス表出傾向の違いを調査した。その結果、日本人学習者のスタンス表出は、書き手自身の観点が強く直接的に反映され、主観性や否定的要素が母語話者以上に前面に出やすいことが示された。一方母語話者は、関係代名詞節や従属節などの多様な文構造を多用することで詳細な内容を伴う主張を行い、さらにそこに一定の推量を行っている可能性が示された。加えて、母語話者は日本人学習者ほど、自身の主観を全面には出さずにスタンス表出を行う傾向を示すことが示された。

RQ2 では、習熟度別日本人学習者(A2、B1_1、B1_2、B2+)と職業属性別母語話者(学生、教師、その他)が多用する語彙文法指標の違いを考察した。対応分析の結果、習熟度別学習者と職業属性別母語話者間の関係性が、「読み手への直接性・間接性の程度」「中立的視点の反映度」の2観点で説明できる可能性が示された。これらを踏まえ、日本人学習者のスタンス表出の特徴を考察すると、(1)母語話者以上に直接的なスタンス表出を行うこと、(2)習熟度の上昇と共に中立的な視点を取り入れる可能性があることが示された。なお、(2)については、母語話者においても、学生から職業従事者(教師・その他)への変化の過程で同様の変化を示すことが示された。これにより、日本人学習者と母語話者の内部的変化の共通性に関しても併せて示された。

RQ3 では、学習者の地域差に着目し、アジアの EFL 圏地域の学習者(日本、インドネシア、中国、台湾、韓国、タイ)と職業属性別母語話者(学生、教師、その他)が多用する語彙文法指標の違いを考察し、日本人学習者のスタンス表出傾向を考察した。RQ2 と同様に対応分析を実施した結果、地域別学習者・母語話者と個別語彙文法指標の関係性が、「読み手への直接性・間接性の程度」「主張の一般性の程度」の2観点で説明可能であることが示された。日本人学習者のスタンス表出傾向に関しては、(1)習熟度差に関係なく読み手に直接的であるという特徴(RQ2)が一定の地域的共通性を有すること、(2)韓国・中国・台湾人学習者と同様に、議論の一般性が低く、書

き手の視点に偏った主張を展開する可能性が高いことが示された。

以上の知見が明らかとなったが、本研究には一定の課題も存在する。以下、課題別に言及する。

1 点目は、本研究が考察対象とした語彙文法指標は、あくまでも MAT が対象とする指標に限定されているということである。そのため、語彙文法指標の区分が結果に影響を与えている可能性も考慮に入れる必要があるだろう。

2 点目は、RQ1 において使用した指標である標準化得点は、あくまでも Biber (1988) における現代英語コーパスに基づくということである。そのため、直近の現代英語の特徴が十分に反映されていないという可能性も考慮に入れる必要があるだろう。

以上の 2 点を本研究全体の課題としたい。

引用文献

- Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.
- Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research*, 1(1), 7–24.
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173–192.
- Hyland, K. (2019). *Metadiscourse: Exploring interaction in writing*. Bloomsbury Academic.
- 飯島真之 (2025). 「日本人英語学習者による 1・2 人称代名詞使用: ICNALE の英作文モジュールを用いたアジアの地域別比較」『中部地区英語教育学会紀要』, 54, 25–32.
- Ishikawa, S. (2023). *The ICNALE guide: An introduction to learner corpus study on Asian learners' L2 English*. Routledge.
- Ishikawa, S. (2024). Lexicogrammar of the L2 English essays written by Asian college students: A corpus-based study. *The Journal of Asia TEFL*, 21(1), 119–138.
- 小林雄一郎 (2010). 「日本人学習者の英作文における人称代名詞について」『言語処理学会第 16 回年次大会発表論文集』, 1074–1077.
- Li, X. (2023). A multi-dimensional analysis of English writing by Chinese EFL learners. *Journal of Language Teaching and Research*, 14(1), 37–47.
- 成田真澄 (2017). 「日本人大学生が産出した英語論述文における主格人称代名詞使用傾向の分析」『東京国際大学論叢 人文・社会学研究』, 2, 1–20.
- Nini, A. (2019a). The multi-dimensional analysis tagger. In T. B. Sardinha & M. V. Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 67–94). Bloomsbury Academic.
- Nini, A. (2019b). *Multidimensional analysis tagger (v.1.3): Manual*.
<https://corpus.bfsu.edu.cn/MATv1.3Manual.pdf.pdf>
- Santorini, B. (1990). *Part-of-speech tagging guidelines for the Penn Treebank project*

(3rd revision). <https://catalog.ldc.upenn.edu/docs/LDC99T42/tagguid1.pdf>

コーパス・その他分析ツール

Ishikawa, S. (2025). Written Essays (v. 2.6). ICNALE: The International Corpus Network of Asian Learners of English. <https://language.sakura.ne.jp/icnale/>

Nini, A. (2019). MAT: Multidimensional Analysis Tagger (v.1.3).
<https://sites.google.com/site/multidimensionaltagger>

清水裕士 (n.d.). HAD (v.18). <https://norimune.net/had>