

# 大学入試センター試験が映し出す英語

## — 電子コーパスとして読む英語問題\*

北 和丈、樫村 真由、伊澤 高志  
佐藤 繭香、瀧口 美佳、土屋 結城

### 1. 研究の背景

日本の英語学習史を顧みる近年の概説書や研究書（e.g. 伊村, 2003; 斎藤, 2007; 江利川, 2011）では、時代ごとの英語学習像を描き出すよすがとして、多くの学習者の共通体験を形作ったと考えられる特定のテキストの存在が指摘されている。教科書であれば明治期の輸入読本である『ナショナル・リーダー』や戦後の英語教育で長きに渡って使用された「ジャック・アンド・ベティ」、参考書であれば「南日」「山貞」として知られる南日恒太郎と山崎貞、「受験英語の神様」伊藤和夫らの著作群がその代表例である。各時代における日本の英語学習者の母数そのものが大きく異なるという但し書きは付くにせよ、いずれもかなりの普及度を誇ったものであるという点についてはほぼ疑う余地もあるまい。

さて、学習者や学習媒体が多様化した現代において、これらに匹敵するほどの規模や割合で人口に膾炙した英語テキストは存在するだろうか。この問いに対する答えとして様々な候補が挙がり、絞り込みが困難になるであろうことは容易に想像がつくとしても、その一覧から大学入学者選抜大学入試センター試験（以下、センター試験）が抜け落ちることはまず考えられない。第1回試験が実施された1990年から現在にかけて、受験者の大半を占める18歳人口が約40%減少している一方で、センター試験の志願者数は微増を続けながら安定して50万人前後を維持し続けている（図1）。こ

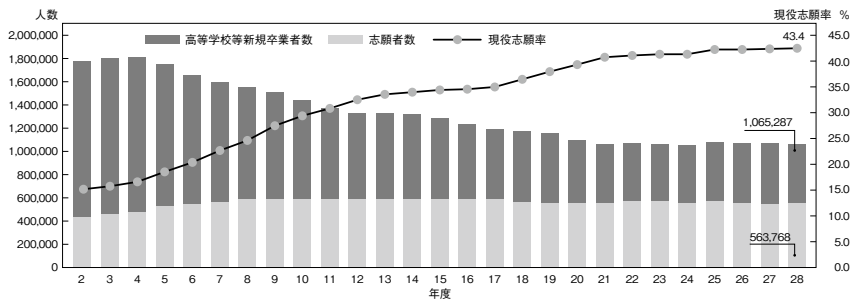


図1 センター試験の志願者数及び現役志願率の推移  
(大学入試センター(2016)より転載)

の志願者のおよそ95%が外国語(英語)を受験する(大学入試センター, 2016, p.18)ことを考慮すると、センター試験の英語問題として提示されるテキストに触れてきた日本の英語学習者は少なく見積もっても1,400万人以上に上ることになる。これに加えて、大学入試の一部であるセンター試験は学習者だけでなくその教育者も綿密に読み込むテキストであること、受験準備の過程で学習者が複数の過去問に取り組む場合が多いことを考え合わせれば、現代の日本における英語像の一部がセンター試験によって形成されてきたという可能性はいよいよ否定しがたいものになる。

このような潜在的な影響力の大きさにもかかわらず、センター試験の英語問題のテキストを、能力測定手段としてではなく内容の側面から検討しようという学術的な試みはこれまでのところほとんどなく、あまつさえ、その功罪すら十分に評価されないまま、2020年をもってセンター試験は30年近い歴史の幕引きを余儀なくされている。この状況に対する抵抗の一つとして企図されたのが、本誌第70号掲載論文(土屋他, 2018)で展開したセンター試験の英語問題に対する文学的読解であり、その執筆者によって構成されるプロジェクトでは、現在も同様の研究手法による批判的精読の取り組みが継続している。

本論文で記述する研究もこのプロジェクトの一環として行われたもので

はあるが、先述した論文とはその視点と手法が異なる。センター試験の英語問題を一種の社会的構築物と見なしたうえで、その内容に表れる思想や文化を詳らかに読み解いていくという方針は共通しているが、本論文ではその作業において、読み手としての人間の役割をいったん保留する。代わりに第一の読み手を務めるのはコンピュータである。センター試験で提示される紙媒体の英語テキストを電子テキスト化したうえでコンピュータに電算処理を行わせ、センター試験の英語が総体として持つ傾向を数値データとして導き出し、その結果を改めて人間の視点から解釈することによって、今後のプロジェクトがさらなる批判的精読を行う際の仮説を形成するというのが、この論文に通底する基本方針である。近年の文学研究では「デジタル・ヒューマニティーズ」や「ポスト・ヒューマニズム」、言語学研究では「コーパス言語学」といった用語で括られるこの種の研究（e.g. Ramsay, 2011; Hunter & Smith, 2012; Jockers, 2013; Moretti, 2013; O'Halloran, 2017）は、ともすれば人間の読み手の直感による精読を軽視・否定するものと捉えられかねない向きもあるが、それは本論文の本意とするところではない。日本の英語学習者が持つ英語像の構築に少なからず寄与した可能性のあるセンター試験の英語テキストをできる限り精緻かつ広範に検討するべく、人間の視点とコンピュータの視点を相補的に活用することがプロジェクト全体の狙いであり、本論文はそのための布石として理解されるべきものだとすることを申し添えておきたい。

## 2. 素材と道具

前節で述べたとおり、本論文はセンター試験の英語問題を電子テキストとして読み解こうとする、言語学で言うところのコーパス分析の試みであるが、ひとえにコーパス分析と称しても、素材としてどのようなコーパスを用い、それをどのような道具で解析するかによって、得られる結果は必然的に変わってくる。とは言え、センター試験の英語テキストに関する何らかの仮説検証を行うのではなく、その前段階としてまず仮説形成を行う

ことを主眼としている本論文の場合、分析対象としてどのような素材や道具が相応しいかを仮説に照らして選定・作成するという選択肢は原理的に存在しない。以上を踏まえ、本節では、プロジェクトとして入手することのできた素材と道具、およびそれを用いてどのような問題に取り組むことが可能となるかについて記述する。

## 2.1 素材

### (1) センター試験英語問題

本論文の主たる分析対象であるセンター試験の英語テキストについては、1990-2017年度分の（リスニングを除く）全筆記問題を電子テキスト化したうえで、必要に応じて詳細な分析ができるよう、年度別・大問別のファイル分けを行った（書誌情報は土屋他(2018)を参照）。コーパス全体としての総語数は約10万語である。一般的なコーパス分析においては、各語の統語的・意味的情報をタグとして付加する場合も多いが、タグ付けという行為そのものが一種の解釈を伴うものである以上、一切の仮説を排した状態での分析を企図する本論文においては趣旨に反するものと判断した。

### (2) COCAサンプル版

コーパス分析においては、分析対象となるコーパス（分析コーパス）の特徴を明確に示すべく、基準となる別コーパス（参照コーパス）との比較が常套手段として行われる。例えば本論文の場合、センター試験という特殊状況での英語使用が持つ特性を炙り出すならば、ジャンルや文脈を問わない一般的な英語使用の傾向との比較が有効に働くはずである。また、日本の学校教育における英語指導がアメリカ英語を基本としている（らしい）ことを考え合わせれば、センター試験に対する参照コーパスもアメリカ英語の有様を代表するようなものであることが望ましい。この目的に合致するコーパスの一つが、ブリガム・ヤング大学の Mark Davies 氏が構築した The Corpus of Contemporary American English (COCA) である。奇しくもセンター試験開始の1990年から現在まで継続的に収集された、総語数5億語超にも及ぶ様々なジャンルのアメリカ英語から成るこのコーパスは、年別・

ジャンル別にファイル分けされているだけでなく、無作為抽出で約170万語にまで圧縮したサンプル版としても入手可能である。データ処理の簡便さを考慮すれば、このサンプル版の使用は選択肢として十分に妥当である（なお、本研究で使用したサンプル版は2012年までのものである）。

### (3) 実用英語技能検定2級問題

比較によるコーパス分析では、有意義な比較を可能にするほどの基盤を共有した参照コーパスを複数用意できれば、得られる結果がより重層的になる可能性がある。また、その共通基盤の具体性が高まれば高まるほど、分析コーパスと参照コーパスは射程を絞り込んだ精緻な形で比較することができるようになる。本論文に即して言えば、センター試験は「日本の高校卒業程度の受験者を想定した大規模な試験」という特性を有しているので、できるだけこれに近い特性を持つコーパスが比較対象として存在するならば、センター試験の英語テキストの全体像に明確な輪郭を与えられる見込みは高まりそうである。この条件を満たす有力候補の一つが、実用英語技能検定（以下、英検）2級の問題である。日本の高校卒業程度の英語力が合格の目安とされているこの試験は、幸いにも主催者が直近3回分の過去問を公開している（日本英語検定協会, 2018）だけでなく、出版各社による過去問題集の販売も長年に渡って行われている（e.g. 旺文社, 2017）ので、それほど困難を伴わずに入手することが可能である。分析対象たるセンター試験の実施期間も考慮に入れた結果、最終的には1989年度第1回～2016年度第2回の（リスニングを除く）全筆記問題を年度別・回別・問題種類別のファイルとして電子テキスト化し、総語数約18万語のコーパスを構築することができた。入学試験であるセンター試験と検定試験である英検との差異に留意する必要はあるものの、水準の近い二つの試験から構築した英語テキストに何らかの顕著な共通点や相違点が見出せるとすれば、それだけでも学術的には興味深い発見と言える。

## 2.2 道具

コーパス言語学の基本は、電子テキストのなかの言葉を「数え」たうえで、

数えたその結果の意味合いを解釈することにあるが、この両方を簡便に行える機能を備えたものとしてしばしば用いられるのが、多機能コンコーダンサである。例えば、入手のしやすさや操作性の良さなどから判断して本論文で使用することにしたAntConc (version 3.4.4)(Anthony, 2016)には、分析コーパスに見られる全ての語を頻度順に一覧表示するWord List Tool、検索語の用例を一覧表示するConcordance Tool、検索語の出現箇所をバーコードとして視覚化できるConcordance Plot Tool、検索語の左右で共起しやすい語をその確率順に一覧表示するCollocates Toolといった複数の機能が備わり、様々な目的に応用可能であることが知られているが、なかんずく本論文の目的に即して最も有効と思われたのがKeyword List Toolである。分析コーパスを参照コーパスと比較した際に顕著な頻度を示す特徴語（キーワード）を、その顕著さ（Keyness：キーワード指数）の順に一覧表示することができるこの機能は、合理的な比較ができると思われる複数のコーパスが手元に揃っていた本研究の状況に合致していた。かくして、この機能によって特定されたキーワードの分析からセンター試験の英語テキストに表れる特徴を垣間見る、という本論文の方向性が選択されることになった。もちろん、この種の分析が直接的に関与するのは語彙の使用状況であって、例えば統語や語用にまで視野を広げるためにはさらに別の方策が必要となるが、本論文はあくまでもセンター試験の英語テキストを読み込むうえでの糸口を探ることに主眼を置くものであり、これだけをもって全てを論じ尽くす意図はないということを改めて強調しておきたい。

### 2.3 語り得ること

以上に記述した三つの素材と一つの道具を駆使すると、例えば下記のような課題に取り組むことが可能になるはずである。第3節では、このそれぞれの課題に対して、本研究からどのような示唆が得られ、今後の研究に向けたどのような仮説が引き出せたのかを詳述していくことになる。

- (1) センター試験はアメリカ英語の一般的な様態をどこまで反映しているか  
先述したとおり、センター試験は日本の英語教育の実情に即して、アメ

リカ英語を基本とした英語テキストを構築しているはずではあるが、それがアメリカ英語一般の様態をどこまで反映したものと言えるかについては定かではない。この問いに答えるための方策の一つが、センター試験を分析コーパス、COCAを参照コーパスとした比較分析である。この二つのコーパス間に顕著な相違点がないとすれば、センター試験はアメリカ英語の姿を極めて忠実に映し出すものということになり、一方で顕著な相違点が見られるとすれば、そこからセンター試験に特有のイデオロギーを嗅ぎ取ることができるかもしれない。

(2) センター試験は日本の同水準の英語試験に比してどのような特徴があるか

ただし、仮にセンター試験がアメリカ英語一般とは異なる様相を呈しているとしても、それが日本の同水準の英語試験に共通する特徴だとすれば、殊更センター試験のみに限った傾向ではないということになる。また、仮に日本の英語試験があまねくアメリカ英語一般からはかけ離れていたとしても、だからと言って日本の英語試験を個別に見た場合に全てが一様な特徴を示すとは限らない。これらの点を確認するうえで、本研究が準備したコーパスからは二つの方法を採用することができる。一つはCOCAを共通の参照コーパスとしたうえでセンター試験と英検2級問題をそれぞれ分析コーパスとする方法、もう一つはセンター試験と英検2級問題を相互に分析コーパス・参照コーパスとして入れ替える方法である。第一の策ではアメリカ英語一般を基準とした両試験の立ち位置、第二の策では日本の英語試験という共通項を踏まえたうえでの両試験の性格が見えてくることが期待される。

(3) センター試験に何らかの経年変化は見られるか

センター試験開始以降の、ほぼ平成という時代と重なる30年近い時間は、センター試験の内容そのものに影響を与えかねない様々な社会的変化を経験している。それが及ぼした影響の規模を把握するべく、センター試験における経年変化の有無を確認するとすれば、一つにはCOCAを参照コーパス、特定の期間別に分割したセンター試験をそれぞれ分析コーパスとする方法、もう一つにはその期間別に分割したセンター試験を分析コーパス・

参照コーパスとして相互比較する方法があり得る。この場合、センター試験の分割基準をどこに定めるかが悩ましいところではあるが、センター試験は各時期の学習指導要領に準拠して作成されていることが知られているので、この学習指導要領が切り替わる時点を大きな契機と見なした分割方針を採ることとした。

### 3. 分析と仮説

#### 3.1 センター試験はアメリカ英語の一般的な様態をどこまで反映しているか

センター試験の英語テキストとアメリカ英語一般の距離を測るべく、COCAを参照コーパスとした場合のセンター試験の特徴語を、キーワード指数の高い順に示したのが表1である<sup>1</sup>。このなかで、頻度（表では“Freq”で示している数値）が高くないもの、試験以外の状況では登場する見込みが低いと考えられるもの<sup>2</sup>、分布が特定の設問のみに偏っているもの（固有名詞など）<sup>3</sup>、単独では意味内容が判然としにくいもの（前置詞や助動詞など）を除外して考えると、残る特徴語からは大きく分けて二つの傾向を指摘することができる。

表1 COCAを参照コーパスとしたセンター試験の特徴語(キーワード指数上位35語)

Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword
1	309	913.738	b	13	25	145.862	cannot	25	27	120.792	friendships
2	114	485.568	japan	14	72	144.352	english	26	110	118.022	however
3	78	260.937	japanese	15	971	136.537	you	27	32	117.865	anna
4	56	243.948	opera	16	97	133.505	friends	28	24	117.506	margaret
5	369	200.15	people	17	30	131.16	tokyo	29	27	113.026	quilt
6	46	199.358	paragraph	18	22	128.359	himawari	30	264	111.777	how
7	34	198.372	emi	19	22	128.359	yuki	31	19	110.855	tsuyoshi
8	32	186.703	corrigan	20	25	127.128	mimi	32	69	109.351	music
9	36	184.479	oranges	21	33	126.379	monkey	33	210	109.252	many
10	160	169.789	children	22	28	126.116	kate	34	369	107.032	can
11	33	159.718	angela	23	21	122.524	sns	35	18	105.021	swip
12	33	159.718	dictionary	24	2988	121.223	to				



一つは、第2位の“japan”、第3位の“japanese”、第17位の“tokyo”から明確に窺える、日本に関する話題への偏向である。これと同じ文脈で重要な情報を与えてくれるのが、第14位の“english”だと言えよう。Concordance Toolで確認した限り、センター試験に登場する“English”は全て「英語」を意味するものとして使用されているが、その英語についてわざわざ“English”という言葉で頻繁に言及しなければならないのは、それが遠くにある対象であって、普段用いている言語ではないからである。つまり、センター試験の英語テキストには、身近な日本の話題を、身近ではない英語で表現するという、日本中心の視点が色濃く滲み出ていると推測できるのである。もちろん、これは「日本」の「英語」試験としては当然の宿命でもあって、仮にこのような日本本位の英語テキストが構築されていたとしても、それをすぐにナショナリズムや国粹主義と結びつけて批判するのが短絡的に過ぎるのは言うまでもない。それでも、上記の特徴語がCOCAとの比較で浮かび上がってきたという観点から強調しておかねばならないのは、センター試験の英語テキストで描かれているのが、本当に日本なのかどうかはともかく、少なくともアメリカとは異なる世界であるらしい、という点である。

では、アメリカではないその世界は、どこに由来するのだろうか。この問いへの手掛かりを与えてくれるのが、表1の特徴語に見られる二つ目の傾向である。第15位の“you”、第16位の“friends”、第25位の“friendships”といった語の顕著さから判断すると、COCAとの比較で見ると、センター試験の英語テキストでは直接的かつ友好的なやり取りが数多く描写されているように見える。また、第10位の“children”に象徴されるように、言及される人物は相対的に年少者が多いらしい。ここで想起されるのは、センター試験の主たる受験者が、一般的に高校卒業の年齢とされる18歳の若者だという点である。つまり、センター試験は、受験者の大半を占めるはずの18歳人口に対する配慮ゆえか、その年齢までに誰もが主体として経験していると想定できる身近なやり取りや出来事を描き込む割合が高くなっている、という仮説が立つのである。この仮説は、先述した日本中心の視点

とも矛盾しない。受験者の大半が日本の高校卒業者なのだとすれば、設問の内容として日本に関連した話題を選択することは、受験者に予備知識を要求する可能性が低い分、試験を行ううえでの不公平を避けることにもなるからである（その配慮の結果として、多少なりとも歪んだ英語像を提示してしまうことの妥当性には大いに議論の余地があるが、それは精読を含めた別の検証を俟たなければならない）。

3.2 センター試験は日本の同水準の英語試験に比してどのような特徴があるところが、センター試験から透けて見えるこのような配慮は、日本におけるほぼ同水準の英語試験であるはずの英検2級問題にはほとんど見られない。それどころか、英検2級問題には、この検定試験に特有のものとする思える二つの傾向が見て取れる。

表2 COCAを参照コーパスとした英検2級問題の特徴語（キーワード指数上位35語）

Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword
1	572	1446.383	b	13	223	280.209	company	25	107	193.760	amount
2	1024	894.097	people	14	149	275.139	animals	26	780	193.507	more
3	258	840.344	scientists	15	1300	270.463	they	27	155	189.952	companies
4	6794	693.516	to	16	290	265.475	help	28	398	189.523	make
5	638	607.922	many	17	97	262.874	ocean	29	69	186.973	drivers
6	379	476.363	use	18	767	243.312	can	30	93	183.578	expensive
7	293	395.628	however	19	80	224.397	electricity	31	593	183.563	new
8	361	392.677	used	20	727	223.518	will	32	114	181.912	environment
9	84	384.962	cannot	21	103	223.241	plants	33	109	177.236	cars
10	175	360.051	countries	22	80	218.493	chemicals	34	75	176.817	traffic
11	1473	339.660	are	23	223	215.634	number	35	259	175.183	money
12	2473	282.460	is	24	186	200.966	able				

表3 英検2級問題を参照コーパスとしたセンター試験の特徴語（キーワード指数上位35語）

Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword
1	1258	500.594	i	13	32	69.616	pat	25	65	53.891	felt
2	971	229.183	you	14	661	68.278	he	26	46	53.194	table
3	914	112.709	was	15	114	66.868	japan	27	32	51.606	anna
4	56	112.581	opera	16	33	63.59	angela	28	133	51.495	it's
5	368	111.054	my	17	51	59.462	author	29	78	51.439	that's
6	1748	100.663	and	18	27	58.738	quilt	30	43	50.429	dance
7	36	78.318	oranges	19	33	58.103	dictionary	31	33	49.731	sat
8	403	77.232	we	20	113	57.704	right	32	211	48.569	did
9	34	73.967	emi	21	243	56.915	me	33	134	48.392	us
10	33	71.791	monkey	22	248	55.186	like	34	97	48.275	friends
11	46	71.465	paragraph	23	72	54.55	oh	35	87	48.22	yes
12	32	69.616	corrigan	24	25	54.387	mimi				

COCAとの比較による特徴語を示した表2から窺える、英検2級問題の第一の、かなり明白な傾向は、第3位の“scientists”をはじめ、第14位以下の“animals”、“ocean”、“electricity”、“plants”、“chemicals”、“number”といった特徴語から幅広く読み取れる、広義の自然科学に関連した話題への偏向である。一方、第二の傾向としては、顕著さの度合いではやや下回るかもしれないが、第13位および第27位の“company”と“companies”、第30位の“expensive”、第35位の“money”といった特徴語から、総じて実利・実用重視の価値観が垣間見える。二つをまとめて「モノ」と「カネ」の英検でも言えそうなこれらの方向性は、表1と改めて見比べれば一目瞭然、センター試験には皆無に近いものである。

このようなCOCAを媒介にした場合の差異を念頭に置いたうえで、英検2級問題とセンター試験とをさらに直接比較した表3<sup>4</sup>を見てみると、センター試験には、第1位の“i(I)”を筆頭に、“you”、“my”、“we”、“me”といった、「わたし」と「あなた」による直接的なやり取りを示唆する代名詞が多用されていることがわかる。これは正に、前項で指摘した「主体として経験していると想定できる身近なやり取り」を特徴づける語群にほかならない。比較で言うならば、「公」の英検2級と「私」のセンター試験、「大人」の英検2

級と「子供」のセンター試験といった仮説が説得力を持ってしまうほどの傾向が窺えるのである。もっとも、これは両試験のどちらが優れているかという問題ではない。強いて言えば、二つの試験はどちらも何らかの形で偏っているが、その偏り方が異なる、ということである。

### 3.3 センター試験に何らかの経年変化は見られるか

さて、総体として上記のような特徴が指摘できるセンター試験は、その30年近い時間のなかで、何らかの変化を遂げてきているのだろうか。第2節 第3項で触れたように、本研究では分析対象としての素材（COCAとセンター試験の英語問題）を、センター試験が作問される際に参照されているはずの高等学校の学習指導要領（以下、「学習指導要領」は高等学校のものを指す）の区分で分割し、その比較によって経年変化の検証を試みた。表4に示すように、学習指導要領には戦後8回の改訂が行われている。これらの各改訂版の施行時期、およびセンター試験受験者が高校で教育を受けた時期を考慮すると、表5のとおり、センター試験の問題は昭和53（1978）年8月告示のものから四つの学習指導要領を元に作られたことになる。こ

表4 高等学校の学習指導要領改訂時期

これまでの学習指導要領	
S22 (1947)年 3月	〔試案〕文部省発行
S26 (1951)年 7月	〔試案〕文部省発行
S30 (1955)年 12月	文部省発行
S35 (1960)年 10月	文部省告示
S45 (1970)年 10月	文部省告示
S53 (1978)年 8月	文部省告示
H元 (1989)年 3月	文部省告示
H11 (1999)年 3月	文部省告示
H21 (2009)年 3月	文部科学省告示

注1 土屋他（2018）より転載

注2 二重線以下がセンター試験作問に係わる

表5 センター試験が参照する学習指導要領

学習指導要領	センター入試の表記年度
S53年度告示版（第1期）	1990 (H2) ～ 1996 (H8)年
H元年度告示版（第2期）	1997 (H9) ～ 2005 (H17)年
H10年度告示版（第3期）	2006 (H18)～ 2014 (H26)年
H20年度告示版（第4期）	2016 (H28)年～

表6 COCAとセンター試験による  
比較のかけ合わせパターン

分析コーパス	参照コーパス	結果
S53年度告示版 センター試験問題	1990～1996年 COCA	表7 (a)
H元年度告示版 センター試験問題	1997～2005年 COCA	表7 (b)
H10年度告示版 センター試験問題	2006～2012年 COCA	表7 (c)

の区分を便宜上、第1期～第4期と呼ぶこととする。

センター試験の経年変化の有無を検証するために用いた方法は二つである。第一の方法は、表5の区分によりCOCAとセンター試験をそれぞれ時期別に分割したうえで、前者を参照コーパス、後者を分析コーパスとして特徴語を探るものである。ただし、本研究で使用したCOCAサンプル版の収集期間は1990年から2012年であり、センター試験の実施期間とはわずかにずれが生じるため、第4期のセンター試験問題は比較対象から除外した。従って、具体的な比較は表6で示すような三つのかけ合わせのパターンで行われ、その結果として得られたのが表7 (a)～(c)である<sup>5</sup>。

一貫してキーワード指数の高い特徴語が必ずしも多いとは言えないこれらの表のなかで、それでもなお際立った傾向として見えてくるのは、本節第1項で触れた日本に関する話題への偏向が、時期を追うごとに強くなっているということである。表7を時系列的に眺めても、“japan”、“japanese”、“tokyo”といった日本にまつわる語の登場頻度が高くなっているのは明らか

表7 分割版COCAを参照コーパスとした分割版センター試験の特徴語  
(キーワード指数上位12語)

(a) 第1期				(b) 第2期				(c) 第3期			
Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword
1	75	277.889	b	1	107	340.775	b	1	46	200.133	japan
2	47	217.609	author	2	46	214.691	japan	2	32	152.355	paragraph
3	32	214.271	corrigan	3	34	213.720	emi	3	55	149.771	english
4	33	187.909	monkey	4	33	198.500	angela	4	68	148.421	b
5	24	160.703	margaret	5	37	160.472	japanese	5	210	136.337	can
6	24	134.399	bats	6	25	148.758	mimi	6	27	132.652	dictionary
7	21	122.745	leisure	7	31	140.299	pat	7	29	123.948	japanese
8	14	78.114	yacht	8	36	129.655	robert	8	32	120.035	anna
9	19	77.246	japan	9	27	127.420	quilt	9	22	114.994	himawari
10	253	76.622	he	10	25	127.214	kate	10	70	111.681	students
11	14	72.201	tofu	11	19	119.432	tsuyoshi	11	21	109.767	sns
12	10	66.960	bordes	12	25	111.580	cheese	12	21	109.767	tokyo

だが、なかでも“japan”の動向は看過しがたい。第1期でも第9位に顔を出しているこの語は、次の第2期においては第2位、第3期に至っては第1位に上り詰めているのである。このことから窺えるのは、センター試験においてはその科目が「外国語（英語）」であっても、経年変化で見ると、扱う題材・話題そのものは日本色が強まっているのではないかということである。

経年変化を捉えるうえで採用したもう一つの方法は、参照されている学習指導要領が異なるセンター試験問題同士の比較である。この比較では、年代が新しいセンター試験問題群に顕著な特徴を確認するべく、古いほうのセンター試験問題群を参照コーパス、新しいほうを分析コーパスに設定したため、結果として表8に示す六つの組み合わせが成立した。表9(a)～(f)がそれぞれの結果である。

六つのかけ合わせパターンのうち、キーワード指数が100を超えるものは、“oranges”、“navel”、“opera”の3語しか存在しなかった。このうちオレンジに関連した“oranges”および“navel”の2語は、アメリカ合衆国の果物輸出入状況推移とオレンジ輸送量の年間推移を読み取る2016年度第4問でのみ使用されている。また“opera”は、ニューヨークの電話ボックスで見つけた木製の猫をめぐる事件を描く1997年度第6問の物語文、オペラの現状と課題に関する2016年度第6問の説明文にのみ現れる語である。局所的な分布を示すこれらの3語以外に顕著な特徴語が見られないとすれば、比較したコーパスには語彙の種類や登場確率について大きな差異がないことになるので、センター試験は使用語の選択に関してかなりの一貫性を維持していることが窺い知れる。これはもちろん、センター試験が学習指導要領に準拠しているという事実を明確に裏付けるものであると同時に、本節第1項で述べた受験者間での公平性に対する配慮が垣間見える結果であるとも言えよう。

表8 センター試験同士による  
比較のかけ合わせパターン

分析コーパス	参照コーパス	結果
H元年度告示版 センター試験問題	S53年度告示版 センター試験問題	表9 (a)
H10年度告示版 センター試験問題	S53年度告示版 センター試験問題	表9 (b)
H20年度告示版 センター試験問題	S53年度告示版 センター試験問題	表9 (c)
H10年度告示版 センター試験問題	H元年度告示版 センター試験問題	表9 (d)
H20年度告示版 センター試験問題	H元年度告示版 センター試験問題	表9 (e)
H20年度告示版 センター試験問題	H10年度告示版 センター試験問題	表9 (f)

表9 分割版センター試験を相互比較した場合の特徴語（キーワード指数上位12語）

(a) 第2期（参照：第1期）				(b) 第3期（参照：第1期）				(c) 第4期（参照：第1期）			
Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword
1	167	45.722	her	1	510	40.716	is	1	72	167.997	oranges
2	206	45.165	she	2	210	31.344	can	2	43	100.332	opera
3	36	36.419	robert	3	35	26.828	music	3	34	79.332	navel
4	34	34.396	emi	4	32	24.529	anna	4	27	49.936	fresh
5	33	33.384	angela	5	70	22.766	students	5	20	46.666	valencia
6	31	31.361	pat	6	117	21.568	or	6	19	44.333	friendships
7	30	30.349	music	7	51	19.579	mr	7	18	41.999	imports
8	27	27.314	quilt	8	55	19.184	english	8	46	38.136	us
9	27	27.314	stress	9	25	19.163	adults	9	18	34.910	singers
10	25	25.291	cheese	10	274	17.712	are	10	14	32.666	market
11	25	25.291	mimi	11	31	17.151	dance	11	16	30.472	tour
12	24	24.279	tommy	12	22	16.863	drinks	12	13	30.333	cultural

(d) 第3期（参照：第2期）				(e) 第4期（参照：第2期）				(f) 第4期（参照：第3期）			
Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword
1	510	52.882	is	1	72	211.994	oranges	1	72	251.492	oranges
2	32	34.315	anna	2	34	100.108	navel	2	43	150.196	opera
3	32	34.315	paragraph	3	43	72.696	opera	3	34	118.760	navel
4	181	25.238	their	4	27	61.557	fresh	4	27	86.064	fresh
5	210	25.106	can	5	20	58.887	valencia	5	20	69.859	valencia
6	70	23.994	students	6	19	55.943	friendships	6	18	62.873	imports
7	22	23.591	himawari	7	18	52.999	imports	7	18	62.873	singers
8	21	22.519	sns	8	18	52.999	singers	8	56	50.929	b
9	115	20.768	has	9	46	51.872	us	9	46	43.383	us
10	55	20.755	english	10	19	39.981	student	10	19	36.616	friendships
11	19	20.374	movie	11	18	37.337	town	11	12	35.248	australia
12	122	20.091	your	12	22	16.863	drinks	12	12	35.248	mexico

一方、全般的には特筆すべき差異の少ない六つの比較のうちで一つだけ目を引くものがあるとすれば、第1期との比較による第2期の特徴語を示した表9 (a) ということになるだろう。というのも、この表の第1位と第2位

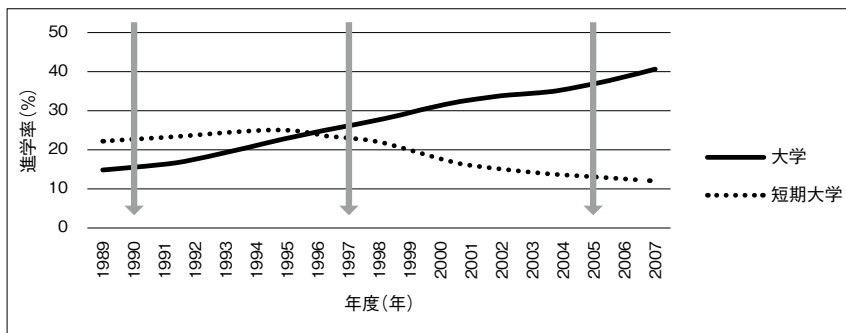


図2 女子高校生の短大・大学進学率推移  
(内閣府男女共同参画局(2016)を元に作成)

がいずれも女性を示す代名詞である“her”および“she”によって占められていることは、平成元(1989)年度告示版の学習指導要領を参照して作成されたであろうセンター試験群に、女性登場人物数(あるいはその言及数)の増加が見られるという可能性を示唆しているからである。

この解釈に関連する外的要因として無視できないのが、日本における女子の大学進学率の上昇である。図2は、1989年から2007年までの期間における女子高校生の短期大学(以下、短大)・大学進学率の推移を示したものである。センター試験が初めて実施された1990年度の女子の短大進学率は22.1%、大学進学率は14.7%であったが、1996年度を境に女子の大学進学率が短大進学率を上回って以降、前者の数値は伸び続けていく(内閣府男女共同参画局, 2016)。そして、女子高等教育の場としての大学がいよいよ重要度を増していくこの時期は、正に平成元(1989)年度告示版の学習指導要領に準拠したセンター試験が実施されていた時期(第2期、1997～2005年)と一致するのである。短大がセンター試験を入試に導入するようになったのは2004年のこと(大学入試センター, 2016, p. 29)で、それまではセンター試験と言えば専ら大学受験予定者を念頭に置いた試験であったのだから、第2期に女子の大学進学率が上昇したということは、取りも直さず同時期における女子のセンター試験受験者の割合増加に直結した可能



表10 短大における女子学生の占める比率推移

年度	2004 (H16)	2005 (H17)	2006 (H18)	2007 (H19)	2008 (H20)	2009 (H21)	2010 (H22)	2011 (H23)
比率 (%)	87.5	87.1	87.6	88.3	88.9	89.1	88.7	88.4
年度	2012 (H24)	2013 (H25)	2014 (H26)	2015 (H27)	2016 (H28)			
比率 (%)	88.4	88.4	88.4	88.5	88.7			

注 文部科学省「学校基本調査」(2009以前), (2010), (2016)を元に作成

性が高い。とすれば、女子受験者の比率増大に鑑み、センター試験の問題が女性登場人物数（あるいはその言及数）の増加という形で呼応したという仮説も、あながち信憑性に欠けるものとは言えないのである（余談ながら、センター試験における女子受験者の存在感そのものは、短大による利用が開始された2004年以降、ますます大きくなっていると思われる。表10からもわかるように、短大在大学生はその9割近くを女子が占めるのである。ただし、表9 (b) ~ (f) を見る限り、その影響がセンター試験の英語問題に直接的な形で表れているかどうかは、疑問符の付くところである）。

#### 4. 仮説と展開

以上に記述したセンター試験英語問題のコーパス分析から形成することのできる仮説は、次の三つに収斂すると言えよう。

- (1) センター試験英語問題は、18歳前後の受験者のほとんどが経験していると想定できるような、直接的かつ友好的なやり取りや出来事を描き込む割合が高い。
- (2) センター試験英語問題では、日本に関する話題への偏向が、開始年度から現在にかけて徐々に強まっている。
- (3) 女子受験者の比率増大に呼応してか、センター試験英語問題において女性登場人物数（あるいはその言及数）が増加した時期がある。

無論、現状では本論文の観察を確認するに足る根拠が出揃っているわけではないが、もしこれらの仮説がそれなりの妥当性を備えているとすれば、

センター試験が提示する英語像は、捨て置くにはあまりに興味深い歪みを伴ったものであると言わざるを得ない。例えば、第一の仮説が正しいとすれば、センター試験を一つの目標として学習してきた多くの日本の学習者には、英語という言語が極めて人畜無害なコミュニケーションから成るユートピア的な世界の構成要素に見えていたとしても不思議ではない。また、第二の仮説が正しいとすれば、英語というのは自己を語り表現する手段の一つであって、他者の理解という負担を被ることなく手に入るありがたい道具なのだという都合の良い「グローバル化」観を、センター試験が体現してきたことにもなり得る。それでも、もし第三の仮説が正しいとすれば、センター試験が描く世界は全く現実を無視したものというわけではなく、良きにつけ悪きにつけ時代ごとの思潮や趨勢を映し出している面もあるのだから、テキスト内外の要因を注視しながら、その曇り具合や屈折率を見極めねばならないはずなのである。

廃止が決定している試験に葬送曲は要らない、という批判には、あらかじめ次のように応えておくことにしよう。その努力を怠ってきたからこそ、英語をめぐる日本の狂騒曲は、鳴り止まないものである。

## 注

- \* 本論文は、日本国際教養学会第7回全国大会における口頭発表「大学入試センター試験が映し出す英語—電子コーパスとして読む英語問題」、イギリス応用言語学会第51回大会（BAAL 2018）におけるポスター発表“Reading English language tests as electronic corpora: A critical/historical enquiry into the Japanese National Center Test”の内容に大幅な加筆・修正を施したものである。
1. 分析に当たっては大文字と小文字の区別を行わないという選択を行った。同じ綴りの単語が文頭と文中で異なる扱いを受けるのを避けるためである。
  2. 第1位の“b”は、下記のような対話において、話し手の片方を特定する記号として登場している例がほとんどであった。  
A: Helen, I heard you are writing a new book.  
B: Well, actually, I'm only thinking of writing one. (1990年度第1問)
  3. 第35位の“swip”は、2011年度第4問に登場する架空の新聞 *Student Weekly*

*International Perspective* の略称である。

4. センター試験を参照コーパスとした英検2級問題の特徴語については、表2と大きく異なる結果が見られなかったため、ここでは割愛する。
5. 掲載をキーワード指数上位12語に限ったのは、どのかけ合わせパターンでも13位以降でのキーワード指数がほぼ100または100未満となり、特徴語としての顕著さを備えているとは言いがたいからである。

#### 引用文献

- Anthony, L. (2016). AntConc (Version 3.4.4) [Computer Software]. Tokyo, Japan: Waseda University. Retrieved from <http://www.laurenceanthony.net/software>.
- Hunter, D. & Smith, R. (2012). Unpackaging the past: ‘CLT’ through *ELTJ* keywords. *ELT Journal*, 66(4), 430-439.
- Jockers, M. L. (2013). *Macroanalysis: Digital methods and literary history*. Urbana: University of Illinois Press.
- Moretti, F. (2013). *Distant reading*. London: Verso.
- O’Halloran, A. (2017). *Posthumanism and deconstructing arguments: Corpora and digitally-driven critical analysis*. London: Routledge.
- Ramsay, S. (2011). *Reading machines: Toward an algorithmic criticism*. Urbana: University of Illinois Press.
- 伊村元道 (2003). 『日本の英語教育200年』東京: 大修館書店.
- 江利川春雄 (2011). 『受験英語と日本人—入試問題と参考書からみる英語学習史』東京: 研究社.
- 旺文社 (編) (2017). 『2017年度版 英検2級 過去6回全問題集』東京: 旺文社.
- 斎藤兆史 (2007). 『日本人と英語—もうひとつの英語百年史』東京: 研究社.
- 大学入試センター (2016). 『独立行政法人大学入試センター要覧』東京: 大学入試センター.
- 土屋結城・榎村真由・北和丈・伊澤高志・瀧口美佳・佐藤繭香 (2018). 「コミュニケーション、経験、言葉—文学テキストとして読む大学入試センター試験」『実践英文学』第70号, 21-38.
- 内閣府男女共同参画局. (2016) 「教育・研究における男女共同参画」『男女共同参画白書 平成27年版』 Retrieved from [http://www.gender.go.jp/about\\_danjo/whitepaper/h27/zentai/html/honpen/b1\\_s06\\_01.html](http://www.gender.go.jp/about_danjo/whitepaper/h27/zentai/html/honpen/b1_s06_01.html) on October 2, 2018.
- 日本英語検定協会 (2018). 「2級の試験内容・過去問」 Retrieved from [http://www.eiken.or.jp/eiken/exam/grade\\_2/](http://www.eiken.or.jp/eiken/exam/grade_2/) on September 24, 2018.

文部科学省（2009以前）.「平成18年度学校基本調査 調査結果の概要（高等教育機関） 学校調査 2短期大学」 Retrieved from [http://www.mext.go.jp/b\\_menu/toukei/001/06121219/002/001/002.htm](http://www.mext.go.jp/b_menu/toukei/001/06121219/002/001/002.htm) on October 6, 2018.

文部科学省（2010）.「学校基本調査—平成22年度（確定値）結果の概要（高等教育機関）」 Retrieved from [http://www.mext.go.jp/b\\_menu/toukei/chousa01/kihon/kekka/k\\_detail/\\_icsFiles/afieldfile/2010/12/21/1300352\\_2.pdf](http://www.mext.go.jp/b_menu/toukei/chousa01/kihon/kekka/k_detail/_icsFiles/afieldfile/2010/12/21/1300352_2.pdf) on October 6, 2018.

文部科学省（2016）.「学校基本調査—平成28年度結果の概要」 Retrieved from [http://www.mext.go.jp/b\\_menu/toukei/chousa01/kihon/kekka/k\\_detail/1375036.htm](http://www.mext.go.jp/b_menu/toukei/chousa01/kihon/kekka/k_detail/1375036.htm) on October 6, 2018.