

科学論文における自己参照マーカー「we」の分析

萩原 明子¹ 小林 薫² 内藤 麻緒³

1. 序論

論文は複雑な内容の研究をひとつの連なった文章で論理的に書き表したものである。ICT が広く普及し、マルチメディアを用いた異なる様相のデータの提示法が可能になった現在においても、学術論文はもっとも信頼される研究成果の公表の方法として認知されている。それゆえ学術分野を問わず国際的に認められるためには、英語を用いて読者、多くの場合はその分野に関わる他の研究者に適切な形で情報を提供することが求められている。著者として論文を作成する際には、読者に誤解させることなく正確に研究の内容を伝えるために文章を整える必要があり、読者としては著者の伝える内容を著者の意図通りに理解することが重要である。すなわち、論文を書いたり読んだりすることは、著者と読者の相互活動なのであり、論文にはその相互活動を円滑にするための言語表現（機能表現）が随所に使用されている。本論では科学論文における機能表現の中でもメタ談話標識として分類されている self-mention（自己提示マーカー）の分析を行い、第二言語使用者への言語教育の可能性を検討した。

1-1. 科学論文における一人称代名詞

研究論文において、一人称代名詞の使用が避けられている時代があったが、ここ数十年、自然科学のみならず社会科学の論文においても we に代表される一人称代名詞の使用が増加してきている（福池, 2019; Hyland & Jiang, 2018）。科学論文を執筆する場合、文法的には一人称代名詞は必ずしも必要なものではなく、現在でも一人称代名詞を使わないように指導する文献（Kirkman, 2005）もあり、実際、一度も使用しないで書かれた論文も少なからずある（後述）。一方、積極的に用いる方が適切であるという（Schultz, 2009）見方もある。

一人称代名詞が使用される理由は一般的には3つ挙げられる。一つには一人称代名詞を使用する能動態の文はそれを必要としない受動態の文より文構造が単純で読みやすく、冗長になるのを防ぐことができる（Strunk and White, 1999）。二つ目として、文の動作主が著者自身であることを示すことで、研究に関わる実験操作や手順において著者の貢献を明確にすることができる。三つ目は、背景、考察などにおいて、当該研究のみならず過去に著者が行った研究結果に関する著者自身の見解を述べる際に、著者の独自性、責任、功績が主張できることである。一人称代名詞が導く文に他の機能を持つメタ言語標識を使用することにより、語調を強めたり弱めたりすることができるため、事象を客観的に捉えより正確な形で示すことが可能になる。このように一人称代名詞を使用することは、文章が読み易くなり著者の貢献が明確になるばかりか、著者自身の主張の度合いを調節し易くなるという利点がある。

科学論文執筆を目標とする英語学習者にとっては、一人称代名詞を論文の中で目にする機会は多いが、実際に使用するためのハードルは高い。日本語では主語を明示しなくても文法的に正しい文（「○○法を用いて細胞数を測定した」等）を書くことができるが、日本語の文章を英語にする場合、

¹生命科学部言語科学研究室 ²東京農業大学生命科学部 ³聖マリアンナ医科大学医学教育文化部門

主語を補う必要があり、適切な主語の用い方を改めて学ぶ必要が生じる。英語で論文を書くときに、具体的に一人称を論文のどこで使えば良いのか、どの語と組み合わせるとより正確な情報を読者に伝えることができるのかを知ることは、客観的な文章を書くための重要なスキルとなる。

1-2. メタ談話標識としての一人称代名詞の機能

一人称代名詞は、メタ談話標識の一つである。メタ談話標識とは文章の命題以外の部分で文章の展開を示しながら書き手と読み手の相互関係を築く機能表現で、*interactive metadiscourse* と *interactional metadiscourse* の二つに分類される (Hyland, 2005)。Interactive metadiscourse が論旨の展開を示す標識であるのに対して、interactional metadiscourse は書き手が自分の主張に対する読者の受け取り方に配慮しながら自分の議論に読者を関わらせるための標識であり、一人称代名詞は、interactional metadiscourse の self mention の一つである (表 1)。一人称代名詞は、研究論文において著者の声を構築する機能を持ち、著者は自分の研究成果を顕示するためにこれを使用する (Camps & Ivanič, 2001; Martinez, 2005)。Interactional metadiscourse には一人称代名詞の他に booster と hedge があり、いずれも文の命題についての書き手の確信の度合いを示すものである。Booster は書き手が命題の正しさについて強い確信を持ち、相対する意見をしりぞける意志を示すのに対して、hedge は書き手が文の命題について他の意見があることを認識していることを示し、読み手を議論に誘う機能がある (Hyland, 2005. pp. 52-53)。更に、一人称代名詞が booster または hedge と共起することもある。Booster と共起する場合は命題に対する著者の高い確信を示すだけでなく、読者を強く説得するはたらきがあり (Hu & Cao, 2015)、逆に hedge と共起する場合は読者に譲歩する姿勢を表明する機能がある。

表 1. メタ談話標識の分類

メタ談話標識	機能	例
Interactive	読み手のテキスト理解を助け、解釈を導く	
Transitions	主節同士の関係を示す	in addition; but; thus; and
Frame markers	テキストの構造を示す	finally; to conclude; my purpose is
Endophoric markers	テキスト内の情報を指す	note above; see Fig; in section 2
Evidential markers	引用文献等を示す	according to X; Z states
Code glosses	命題の例示や言い換えを導く	namely; e.g.; such as; in other words
Interactional	読み手を関わらせる	
Hedges	語調を弱め意見交換を始める	might; perhaps; possible; about
Boosters	主張を強め意見交換を終了する	in fact; definitely; it is clear that
Attitude markers	命題に対する書き手の感情を表現する	unfortunately; I agree; surprisingly
Self mentions	書き手自身に言及する	I; we; my; me; mine; the author
Engagement markers	読み手に呼びかける	consider; note; you can see that

注：Hyland (2005, p. 49) による。

一人称代名詞と共起する booster と hedge については、Walková (2019) が言語学分野の三つのコ

ーパス（スロバキア語母語話者が母語で書いた論文、同母語話者が英語で書いた論文、英語母語話者が英語で書いた論文）において、一人称代名詞の使用を分析した。Walková は、英語で書かれた論文の場合は *we* のもつ自己言及の度合いが、*booster* と共起する場合は強くなり *hedge* と共起する場合は低くなることを示し、スロバキア母語話者の英語論文における一人称代名詞の使用は母語の影響を多少なりとも受けると結論付けた。また、Wang and Zeng (2021) は、4 つの科学分野において、中国語母語話者の学生が書いた論文と同母語話者の研究者が書いた論文で一人称代名詞が *booster* と *hedge* と共起するパターンを分析した。結果として、4 つの分野間および使用目的（目的設定、手順の説明、結果の提示、主張を述べる、先行研究に言及する）で差があることを示し、全体的に両者の使用パターンが類似することに基づき、学生は一人称代名詞と *booster/hedge* の使用パターンを学習していると考察した。これらの二つの論文は一人称代名詞と共起する *booster* と *hedge* の論文における使用パターンは分野によって違うこと、著者の母語が英語ではない場合は母語の影響を受けること、使用パターンが学習可能であることを示唆している。

これまでの研究で、一人称代名詞のメタ談話標識としての機能、その機能を *booster* と *hedge* が調節すること、論文においては一人称代名詞、*booster* と *hedge* の使用パターンは著者の母語の影響を受けるが、学習も可能であることが示された。しかし、一人称代名詞と *booster, hedge* を組み合わせたパターンをコーパス分析した研究はまだ少なく、特に生命科学分野の論文に特化した、著者が日本語母語話者であるコーパスを論文のセクション別に分析した研究はまだ無い。

1-3. 本論の目的

日本において科学論文作成を指導するための基礎研究の必要性を鑑み、本論の目的は、生命科学分野の英語論文における一人称の使用について、論文著者の言語的な背景をもとにコーパス研究の手法を用い、定量的に分析することである。国際ジャーナルに掲載された論文を著者の所属機関の所在地別に欧米（インドヨーロッパ語圏）、東アジア（日本語以外の漢字文化圏）、日本の3地域に分類し、一人称代名詞の中でも最も頻度が多い複数形の主格 *we* の使用頻度¹を論文のセクション（abstract, introduction, methods, results, discussion）ごとに比較し、さらに共起する動詞のメタ言語標識（特に *booster* と *hedge*）を分析した。

2. コーパス構築と *we* の抽出および分析方法

論文中の一人称の使用は限定的であり、すべての著者が使用しているわけではないことが考えられるため、本研究においてはコーパスを構築し、コーパス中に出現した *we* の頻度を比較した。コーパスは、インパクトファクターを考慮し、生命科学関連分野を12の分野に分類し、それぞれ40編ずつ、合計480編収集することを目標とした。言語間にインパクトファクターや内容による偏りが出ないように、著者の言語背景を3つに分け、ひとつのジャーナルからそれぞれ1編ずつ4編の論文を選択した。分析の対象としたコーパスは約200万語で構成され、含まれる語彙はレマをベー

¹ 本研究のコーパスに出現した一人称代名詞の頻度は *we* が 10,488、*our* が 3173、*us* が 383 だった。単著の論文は一つだけだったため、一人称単数の代名詞も対象から外した。

スにすると約 6 万 7 千語 (表 2) だった。本論で分析の対象とした論文は 478 編²とし、それぞれの論文を IMRaD (Introduction, Methods, Results, Discussion) のセクションに分け、セクションごとにテキストファイルを作成した。

表 2. コーパスの概要

コーパス	論文数	異なり語数	総語数
E (欧米)	240	33,910	1,083,083
A (東アジア)	119	20,428	453,392
J (日本)	119	19,781	437,685
コーパス全体	478	67,274	1,974,160

コーパス分析に使用した文献は基本的にインターネット上で公開されているもの(open access)とし、一般的に使用する表計算やワードプロセッサ以外に、(1) Preview, Text edit (Apple 社) : pdf から UTF-8 のコードに変換、(2) WordSmith Tools (v.10) (Lexical Analysis Software 社 / オックスフォード大学), AntConc (v.3.4.4m) (Laurence Anthony) : コーパス分析、(3) FileMaker Pro. (v.19) (クラリス社) : データ整理、(4) js-Star XR : 統計分析、の 4 種類のソフトウェアを使用した。文字コードは UTF-8 を用いた。

このコーパスから、WordSmith Tools を用いてワードリストを作成し、一人称の出現数を確認した。その中で最も頻度の多い一人称複数主格の *we* を WordSmith Tools の concordance の機能を用いて、データとして抽出した。Concordance のデータを FileMaker Pro を用いて、著者の言語背景と論文のセクションとで分類できる形のデータとし、*we* の頻度と *we* を主語とする動詞を抽出した。本論では、「コーパス」は、データ化した全ての論文データの総称を示し、「サブコーパス」はコーパスを言語背景 (E は欧米の研究者、A は日本以外の東アジアの研究者、J は日本の研究者によるもの) で分けたものを示す。「論文」は論文データ全体を指し、「セクション」は IMRaD (Introduction, Methods, Results, Discussion) の各セクションを指し、「テキスト」はセクション毎に分けられた電子ファイルのことを指すこととする。

3. 科学論文の各セクションにおける *we* の頻度

コーパスに含まれていた *we* の総数は 10,488 (総語数のうち 0.5%) であり、*we* を一度も使用していなかった論文は 26 編あった。*we* は論文の全てのセクションに使用例があったが、セクションおよび著者によって頻度のばらつきがみられた。論文各セクションで *we* を使用したテキストが含まれている比率 (表 3) を見たところ、Introduction および Discussion のセクションで *we* が使用されたテキストが多く、Methods では *we* が含まれていないものの方が多い結果になった。言語背景を見ても、J がそれぞれのセクションの中で *we* を使う比率が最も高く Methods でも 55.3% の使用比率があったが、E は *we* の使用率が相対的に低い傾向が見られた。すなわち、論文を書くときに、全てのセクションで一度でも *we* を使う比率が J においては他地域より多いことが示された。実際、論文全体で *we* を使用する比率も J が最も高く、次は A で、E では *we* を全く使用しない論文を書く

² 今回、データを精査する過程でテキストファイル作成上の問題が含まれていたため、東アジアの研究者によるもの (A) と日本人の研究者によるもの (J) から、それぞれ 1 編の論文を分析の対象から外した。

著者が 7.5%を占めていた (表 3)。

表 3. 各セクションで we を含むテキストの数と比率 (we を含むテキスト数/テキスト総数%)

Section / text	All	E	A	J
Entire article	452/478 (94.6%)	222 /240 (92.5%)	114 /119 (95.8%)	116/119 (97.5%)
Abstract	326/478 (68.2%)	155/240 (64.6%)	77/119 (64.7%)	94/119 (79.0%)
Introduction	400/478 (83.7%)	192/240 (80.0%)	102/119 (85.7%)	106/119 (89.1%)
Methods	168/350 (48.0%)	89/184 (48.4%)	32/81 (39.5%)	47/85 (55.3%)
Results	300/420 (71.4%)	144/214 (67.3%)	78/104 (75.0%)	78/102 (76.5%)
Discussion	446/510 (87.5%)	220/255 (86.2%)	112/130 (86.1%)	114/125 (91.2%)
計	1640/2236 (73.3%)	785/1133 (69.3%)	396/553 (71.6%)	430/550 (78.1%)

注：Discussion には Conclusion のファイルが含まれているため母数が論文数を超えている。

それぞれのセクションにおける we の使用頻度を 10 万語あたりで比較 (表 4) してみたところ、Methods 以外の全てのセクションにおいて J が we を最も頻繁に使用していた。全てのセクションで we の頻度が最も少なかったのは A によるものだった。

表 4. それぞれのセクションにおける we の頻度 (/10 万語)

Section / text	All	E	A	J
Abstract	669 (618.4)	357 (632.6)	124 (472.7)	188 (736.8)
Introduction	1,060 (380.7)	556 (377.7)	236 (352.4)	268 (417.3)
Methods	1,407 (352.4)	929 (396.4)	185 (235.3)	293 (340.4)
Results	4,569 (766.4)	2,430 (770.7)	1,092 (733.4)	1,047 (793.5)
Discussion	2,783 (464.6)	1,521 (462.0)	577 (432.6)	685 (502.2)
計	10,488 (529.4)	5,793 (535.2)	2,214 (487.4)	2,481 (558.6)

これらの結果では、we を使用する著者は J が多く頻度も高かったが、E では使用する著者は比較的少ないが、使用されているテキストにおいては we の頻度は必ずしも低い訳ではないことが示された。著者による使用頻度の差が大きいことが考えられるため、we が使用されていたテキストごとの使用頻度を言語背景で比較した (表 5)。ここでは、E の方が 1 テキストあたり、we を多く使用していることがわかった。三つのグループにおいて頻度の平均にくらべて標準偏差が大きく、テキスト中の we の頻度に大きなばらつきがあることが示唆される。さらに、中間値が A と J では大きく異なり、A において we の使用頻度が少ない論文の数が相対的に多いことが示された。

表 5. 1 テキストあたりの we の平均頻度と標準偏差

We の頻度	E	A	J
中間値(median)	18	13	17
平均値(mean)	7.14	5.45	5.51
標準偏差(SD)	10.55	8.06	7.85

論文全体の we の度数分布を調べたところ (表 6) 欧米の著者とアジアの著者には大きな違いが見られた。アジア (A, J) では、論文全体で 61 回以上 we の使用が見られたのは全体の 3.4%~4.4%に過ぎないが、欧米では全体の 11.3%だった。一方、we を 20 回以下の頻度で使用している論文は A では 63.2%、J では 61.2%であったが、E では 54.0%と相対的に少なかった。

以上、we の分布は論文のセクションごとに異なり、論文によって we の頻度には大きなばらつきが見られること、および、著者の言語背景が we の使用に影響を与えていることが示唆された。

表 6. 言語背景による論文の we の度数分布と相対度数 (%)

度数 we	E	相対度数	A	相対度数	J	相対度数
1-10	72	32.4	41	36.0	38	32.8
11-20	48	21.6	31	27.2	33	28.4
21-30	31	14.0	19	16.7	16	13.8
31-40	26	11.7	11	9.6	16	13.8
41-50	8	3.6	5	4.4	3	2.6
51-60	12	5.4	2	1.8	6	5.2
61-	25	11.3	5	4.4	4	3.4
計	222	100.0	114	100.0	116	100.0

注：論文中に we の使用のないものは分析の対象外とした。

4. We と共起する動詞 (booster と hedge)

主格の we は動詞を伴うため、メタ言語標識を使用して we と共起する動詞を分析することを試みた。コーパス中の we の使用頻度合計は 10,488 で、共起する動詞は 609 種類あった。メタ言語標識に分類される動詞は 51 種類(4,190 語, 42%)あり、その他は 558 種類 (6,050 語, 58%) あった。頻度数の割合から we はメタ言語標識と共起しやすい傾向があると考えられる。

サブコーパス内での頻度の順 (表 7) を見ると、we と共起しやすい動詞の種類は言語背景によって大きな違いはないようである。メタ言語標識の booster と hedge の動詞では booster は 8 種類 1567 語、hedge は 8 種類 180 語あり booster のほうが多く使用されていた。他に attitude marker, code gloss, engagement marker, frame maker があり、そのうち engagement marker(29 種類, 2,634 語)に属する動詞の占める割合が大きかったが、engagement marker の特性上それらがメタ言語標識として使用されているかどうかは文脈に依存するため、今回は booster と hedge のみ分析対象とした。

表 8 はコーパスに含まれる全ての booster/hedge 動詞のセクション別の出現頻度をまとめたものである。セクションによって使用される動詞に偏りがあり、show や find など特定の動詞の頻度が高い。

表 7. コーパス内の“we”と共起する動詞 (コーパスおよびサブコーパス内の順位)

動詞 (レマ)	All	E	A	J
find (booster)	1	1	1	1
use (engagement m)	2	2	2	3
observe (engagement m)	3	3	5	9
examine	4	10	2	2
perform	5	6	4	4
show (booster)	6	4	9	7
identify	7	5	8	5
analyze (engagement m)	8	7	10	6
demonstrate (booster)	9	9	6	10
test	10	8	(11)	(27)
investigate	11	(12)	7	8

表 8. コーパス内の“we”と共起する booster 動詞と hedge 動詞

		Abstract	Introduction	Methods	Results	Discussion	計
Booster	believe			1	2	39	42
	demonstrate	30	45		17	132	224
	establish	5	6	2	15	15	43
	find	63	53	15	415	325	871
	know		5	3	1	14	23
	realize			1		4	5
	show	68	69	3	49	165	354
	think				2	3	5
Booster 計							1567
Hedge	argue	1		1	2	7	11
	assume	3	1	13	16	20	53
	estimate	2	1	12	23	12	50
	feel		1			3	4
	postulate		2		3	6	11
	suggest	6	3		3	24	36
	suppose					3	3
	suspect		2		5	5	12
Hedge 計							178

5. 考察

英語で書かれた同等分野の論文を言語背景で分析したところ、we を使用する文体は広く受け入れられている傾向があるものの、欧米においては使用しない文体を支持する著者の比率が高かった。一方、第二言語話者の研究者に関しては、独自の文体を使用するより、その分野で一般的によく使用されている we を使用する文体を取り入れて論文を執筆する傾向が見られた。日本と他の東アジア圏（主に中国）と比べると、日本語においてはフォーマルな表現として論文で「我々」が頻繁に使われるため、主語抜き日本語から英語に置き換える時に言語転移により we の頻度が高い可能性がある。主語を省略しない中国語では、複数形の一人称にはフォーマルな表現と日常の表現に違いがないため、日本語ほど we の使用頻度が多くはない可能性がある。このように、母語の影響が推測されることは Walkova (2019) の考察と一致するが、今後それぞれの言語で書かれた論文における一人称の使用を分析する必要がある。

セクションごとの we の分布は、文体として使用が確立されているのは Introduction と Discussion 部分で、Results での頻度が高いのは、we を多用する著者が特に Results 部分で多く使用しているからであり、Results で we を使用しない著者も多い。とくに言語背景が欧米の場合、we を論文中で頻用する著者の比率が高く、それらの著者は特に語用論的な意図で we を使用するのではなく、純粋に主語を明確にし、単純化する目的で使用している可能性は否めない。ただ we と共起する動詞を見ると、find が Results でとりわけ多く使われており、これは、特に客観的に記述される Results セクションにおいてもおそらく著者の貢献を強調するために使用されている可能性を示している。今後、定性的な分析を用いて we を高頻度使用する論文のパターンを見極める必要がある。

アブストラクトの結論部分で用いられる定型表現 our results [suggest / demonstrate / show] that (引用) のパターンに含まれる booster, hedge 動詞は、今回のコーパスにおいても Abstract, Introduction,

Discussion に多く含まれ、研究の結果の重要性を記述するために頻繁に用いられることが実証された。今後は論文で使用される we と共起するこれらの動詞による定型表現を同定し、教育に応用することが可能になる。

本論では、欧米、アジア、日本の地域別に生命科学分野の論文での we の使用頻度を論文のセクション (Abstract, Introduction, Methods, Results, Discussion) ごとに比較し、さらに共起する動詞のメタ言語標識 (特に booster と hedge) を分析した。その結果、欧米地域には we を全く使わない著者も頻用する著者も多く、アジアと日本の著者では欧米とは異なる we の使用傾向が見られ、母語の影響が示唆される。We と共起する booster および hedge については、サンプル数が少ないこともあり地域差による違いは検出できなかった。以上から生命科学分野の英語教育では、we の使用セクションとして確立されている Introduction と Discussion での使用パターンを、共起する hedge, booster と併せて学習させることが論文執筆の準備として有効であろう。

参考文献

- 福池深月 (2019). 「学術研究論文の abstract 中に見られる self-mentions マーカー, “we”の使用」東京薬科大学大学院生命科学研究所修士論文 (未公開)
- Camps, D. & Ivanič, R. (2001). I am how I sound. Voice as self representation in L2 writing. *Journal of Second Language Writing*, 10(1-2), 3-33.
- Hu, G., & Cao, F. (2015). Disciplinary and paradigmatic influences on interactional metadiscourse in research articles. *English for Specific Purposes*, 39, 12-25.
- Hyland, K., & Jiang, F. K. (2018). “In this paper we suggest”: Changing patterns of disciplinary metadiscourse. *English for Specific Purposes*, 51, 18-30. <https://doi.org/10.1016/j.esp.2018.02.001>
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. New York: Continuum.
- Kirkman, J. (2005). *Good Style: Writing for Science and Technology*, 2nd ed. London: Routledge.
- Martínez, I. A. (2005). Native and non-native writers’ use of first person pronouns in the different sections of biology research articles in English. *Journal of Second Language Writing*, 14(3), 174-190. <https://doi.org/10.1016/j.jslw.2005.06.001>
- Schultz, D. M. (2009). *Eloquent science: A practical guide to becoming a better writer, speaker and atmospheric scientist*. Boston: American Meteorological Society.
- Strunk, W. & White, E. B. (2000). *The Elements of Style*. 4th ed. New York: Longman.
- Walková, M. (2019). A three-dimensional model of personal self-mention in research papers. *English for Specific Purposes*. 53(1):60-73
- Wang, J., & Zeng, L. (2021). Disciplinary recognized self-presence: Self-mention used with hedges and boosters in PhD students’ research writing. *SAGE Open*, 11(2), <https://doi.org/10.1177/21582440211005454>