科学系高校生のための ESP 開発基礎研究: 語彙の観点から

山岡大基

本研究は、高校生が自然科学分野で論文、特に英文要旨を書く際に必要となる英語表現を明らかにすることで、自然科学のための English for Specific Purposes (ESP) 教育の教材およびカリキュラムを開発していく際の基礎資料を得ることを目的とする。広島大学附属高等学校の生徒による2015~2022 年度分の課題研究論文の英文要旨をコーパス化し、使われている語やコロケーションの特徴を分析した。また、汎用コーパスおよび教科書コーパスと比較することで、高校生の書く自然科学論文に見られる英語表現上の特徴を抽出した。その結果、自然科学論文の、主観的な意見を述べるよりも事実を説明する性質により、高校生が English for General Purposes (EGP) で書くのとは異なるディスコースが作られていることがわかった。また、英語力に制約のある高校生にとって使いやすい表現の存在が示唆された。

1. 目的

本研究の目的は、English for Specific Purposes (ESP: 特定目的のための英語教育)の観点から、高校生が自然科学分野での課題研究を論文化する際に、特に論文要旨 (abstract) を英語で執筆するのに、どのような表現を必要としているかを探り、今後の教材およびカリキュラム開発のための基礎資料を得ることである。

まず、本研究が求められる背景に、次のような事情がある。

広島大学附属高等学校(以下「本校」)は、平成15年度(2003年度)より4期にわたりスーパーサイエンスハイスクール(以下「SSH」)の指定を受け、先進的な理数系教育を通じた科学技術人材の育成に取り組んでいる。カリキュラムの中核を成すのは生徒による課題研究であり、自分たちで設定した研究課題について約2年をかけて科学的に探究する。その成果は生徒自身の手で論文にまとめ、本校として、毎年度『課題研究論文集』の形で公表している。

第3期にあたる平成24年度(2012年度)からは、 国際性の育成という観点から、論文要旨を英語でも 執筆することとしている。執筆作業は、概略として は次のようなプロセスで進む。

- 1) 生徒が和文で要旨を作成する
- 2) 研究の指導教員が内容について指導し、生徒が和文要旨を完成させる
- 3) 生徒が和文要旨をもとに英文要旨を作成する
- 4) 英語科教員の指導のもと、生徒が英文要旨を

完成させる

しかし、実際はこのように直線的に進むことはまれで、最終的な英文要旨が仕上がるまでには、1) ~ 4) を何度か往復するのが普通である。

そのようになる理由はいくつかある。まず、教員 側の要因として、論文の内容が英語科教員にとって は高度であり、生徒が言い表したい内容を適確に表 す英語表現を確定するのにコストがかかることが挙 げられる。英語科教員は, 英文要旨の添削をする前 に、まず論文の内容を正確に理解しなければならな い。また、生徒の日本語による表現も常に十分とは 言えず. 限定された専門的知識を持たない読者には 意味が伝わりにくい文言が不用意に用いられている ことも少なくない。そのようなことから、英語科教 員と生徒の間で「ここはどういう意味?」「どのよ うな実験だったの? | といったやり取りを繰り返す 中で、生徒が本当に言い表したい内容を明らかにし、 それに応じて英語科教員も適確な英語表現を提案す る、というプロセスが必要になる。場合によっては、 英語科教員が研究の指導教員から研究内容について レクチャーを受けることもある。

また、生徒側の要因としては、論文要旨や論文本体に含める内容の精選が難しいことが挙げられる。 多くの場合、論文の主旨を伝えるには不要な情報までも含めようとする傾向がある。これは致し方ないことで、生徒にとって課題研究は高校生活において相当な時間と労力を費やして取り組むものであるため、研究に着手した当初からの紆余曲折すべてがかれらにとっては大切な足跡であり、それを、集大成 としての1本の論文に書き残したいと思うのは無理 からぬことである。

しかし、当初の目論見通りに研究が進むことは少なく、研究の途中で実験の内容を変更したり、場合によっては研究目的そのものを設定し直したりすることすらある。そういった紆余曲折は科学論文としてはノイズであり、そういった、論文の主旨に沿わないノイズを削ぎ落とすよう生徒に促すことも、要旨執筆の指導においてはしばしば必要になる。

このように、科学論文にふさわしい英語表現を求めることには、いわゆる英作文の添削とは別種のノウハウが必要である。科学論文を英語で書くことについては、既に大学以上の学生・研究者を対象にした概説書やマニュアルの類が多数出版されており、本校学校設定科目「サイエンスコミュニケーション」等においてもそれらを参考にして指導を行っている。しかし、高校生であることから、基盤となる英語力には限りがあり、あまり多くの表現を使いこなすのは難しい。また、本校生徒の課題研究では実生活に根差したテーマを扱うことも多く、そのような応用的・分野複合的な内容については、一般的なマニュアルだけでは対応しきれない部分もある。

そのような条件を考慮すると、本校の生徒が課題 研究論文、特に要旨を執筆するために必要とする表 現を明らかにすることができれば、上述の英文要旨 作成のコストを減ずることが期待される。

2. 方法と結果

前項の目的を達するため、本研究では、本校 SSH 第 3 期以降の課題研究論文の英文要旨をコーパス $^{1)}$ 化し分析することとした。本研究では SS コース(第 3 期)・AS コース(第 4 期)の生徒によるもの 81 本を対象とした。このことから、本研究で用いるコーパスを、便宜的に「AS_Abstract_Eng_2015-2022 (ASAE1522)」と呼ぶこととする。

コーパスの管理と分析には $AntConc^{2)}$ を使用し、以下の分析を行った。

- (1) 語の出現頻度
- (2) コロケーション³⁾ の特徴
- (3) 出現語について汎用コーパスとの比較
- (4) 出現語について教科書コーパスとの比較

以下、順に詳述する。

(1) 語の出現頻度

全81本の英文要旨において、ある語が何回使われ

ているかを数え上げた。レンマ化 (lemmatization) 4) を行っていない状態での総語数 (token) と異なり語数 (type) 5) は以下の通りであった。

総語数:11,107語 異なり語数:2,122語

表 1. 英文要旨における語の出現頻度と広がり

順位	出現語	頻度	広がり
5	we	253	76
15	which	81	51
18	study	71	52
24	method	46	21
29	found	41	32
30	result	41	32
31	different	40	27
32	experiments	37	28
33	solution	36	15
34	water	36	14
36	experiment	34	21
41	using	31	23
42	zebrafish	31	7
44	number	29	15
45	conducted	28	24
46	results	28	22
47	between	27	21
50	purpose	27	25
53	surface	26	12
56	used	24	21
57	find	23	17
61	bubbles	22	3
62	conditions	22	16
63	showed	22	19
64	temperature	22	11
66	amount	20	15
67	metal	50	9
70	model	19	13
71	research	19	16
72	strength	19	6
73	structure	19	10
74	however	18	17
79	time	17	10
80	based	16	14
81	because	16	11
85	order	16	15
87	process	16	10
88	clarify (clarified 3, 3)	15	15
90	observed	15	11
91	patterns	15	6
97	granites	13	3
99	revealed	13	12

出現頻度の最上位層は the, of, and といった汎用的な機能語が占めるため、単純な頻度のみに着目することは、本研究における分析としてはさして有意義ではない。したがって、頻度リストから特徴的と思われる語を抜粋し、ASAE1522 全体の中で出現する回数(頻度)と、その語が出現する論文の数(広がり)とともに表1に示す。

科学論文であることから、study、method、result(s), experiment(s)等が高頻度で出現するのは当然のことである。また、zebrafish、bubble、metalといった語は比較的高頻度で出現しているが、広がりは小さく、特定の研究テーマの論文で偏って使用されていることがわかる。

そのような語を除き、特徴的と思われる4つの語 に着目する。

1) which (頻度 81. 広がり 51)

81 例すべてが関係代名詞としての用法である。科学論文では、実験装置や試料等について詳細に説明する必要があり、また、独自の手法を用いる場合などには既存の名詞では十分に説明できないことも多いことから、関係代名詞の需要が高いと考えられる。

2) however (頻度 18, 広がり 17)

18 例すべてが逆接の接続副詞としての用法である。この用法における同義語である but は7 例 (頻度7, 広がり7うち however も使用されているもの1)あり,これと合わせると25 例 (広がり23)になる。逆接の論理展開が使われているのが81本の論文中23本という割合を一概に評価することはできないが,要旨といえども,先行研究と当該研究の矛盾や,実験結果についての予想と実際の違いなど,相反する内容を整理して述べねばならない場合があるということである。

3) because (頻度 16, 広がり 11)

因果を表す接続詞である。文構造は異なるが同義で用いられている so は 6 例 (広がり 6 うち because も使用されているもの 2) あった。これと合わせると,因果関係を明示する表現が使われているのが 81本の論文中 15 本ということになる。

科学論文とは、基本的に事実の説明の積み重ねによって成り立つものであり、事実と事実の関係が明らかである場合には、それを明示する標識を必要としない。その意味で、「論文」とはいっても、テクスト・タイプとしては描写文や説明文の要素が強いと言える。

それに対して、 高校生が他に書くことの多い、 大

学入試等のいわゆる自由英作文では、人文・社会科学系のテーマについて自分の主張を述べ、それを読み手に納得してもらおうとする「意見文」が多い。この場合、必ずしも自明ではない事実関係を論理的に整理する必要があるため、因果関係を明示することが必要になる。この相違は、ESPの観点からは注目に値する。

4) solution (頻度 36, 広がり 15)

この語は CEFR ⁶⁾ では A2 レベルとされ、また近年では日本語でも外来語として「ソリューション」が定着していることから、高校生にとっては平易な部類に属する。しかし、ASAE1522 でこの語が使われている論文は 15 本に留まる。これは、この語が自然科学の文脈では、一般的な「解決(策)」ではなく、数学における「解」もしくは化学における「溶液」という意味で理解するのがふさわしいことによる。ASAE1522 ではまさにその通りの用例となっており、平易な一般語というよりは、特定の研究分野と結びつく専門用語の性格が強い。

同様に treatment も CEFR では B1 レベルであるが、一般的には「治療」あるいは「扱い」といった訳語があてられるのに対し、科学的文脈では、化学的な「処理」や心理学実験における「処遇」といった意味で用いられ、やはり専門用語に属する。この語は、ASAE1522 では 1 例だけ、the treatment of the silica-gel adsorbed in copper ions という化学の文脈で使われている。

このことは、English for General Purposes(EGP: 一般目的のための英語教育)ではそれほどの高難度語としては扱わない語句も、ESPにおいては、あらためて専門用語として扱う必要性を示唆している。

(2) コロケーションの特徴

出現頻度の高い語が、どのようなコロケーションで使われているかを観察したところ、いくつか特徴的なものが見られた。

1) result/results

result という語は、科学論文において実験結果等を述べる文脈では results という複数形で用いられるのが通例である。ASAE1522 でも、次のような用例が見られ、通例に沿う用法になっている。

the results of the experiments the results of ion orbit simulation the results of these observations the results showed that… the results suggested that...
the results corresponded to...

the results の後に of 句を伴って実験や調査について 述べるか, the results showed that…のように, 動 詞と that 節を伴って, 実験・調査の結果から導かれ る結論を述べる文脈を作っている場合が多い。

単数形の result でも、次のように複数形 results と同様の意味で使われている場合がある。

the result of our first experiment the result of the observation

これらの場合、単一の実験・調査結果を指しているが、科学論文の慣用上は複数形 results にしても文意を損ねることはないと思われる。

一方、単数形の result のコロケーションで最も多いのは as a result である(18/41 例)。要旨では論文の結論にも言及するため、実験・調査の結果を明示する標識として高校生にとって平易な as a resultが選好された可能性が考えられる。

2) found

この語は、we found that のコロケーションでの 出現頻度が高い。we also found that、we found out that や that 節の省略を含めると 27/41 例を占める。 次いで it was found that が 4/41 例となっており、 主に研究結果を述べる文脈で用いられている。科学 論文において I や we のような一人称主語を用いる ことについては、これを避けるべきという伝統的な 立場から積極的に容認する立場まで異なる考え方が あり、本校では容認する方針で指導していることが この結果の背景にあると考えられる。そして、表現 の平明さから、結果を導入する標識として we found that が選好されたということであろう。

3) using

「使う」の意味での use は使用例が多く,原形 use が 1 例(広がり 1),過去形・過去分詞 used が 24 例(広がり 21),現在分詞・動名詞 using が 31 例(広がり 23)であった。語形としては using が最も 多く,中でも by using のコロケーションが 16 例と 半数を占めた。また,分詞構文として using が使われている中にも by using と置き換えが可能な例があり,by using のコロケーションは生徒にとって使い 勝手の良い表現であるように思われる。

他に、ASAE1522で見られた特徴的なコロケー

ションとしては、addition と comparison に関する ものが挙げられる。

addition には「足し算」の意味があり、数学の 文脈ではそのように使われることも想定されるが、 ASAE1522 ではその用例は 1/10 例に留まり、残り in addition が 8 例 2 in addition to が 1 例 2 のあり、 特に追加を表す標識として in addition が選好されて いる実態がある。

逆に comparison については EGP では in comparison to/with のコロケーションで出会うことも多いが, ASAE1522 ではそれは 1/6 例に留まり、残り 5 例はすべて comparison value 「比較値」のコロケーションで 1 本の論文において集中して使われている。

コロケーションについても, ESP では, ある意味 で EGP からの「学び直し」を促すことが重要にな るように思われる。

(3) 出現語について汎用コーパスとの比較

AntConcでは、分析対象のコーパスを別のコーパスと比較したときに、統計的処理に基づき分析対象のコーパスに特徴的に現れる語を示す"Keyword"機能がある。この機能では、その語がれほど「特徴的」であるかが"Keyness"という指標で示され、数値が高いほど、その語がそのコーパスに特徴的に現れることを示す。また、AntConcでは、初期設定でアメリカ英語およびイギリス英語の汎用コーパス⁷⁾が用意されており、自作のコーパスをそれらと比較することができる。

本研究では、この機能を利用して ASAE1522 の特 徴を明らかにすることを試みた。比較する対象は、アメリカ英語の汎用コーパス (AmE06) とした。これは、日本の中学校・高校で使用している英語教材がおおむねアメリカ英語をもとにしており、生徒が触れる英語もアメリカ英語が多いことによる。

ASAE1522 を AmE06 と比べた際に Keyness の高かった上位 30 語を表 2 に示す (短縮語は除外した)。一人称主語 we については (1) でも言及したが,汎用コーパスとの比較においても有意に多く使われていることがわかる。これは,本校では課題研究に生徒がグループとして取り組むことが多く,また,科学論文では事実の報告が主となることから表れる傾向と考えられる。

また,前置詞 of が上位にあることも注目に値する。 of は英語の前置詞の中で特に幅広い用法を持ち, どのような分野であっても使われることが多いが, 科学論文では, まさにその多様な用法が, 複雑な内容を正確に伝えるのに活かされているということかもしれない。 of あるいはこれに類する機能語を正確に

表2. 英文要旨において特徴的に現れる語(対汎用コーパス)

mz /1.	th ru ≕r	वद के	
順位	出現語	頻度	Keyness
1	we	253	550.090
2	zebrafish	31	280.868
3	study	71	275.462
4	method	46	245.546
5	experiments	37	225.727
6	solution	36	213.602
7	experiment	34	209.290
8	bubbles	22	164.687
9	result	41	157.063
10	conducted	28	145.642
11	of	555	128.752
12	granites	13	117.763
13	surface	26	113.096
14	clarify	15	110.885
15	purpose	27	110.178
16	temperature	22	109.690
17	gel	14	100.241
18	alginate	12	97.263
18	ions	12	97.263
20	different	40	97.211
21	results	28	94.312
22	droplets	12	93.756
23	metal	20	91.963
24	fibers	12	90.795
25	found	41	89.958
26	which	81	88.967
27	conditions	22	85.589
27	showed	22	85.589
29	water	36	85.426
30	indigo	10	83.904
30	silica	10	83.904

運用できる力を養っておくことは、当然ながら EGP でも重要なことではあるが、ESP の観点からも重点的に扱う価値があると考えられる。

他に特徴的なこととしては、これも当然のことであるが、solution、temperature、ionsといった自然科学分野で広く使われる語や、zebrafish、granites、alginate、indigo、silicaといった、特定の研究テーマと結びついた語が多いことが挙げられる。生徒が課題研究のテーマに選ぶ分野は多岐にわたり、前節で述べたように応用的・複合的なテーマも少なくないことから、ESPにおいて生徒が必要とする語彙をすべてカバーすることは不可能である。しかし、生徒が課題研究について表現するための語句に、研究テーマの壁を越えてある程度の共通性があるのであれば、自然科学分野で広く使われる語については計画的に習得できるように手当てするメリットはある

表3. 英文要旨において特徴的に現れる語(対教科書コーパス)

順位	出現語	頻度	Keyness
1	we	253	261.710
2	of	555	162.852
3	study	71	117.372
4	which	81	71.802
5	method	46	58.803
6	zebrafish	31	58.323
7	result	41	58.219
8	experiment	34	55.879
9	conducted	28	52.674
10	experiments	37	51.277
11	solution	36	49.552
12	surface	26	48.909
13	the	893	48.479
14	this	118	43.225
15	purpose	27	43.152
16	bubbles	22	41.379
17	is	159	41.228
18	water	36	36.452
19	model	19	35.734
19	structure	19	35.734
21	our	44	34.446
22	using	31	34.255
23	conditions	22	34.142
23	temperature	22	34.142
25	results	28	32.516
26	different	40	30.953
27	by	106	29.744
28	clarify	15	28.207
28	patterns	15	28.207
30	gel	14	26.326
30	silica	10	83.904

ように思われる。

(4) 出現語について教科書コーパスとの比較

前項では汎用コーパスとの比較において ASAE1522の特徴を明らかにすることを試みたが、 生徒にとってより身近な英語のリソースとして教科 書を取り上げ、そこで出会う英語と課題研究論文の 英文要旨で必要となる英語の異同も分析することと した。

比較対象のコーパスとしては、「英語コミュニケーション I」および「コミュニケーション英語 II」の検定教科書 *ELEMENT English Communication I* および同 II (啓林館)の本文を用いた。このコーパスを便宜的に「ELECom」と呼ぶこととする。ASAE1522と ELECom の比較の結果を表 3 に示す。

上位に入る語は前項の汎用コーパスと比較した場

合と共通しているものが多い。上位30語の中では21語が共通しており、それ以外の語でも、30位をやや下回るところに位置するなど、総じて大きな違いは見られない。対汎用コーパス(表2)と対教科書コーパス(表3)の間ではKeyness指標の値に差があるが、コーパスのサイズが異なるため単純な比較はできない。

参考までに、AmE06を比較対象として ELECom について同様の分析を行ってみたところ、ELECom で特 徴 的 に 表 れ る の は euglena、asahi、momo、Jimmy、Nikkei、Watson といった、特定の文章のトピックに関わる語が多く、それ以外で両コーパスに何らかの特徴が広く見られるということはなかった。

3. 考察と今後の展望

本校生徒が英文要旨を書くのに必要な英語表現を 知るという本研究の目的に対しては,前節の分析で ある程度の知見を得ることができた。一方で,本研 究には乗り越えられていない限界や課題が残されて いる。

(1) 教員の指導の影響

本研究では最終的な成果物である課題研究論文の 英文要旨をコーパス化して分析した。しかし、それ らはすべてが生徒の手によるものではなく、1. で 述べたように、英語科教員が生徒やその指導教員と やり取りをしながら作り上げたものである。した がって、すでにそこには、個々の論文に関わった英 語科教員の表現上の好みが反映されており、中立的 なデータではない。より生徒が言い表したい「生の 声」に近い表現を収集するには、和文要旨や本文(和 文)も分析の対象とするのが良いと思われる。そし て、AIによる自動翻訳なども活用して、翻訳者のバ イアスを極力排除したときにどのような表現が特徴 的に現れるかを分析してみたい。

(2) サンプルの偏り

本研究では全81本の論文を分析対象としたが、本校における課題研究テーマは、大別しても物理・化学・生物・地学・数学といった領域に分かれ、個別のテーマになると、さらに細分化される。したがって、未来の生徒が必要とする英語表現をカバーするには、参照するサンプルに偏りがあり、サンプルサイズ自体もまだ小さい。GSコースの論文のうち自然科学分野のものも分析対象に含めるなど、より広範なサンプルを用いることが必要となろう。

(3) GS コースとの比較

本校において課題研究論文に関わる英語指導を安定して持続させていくには、GSコースの論文の分析も欠かせない。GSコースでは研究分野に人文科学・社会科学も含まれるため、研究テーマがより多岐にわたる。ASコースとGSコースで課題研究論文において必要とされる英語表現にどのような違いや共通点があるのかは、本校に適したESP開発において重要な意味を持つ。

(4) 高校生にとって使いやすい表現の選定

前節での分析において、たとえば追加を表す標識として、Besides や Furthermore、Similarly といった語もある中で、In addition が選好されている実態が明らかになった。これらの語には、それぞれ独自のニュアンスがあり、科学論文においても必要に応じて使い分けられるものではあるが、高校生は、多様な語を使い分けるほどに英語に熟達していないため、かれらにとって平易な In addition を方略的に用いていると考えられる。

このように、従来のマニュアル類が示すよりも、より限定的な語彙を多用途に「使いまわす」のが高校生の特徴である。この傾向を指導にも活かし、高校生にとって扱いやすく、ある程度汎用的に使える平易な表現を選定し、体系的に学習させることで、生徒の英文要旨作成が、より容易になることが期待される。

これについては、上記(2)(3)とも密接に関わることであり、さらに多くのサンプルをもとにして表現を抽出していく必要がある。

これらの課題については、今後も引き続き多角的に研究を進め、高校生が科学研究について英語で口頭発表や論文執筆するのを支援する ESP プログラムおよび教材の開発を進めたい。

注

- 1) コーパス (corpus) とは、言語の実際の使用例を蓄積したデータベースであり、分野による頻出語の偏りや時系列での用法の変化など、さまざまな角度から言語の実態を明らかにするのに用いられる。
- 2) AntConc は、Laurence Anthony 氏(早稲田大学)が開発した、コーパスの構築とそれを用いた分析ができるソフトウェア(コンコーダンサーconcordancer)である。本研究ではバージョン 4.1.4 を用いた。
- 3) コロケーションとは、ある語が他のどの語と共

- に使われること(共起)が多いかという, 語と語の結び付きやすさである。
- 4) レンマとは、複数の語形を持つ語のうち、辞書の見出し語になる基本的な形である。たとえばhave、has、had、havingのレンマはhaveである。レンマ化とは、コーパスに現れるこれら異なる語形を1つのレンマとして数え上げるための処理である。本研究では、語形の違いによる用法の違いにも着目したかったため、レンマ化を行わなかった。
- 5) 総語数とは、あるデータに含まれる語数を数える際に、同じ語が複数回使われていてもそれぞれ 1語と数える場合の語数であり、異なり語数とは、 同じ語は何度使われても1語と数える場合の語 数である。例えば、The black dog and the black

- cat played with the black stones. という文では, 総語数は12, 異なり語数は8となる。
- 6) CEFR とは Common European Framework Reference for Languages「ヨーロッパ言語共通参照枠」の一般的略称である。これは、EU が複言語主義(plurilingualism)に基づき、複数の言語の習熟度や運用能力を同一の基準で評価するための枠組みである。最も基礎の A1 レベルから最も熟達した C2 レベルまでの 6 段階に分かれている。
- 7) 汎用コーパスとは、コーパスのうち、できるだけ特定のジャンル(新聞・小説・ラジオ放送など)や分野(自然科学・政治・芸能など)に偏らないようにまんべんなく言語データを収集したものである。本研究で用いた AmE06 は、2005 年~2007年のアメリカ英語を収集したコーパスである。

Developing English for Specific Purposes Course for Natural Science at the Secondary Level:

A Vocabulary Perspective

Taiki YAMAOKA

Abstract:

This study aims to obtain data about English expressions which high school students need in writing English academic papers on natural sciences, especially their abstracts. This would lead to developing teaching materials and designing courses for natural sciences in English for Specific Purposes (ESP). The author created a corpus from the English abstracts of research papers written by students at Hiroshima University High School from 2015-2022, from which he analyzed the characteristics of words and collocations. In addition, the author compared the corpus to a general corpus and one created from English textbooks to find characteristic expressions preferred by high school students in writing natural science papers. The results suggest that high school students produce different types of discourse in their science papers than they do in English for General Purposes (EGP) courses. Specifically, they primarily describe facts in science papers rather than express their opinions as they do in EGP writing. It was also suggested that high school students prefer certain expressions due to limitations in their English proficiency.

資料 AntConc の画面イメージ (KWIC 表示の例)

