

大学入試センター試験 32 年のコーパス分析による 通時的変化と近隣のアジア諸国との共時的比較研究

柏木 哲也

1. はじめに

本年度から大学入試共通テストが開始された。昨年度まで行われていた大学入試センター試験（以後センター試験）を受け継ぐ国家規模での公的な試験である。そこで本研究では、センター試験の語彙、文法特性を通時的にコーパス分析し、どのような変化が見られたのかを探索的に分析することを目的とし、大学入試共通テスト（以後共通テスト）とセンター試験の相違点を踏まえ今後のありかたを議論する。サブコーパスとして、センター試験における長文読解問題を3つの期間に分け①1988年から1999年まで、②2000年から2009年まで、③2010年から2020年まで、及び④2021年度の共通テスト、⑤韓国のCSAT（College Scholastic Ability Test）、⑥台湾の学科能力測驗（T1）、⑦台湾の指定科目考試（T2）、⑧TOEFL用市販問題集から合計8つを作成した。語彙関係の調査項目は、標準化されたタイプ・トークン比、平均文長、平均語長、平均語彙難度（JACET 8000基準）、リーダビリティ値であり、文法関係は、CLAWS（ランカスター大）のParsingによるものを比較した。結果として、①から④までは語彙関係の調査項目のほぼ全てに経年ごとの数値の上昇が見られ、英文が難化していることが判明した。平均文長の増加は、文の複雑さの上昇を意味し、平均単語長の上昇は、単語難易度の上昇を意味する。リーダビリティ値の上昇も英文全体の難易度が上昇していることを示唆する。また、①と③を比較した文法項目の比較調査において、英文の質が変化している兆候が多く発見され、柏木（2019）に類似した結果が示唆されていることが確認された。最後に、一年分ではあるが共通テストの調査結果と照らし合わせ、今後の入試問題がどのようなべきかを議論する。

2. 分析方法

先に述べたようにセンター試験における長文読解問題を① 1988年から1999年まで② 2000年から2009年まで③ 2010年から2020年までの3つの期間に分け、サブコーパスを作成した。また④ 2021年度の共通テストの設問から4つ目のサブコーパスを、更に国外の同類の試験として⑤韓国の大学修学能力試験（CSAT）⑥台湾の学科能力測驗（T1）、⑦台湾の指定科目考試（T2）、⑧ TOEFLの市販対策問題集9冊（TFL）もサブコーパス化した。台湾の「学科能力測驗」は高1と高2の学習内容に限定され、大学での授業を受けるための基本的な知識を測るために行われる。「指定科目考試」は高1から高3までの学習内容が含まれ、教科知識と大学が要求する教科知識、応用能力を測るために行われる。TOEFLを対照コーパスに選んだ理由は、近年文部科学省が大学入試問題の代替項目の案として提示したためである。また、⑤、⑥、⑦については最近20年間実施されたものからランダムにサンプリングを行ってサブコーパス化したため採用期間については若干のズレが生じている。以上合計8つのサブコーパス間の比較を行い、語彙関係の調査項目は、標準化されたタイプ・トークン比、平均文長、平均語長、平均語彙難度（JACET 8000基準）、リーダビリティ値であり、文法関係は Parsing によるサブコーパス①と③の比較を行い20世紀と21世紀を跨いでの変化を探った。特に共時的分析においては、各コーパス間で対数尤度検定を用いて出現頻度の異なる名詞を抽出し、問題で扱う題材にどのような傾向が見られるのかを概略調査すると同時にランダムサンプリングから入試問題における代表的な英文に見られる特徴を探った。

コーパス化の際、文法問題、語法問題、会話文問題、内容真偽問題、本文の書き換え、設問文などを除外し、長文読解問題の本文のみをコーパス化し、脚注のついた語も含めた。但し会話文問題でも日常レベルのやり取りではなく説明調の文脈を持つと判断したものについては、コーパスに組み入れた。近年のセンター試験においては、以前のような長文然とした形式ばかりではなく対話形式の中にかなり文語的で内容の複雑な英文を含めることが増えているためである。

また語彙難度やリーダビリティを調べるために、オンライン測定を染谷先生(青山学院大学)のページで行った。(http://someya-net.com/wlc/index_J.html)

3.1 統計値

表 1. コーパスデータ

項目\サブコーパス	① 1988- 1999本試	② 2000- 2009本試	③ 2010- 2020本試	④ 2021共 通テ問4	⑤ CSAT	⑥ T1	⑦ T2	⑧ TFL
総語数	121971	99795	179641	9903	53361	18663	20444	83300
タイプ (異なり語数)	22232	17815	31523	612	7252	3799	4407	10595
異なり語率	13.3406734	15.58196	13.15772	36.88969	13.59	20.92191	22.15686	12.87269
標準化された異なり語 率 (Standardized TTR)	38.333321	38.71765	40.65667	42.9	47.528	45.55556	47.08947	47.21706
標準偏差	57.5765305	57.18053	56.56506		50.268	50.25329	48.93696	51.74743
ベースス	1000	1000	1000	1000	1000	1000	1000	1000
平均語長	4.23993015	4.373221	4.530896	4.618445	4.662	4.53062	4.715384	5.014908
標準偏差	2.29555798	2.450778	2.564175	2.637243	2.555	2.404862	2.568209	2.71432
平均文長	15.5145893	16.62596	18.07106	21.82895	19.188	17.13019	17.43208	20.58679
標準偏差	27.6734982	29.82576	31.02495	31.86208	8.451986	8.879745	9.210598	10.77605

表 1 は、サブコーパス 8 つの語彙とテキストの複雑性を表したワードリストである。表中のベースス (basis) とは、コーパス分析をするさいの語彙計算においてある程度の「まとまり」作ることにより、繰り返し計算による語彙使用の偏りを補正するためのものである。一般的には各設問の語数を大きく下回らない範囲で設定を行う。

さて共時的な比較対照を行ってみよう。標準化された異なり語率 (Standardized Type-token Ratio : 以後 S.TTR) においては、⑤韓国の CSAT、⑦台湾の指定科目考試 (T2) 更に TFL においてもほぼ同じような値 (47 点台) が出ている一方で日本のセンター試験は、上昇しているとはいえ 40 点台に止まっており、高校 2 年生までを対照にした台湾の「学科能力測驗」の 45 点にも届いていない。その意味ではかなり基礎的な能力を測る目的で作成された試験と言うことができよう。通常異なり語率は英文の複雑さや難解さを表す尺度と考えられているが、入試問題分析の場合は、一題の長文の長さも考慮すべき要因と考えられる。すなわち中文 (300 語程度) が多い構成の場合は、1000 語の中の名詞や形容詞には異なる話題を扱った語が多く占める可能性が高く、異なり語率も上昇しがちになるため各問題の量の違いも S.TTR に関係する。

もう一つの指標である平均文長に着目してみよう。こちらも①から③まで増加の一途をたどり韓国や台湾の値とほぼ同じかやや多い値になっている。入試問題の場合は T-Unit を出さずとも英文の長さが文構造の複雑さに結び付くことが多く問題難度に関係する場合が多い。

更に平均語長にも目を向けてみる。こちらも平均文長と同様に右肩上がりになっ

ているが、それでも韓国、台湾のレベルには達していない。ここからも⑥と同じく基礎力の判定を目的とした問題であると推察できる。

次の2表では、別のオンライン解析により JACET8000 による単語難度とリーダビリティを示したものである。

表 2. 語彙特性とリーダビリティ

	① 1988- 1999本試	② 2000- 2009本試	③ 2010- 2020本試	④ 2021共 通テスト	⑤ CSAT	⑥ T1	⑦ T2	⑧ TFL
総語数	20915	15731	28998	1519	52839	17622	19267	81853
総文数	1413	997	1789	95	2794	1076	1131	4028
平均1文語数	14.5	15.8	16.5	16.1	18.7	16.5	17.2	20.2
平均語長	4.6	4.7	4.8	5	4.8	4.7	4.9	5.1
平均語彙レベル	1.14812	1.2031	1.21567	1.31402	1.31692	1.25718	1.32724	1.41002
分散	0.7597	0.95188	0.91741	1.45965	1.28516	1.19671	1.43218	1.77438
標準偏差	0.87161	0.97565	0.95782	1.20816	1.13365	1.09394	1.19674	1.33206
リーダビリティ値								
Automated Readability Index (ARI)	7.5	8.6	9.4	10.2	10.5	9	10.2	12.7
Coleman-Liau Index (CLI)	9.2	10	10.7	11.8	10.9	10.1	11.3	12.8

読みやすさの程度を数値で表したものがリーダビリティスコアである。本稿では、ARI と CLI を採用している。計算方法は以下の計算式による。

$$\text{ARI} : 4.71 \times (\text{文字数} / \text{単語数}) + (0.5 \times 1 \text{ センテンスあたりの平均単語数}) - 21.43$$

$$\text{CLI} : (0.0588 \times 100 \text{ 単語あたりの平均文字数}) - (0.296 \times 100 \text{ 単語あたりの平均センテンス数}) - 15.8$$

基本的に ARI は、1 語の長さと同語数が考慮され、CLI は、1 語の長さと同平均文長が主な影響要因になっている。またリーダビリティ値は、母語話者の学習学年に相当すると考えられていることにより③においては、小学校 9.4 年生即ち母語話者の中学 3 年に相当する文章であり、⑤、⑦の韓国、台湾の値はそれより更に 1 点（1 歳相当）ほど高い。即ち高校一年生レベルの英文であると考えられる。

平均語彙レベルは、0.1 上がるだけで相当イメージとしては異なるという感覚であるので日本のセンター試験の 2000 年代の語彙レベルの上昇はかなり大きな変化と受け止められる。

表 3.1 でも部分的に見られたが、①から③までの通時的変化において平均文長、平均語長、平均語彙レベル、リーダビリティ値のいずれも経年と共に上昇している点から作成者側の問題の難度の上昇を図った意図が見てとれる。更に④の共通テストの 4 番の問題においてもそれと共通した傾向が観察できる。

3.2 語彙レベル別占有率

表 3、表 4 は JACET8000 を基にした語彙レベル別の占有率をサブコーパス別に表している。まず①～④においては語彙レベル 1,000 の占める割合は年々下降し、その代わり語彙レベル 2,000 から 4,000 の占める割合が上昇している傾向は顕著である。③と韓国や台湾と比較してみると、レベル 1,000 の占有率は最も高く、逆に基礎的英語力の要と言われるレベル 2,000~4,000 の合計値は最も低くなっている。高校 2 年生までを対照にした台湾の「学科能力測驗」(⑥) とほぼ同じ値である点から同じような難易度であることが伺える。

表 3. 語彙レベル (1)

語彙レベル	① 1988-1999		② 2000-2009		③ 2010-2020		④ 2021	
	頻度	%	頻度	%	頻度	%	頻度	%
1,000	16917	80.88	13046	78.74	22587	77.89	1103	72.61
2,000	1326	6.34	1150	6.94	2368	8.17	134	8.82
3,000	574	2.74	470	2.84	919	3.17	47	3.09
4,000	127	0.61	164	0.99	392	1.35	32	2.11
5,000	136	0.65	138	0.83	178	0.61	26	1.71
6,000	87	0.42	60	0.36	216	0.74	9	0.59
7,000	84	0.4	41	0.25	102	0.35	12	0.79
8,000	53	0.25	49	0.3	88	0.3	11	0.72
その他	1611	7.7	1451	8.76	2148	7.41	145	9.55
合計	20915	100	16569	100	28998	100	52839	100

表 4. 語彙レベル (2)

語彙レベル	⑤ CSAT		⑥ T1		⑦ T2		⑧ TFL	
	頻度	%	頻度	%	頻度	%	頻度	%
1,000	40658	76.95	13431	76.22	14075	73.05	57955	70.8
2,000	4144	7.84	1318	7.48	1550	8.04	6741	8.24
3,000	1807	3.42	598	3.39	729	3.78	3431	4.19
4,000	1091	2.06	279	1.58	481	2.5	2490	3.04
5,000	704	1.33	173	0.98	268	1.39	1352	1.65
6,000	451	0.85	129	0.73	180	0.93	946	1.16
7,000	324	0.61	126	0.72	158	0.82	832	1.02
8,000	295	0.56	82	0.47	95	0.49	683	0.83
その他	3365	6.37	1486	8.43	1731	8.98	7423	9.07
合計	17622	100	19267	100	81853	100	81853	100

3.3 本試験と追試験

次に本試験と追試験を比較してみよう。左側3列が本試験、4列目から6列目が追試験のコーパス統計値である。③の本試以外はS.TTRにおいても平均文長においても全て追試験の数値が上回っている。英文の質と複雑さにおいては追試験の方が難しいと言えるかもしれない。しかも着目すべきは、本試においても追試においても実施時期が新しくなるほどS.TTRも平均文長も数値が上昇している点である。これは経年と共に英文の質が上がっており、読みにくい英文に変化していることを示唆している。

表5 本試と追試の統計値

項目\コーパス	① 1988-1999本試	② 2000-2009本試	③ 2010-2020本試	1988-1999追試	2000-2009追試	2010-2020追試
総語数	121971	99795	179641	121028	94456	180926
タイプ(異なり語数)	22232	17815	31523	21892	16785	32364
異なり語率	13.34067	15.58196	13.15772	13.90362	16.59776	12.99798
標準化された異なり語率(S.TTR)	38.33333	38.71765	40.65667	39.2	39.475	39.99677
標準偏差	57.57653	57.18053	56.56506	55.65113	55.97912	57.39393
ベース	1000	1000	1000	1000	1000	1000
平均語長	4.23993	4.373221	4.530896	4.276723	4.393057	4.466245
標準偏差	2.295558	2.450778	2.564175	2.319953	2.441283	2.525188
平均文長	15.51459	16.62596	18.07106	15.02305	17.36853	18.45438
標準偏差	27.6735	29.82576	31.02495	25.90461	30.29115	31.86208

3.4 文法解析による頻度比の増減

以下に上げる表6、7、8は、センター試験最初の11年間(①)と最後の11年間(③)の文法パーザーによる出現頻度率の20%以上異なる項目を、頻度率の低い方から減少と増加に分けて提示したものである。

表 6 出現頻度比の低い項目 (1)

10-20	[10/88 比率]	
文法項目	出現頻度比	代表的事例
RPK	20	about
NP1@	33.33	Nominal noun = 特定事象の回避
PN	33.33	everyone, someone = 代名詞の減少
PPX221	33.33	each
PPX222	33.33	other
VHD	33.33	had (完了、had to, 過去均等)
VVGK	33.33	be going to
PPHS1	36.41	he, she = 3 人称単数代名詞の減少
VDD	38.89	did (人、もの半々)
PPHO1	39.02	her, him = 3 人称単数代名詞の減少
UH	40	yeah, oh, hi, no, huh, mmm, wow
PPX1	41.67	itself, herself, himself, myself..
VVD@	42.86	過去時制動詞
VBDZ	48.7	was = 過去形、単数の減少
CS21	50	as,
CS22	50	if, though, that
CSW	50	if, whether
NN1%	50	gerund
VHN@	50	had
CSN	52.94	than = 比較級減少
VVD	55.23	過去時制動詞
DAR	56	more, less (ad, a) = 限定詞減少
RL	56	together, away, home, aside, there, here, abroad, downstairs, above, apart = 副詞全般の減少

まず減少項目の顕著な点として、代名詞全般、過去時制動詞、副詞全般、固有名詞、更に限定詞 (determiner) の減少である。これらは、人と過去を中心とした物語文から現在を中心とした説明文へ、修辭的表現から客觀的事象を述べる經驗主義への転換を示唆している。

表 7 出現頻度比の低い項目 (2)

10-20	[10/88 比率]	
文法項目	出現頻度比	代表的事例
VDI	58.33	do (不定詞、助動詞と共起)
VHI	60	have (助動詞、本動詞半々)
VVN@	60	過去分詞 (完了、受動)
DA1	62.5	much, little (ad, a) = 限定詞減少
RGR	65	less, more = 比較級の減少
VBN	65.22	been (完了受動態)
PPH1	66.07	
PPX2	66.67	themselves, ourselves
VDN	66.67	done (完了、受動)
VHG	66.67	having
GE	68.25	apostrophy s
CS@	71.43	so, once, 副詞 (句) 減少
RRR	71.43	more, better, later, sooner, longer,
VBM	71.43	am
VD0	72	do (現在形、疑問、否定)
NP1	72.8	Nominal noun = 特定事象の回避
XX	73.12	not
EX	73.91	there (null s) = 日本的「ある」の減少
RT	74.42	today, then, now, again, most,
NN1@	75	gerund
RA	75	ago, BC
RP	75	up, out, back, in, down, off, along = 副詞全般の減少
PPHS2	76.39	they
RG	76.92	too, about, around, so, more, how, = 副詞全般の減少
JK	77.78	be able to
RRQ	78.72	why, how, when, where
IW	79.1	with (prep) = 孤立化
TO	82.1	不定詞=人間、推量の減少

更に減少幅は小さくなるが、there is / are 構文と言われる存在表現、動名詞、am は日本の表現からの直訳によく見られるが、これらの減少は英語的視点や表現の反映が増加していることの裏返しと考えられる。同時に限定詞や比較級の減少はいわゆる修辭的熟語表現 (no less than, no more than など) の減少を意味し、より直接的でわかりやすい文法体系への変換と思われる。

今度は逆に出現頻度比が増加している項目を見てみよう。表 8 は、①:③の出現頻度比が 2 割以上上昇している項目である。

表 8 出現頻度比の高い項目

10-20	[10/88 比率]	
文法項目	出現頻度比	代表的事例
PPY	121.21	you = 会話文、一般人称代名詞
JJ	122.98	adj = 母語話者の描写
DB	123.53	all, both, half, (ad, a, pron) = 複数名詞による一般化
MC2	125	30s, 50s, 60s, 70s, = 実証的
VVG	127.61	分詞、動名詞
IO	129.03	of (prep)
MC	130.85	number (question & adj) = データ重視
VH0	137.93	have (完了高率)
VBR	139.58	are = 複数、一般化の増加
VV0@	140	命令?
VV0	140.87	他動詞全般
IF	141.54	for (prep)
MC1	150	one
PPIO2	150	us= 目的格 会話文= フォームとレトリックの競合
VDZ	150	does (否定文率高)
NN2	153.14	noun (plurals) = 複数名詞による一般化
VVZ	168.63	3 人称単数動詞
VBZ	172.46	is (無生物主語)
RGT	175	most, least = 最上級の増加
DD2	185.71	these, those, (a, pron) = 複数名詞による一般化
NN2%	200	remains (noun, plurals)
JJT	212.5	superlative (a) = 最上級増加
NNU	241.67	percent, miles, meters = データ重視
VBG	250	being (進行受動態、動名詞) = 不定詞の減少と対照
MCMC	300	---
REX21	300	for example, that is = 言い換えロジカルコネクター
REX22	300	for example, that is = 言い換えロジカルコネクター
RP@	300	in, under, on,
RRT	300	best, lowest, most,
VVZ@	400	3 人称単数動詞
ZZ1	430.77	alphabet**
NPM1	1100	月名 (March, June...) = 時間強調

減少傾向のある単語として限定詞を挙げたが、all, both, half という複数を扱う限定詞は例外的に増加している。これは個別の事象から、一般化を図ろうというロジックに則った語使用であると推察できる。同様の傾向が、VBR (are) や DD2 (these, those) に見て取れる。また英語ロジックを示す例として、最上級を表す RGT, JJT, RRT が高頻度比を示している点が注目される。最上級表現は、特に「最も～な」という日本語的発想が数値的尺度に頼り、日本人の産出する英語において事物を表

現する時に、感覚的な形容詞に余り最上級を使用しない傾向が強いように思える。日本語では余り見られないような one of the best... という形式も含み、英語的レトリックと言えよう。

また場所 (where) を中心とする日本語発想に対して、時間 (when) を前面に押し出す英語発想を表す月名や MC2、更にデータを重視する MC (数値) も多用されている。同時に読み手を示唆する “you”, 他動詞や前置詞の目的語としての “us” は、日本語発想とは異なる英語的視点を反映した語使用と言えらるう。

3.5 対数尤度検定

Wordsmith の log likelihood の対数を取り、2つのサブコーパス間で有意な差がでている語を調べ、それによりどのような隠れた傾向が見てとれるのかを探索してみた。その結果、①や③と韓国、台湾のサブコーパス間に余り特徴だった語彙特性は見られなかったが、①と③の通時的比較のサブコーパス間に以下のような語が観察された。

表 9 ①と③の間の差のある語

88-20		20-88		
	Key word	Keyness	Key word	Keyness
1	HE	194.8863	STUDENTS	58.99905
2	ROBERT	63.6033	IS	49.0678
3	WAS	62.83901	ORANGES	38.4448
4	HIS	62.22935	ROUTES	35.23977
5	MONKEY	58.30041	VENDING	33.10321
6	CORRIGAN	56.53288	MUSIC	31.13801
7	HAD	56.04497	HEALTH	29.89857
8	AUTHOR	55.45898	CAN	28.55973
9	PICTURE	52.05232	FOLLOWING	27.83908
10	HOUSE	46.11662	CASTLE	27.76227
11	BATS	44.16169	PASSAGE	25.62608
12	THE	42.57632	AREAS	25.62608
13	WOMEN	42.44289	FRIENDSHIPS	24.55803
14	MARGARET	42.39459	ADULTS	23.96283
15	MIMI	38.44116		
16	HIM	35.97478		
17	BORDES	35.32671		
18	FISHING	33.55988		
19	DR	33.55988		
20	MEN	33.56		

88-20 列は、サブコーパスの①から③を見た場合の頻度率の統計的に有意に高い語の 20 語であり、20-88 列は逆から見た場合の 14 語である。(キーネスの最低ライン = 0.23) 左列に多い語は、物語文で身近の描写に使用される平易な語であり、3 人称単数の代名詞が目立つ。一方、右列にはやや抽象的な名詞、しかも複数形が多く、一般化された事象を扱った論説や説明文のジャンルが予想される。

4. 入試問題例

ここでは、各国の最近の入試問題の例をいくつか挙げ、その特徴を概観してみたい。

4.1 日本

Ice hockey is a team sport enjoyed by a wide variety of people around the world. The object of the sport is to move a hard rubber disk called a “puck” into the other team’s net with a hockey stick. Two teams with six players on each team engage in this fast-paced sport on a hard and slippery ice rink. Players may reach a speed of 30 kilometers per hour sending the puck into the air. At this pace, both the players and the puck can be a cause of serious danger.

The speed of the sport and the slippery surface of the ice rink make it easy for players to fall down or bump into each other resulting in a variety of injuries. In an attempt to protect players, equipment such as helmets, gloves, and pads for the shoulders, elbows, and legs, has been introduced over the years. Despite these efforts, ice hockey has a high rate of concussions.

A concussion is an injury to the brain that affects the way it functions; it is caused by either direct or indirect impact to the head, face, neck, or elsewhere and can sometimes cause temporary loss of consciousness. In less serious cases, for a short time, players may be unable to walk straight or see clearly, or they may experience ringing in the ears. Some believe they just have a slight headache and do not realize they have injured their brains.

In addition to not realizing the seriousness of the injury, players tend to worry about what their coach will think. In the past, coaches preferred tough players who played in spite of the pain. In other words, while it would seem logical for an injured player to stop playing

after getting hurt, many did not. Recently, however, it has been found that concussions can have serious effects that last a lifetime. People with a history of concussion may have trouble concentrating or sleeping. Moreover, they may suffer from psychological problems such as depression and mood changes. In some cases, players may develop smell and taste disorders.

The National Hockey League (NHL), consisting of teams in Canada and the United States, has been making stricter rules and guidelines to deal with concussions. For example, in 2001, the NHL introduced the wearing of visors — pieces of clear plastic attached to the helmet that protect the face. At first, it was optional and many players chose not to wear them. Since 2013, however, it has been required. In addition, in 2004, the NHL began to give more severe penalties, such as suspensions and fines, to players who hit another player in the head deliberately.

The NHL also introduced a concussion spotters system in 2015. In this system, NHL officials with access to live streaming and video replay watch for visible indications of concussion during each game. At first, two concussion spotters, who had no medical training, monitored the game in the arena. The following year, one to four concussion spotters with medical training were added. They monitored each game from the League's head office in New York. If a spotter thinks that a player has suffered a concussion, the player is removed from the game and is taken to a “quiet room” for an examination by a medical doctor. The player is not allowed to return to the game until the doctor gives permission.

The NHL has made much progress in making ice hockey a safer sport. As more is learned about the causes and effects of concussions, the NHL will surely take further measures to ensure player safety. Better safety might lead to an increase in the number of ice hockey players and fans.

これは 2021 年の大学入学共通テストの問題のひとつである。614 語という長さがあり、限られた語彙レベルの語を使いスポーツ競技の際に起こり得る身体的ダメージを説明した文章である。

動詞句、副詞、不定詞など高校の既習事項をバランスよく配置したよく練られた英文と考えられる。語彙レベルもレベル 5 以上の語はほとんどなく、限られた範囲のレベルから出題されている。ただ限定詞、不定代名詞が少なく、わかりやすい反

面、結束性の理解や主張と具体例の判別などは読解できているのかが検査しにくいと思われる。

4.2 韓国

We might describe science that has no known practical value as basic science or basic research. Our exploration of worlds such as Jupiter would be called basic science, and it is easy to argue that basic science is not worth the effort and expense because it has no known practical use. Of course, the problem is that we have no way of knowing what knowledge will be of use until we acquire that knowledge. In the middle of the 19th century, Queen Victoria is supposed to have asked physicist Michael Faraday what good his experiments with electricity and magnetism were. He answered, “Madam, what good is a baby?” Of course, Faraday’s experiments were the beginning of the electronic age. Many of the practical uses of scientific knowledge that fill our world—transistors, vaccines, plastics—began as basic research. Basic scientific research provides the raw materials that technology and engineering use to solve problems.

Basic scientific research has yet one more important use that is so valuable it seems an insult to refer to it as merely functional. Science is the study of nature, and as we learn more about how nature works, we learn more about what our existence in this universe means for us. The seemingly knowledge we gain from space probes to other worlds tells us about our planet and our own role in the scheme of-nature. Science tells us where we are and what we are, and that knowledge is beyond value.

2017 年に行われた College Scholastic Ability Test (大学修学能力試験) (⑤) より引用した。英語による内容選択がこの後に問われ、小問を含めた合計総語数は 3500~3700 語である。代名詞が少なく、語彙難度が高い。70 分で 45 問。内 17 問はリスニングテスト。相当な速読力が要求される問題である。

4.3 台湾

Imagine two bottlenose dolphins swimming in the Gulf of Mexico. You hear a series of clicks, whistles, and whines coming from each, much like a conversation. We can’t be sure what they are discussing, but scientists do believe dolphins call each other by “name.”

A recent study suggests the marine mammals not only produce their own unique

“signature whistles,” but they also recognize and mimic whistles of other dolphins they are close to and want to see again. It seems that dolphins can call those they know by mimicking their distinct whistles. “They’re abstract names,” said Randall Wells, one of the authors of the study.

To conduct the study, the researchers listened to recordings of about 250 wild bottlenose dolphins made around Florida’s Sarasota Bay from 1984 to 2009, and four captive dolphins at a nearby aquarium.

Some wild dolphins were briefly captured and held in separate nets by the research team, allowing them to hear but not see each other. Researchers found that dolphins familiar with each other would mimic the whistle of another in that group when they were separated. Most of this took place among mothers and calves, or among males who were close associates, suggesting it was affiliative and not aggressive—somewhat like calling out the name of a missing child or friend. Whistle copying of this sort was not found in dolphins that happened to cross paths in the wild.

This use of vocal copying is similar to its use in human language, where the maintenance of social bonds appears to be more important than the immediate defense of resources. This helps differentiate dolphins’ vocal learning from that of birds, which tend to address one another in a more “aggressive context.”

If confirmed, this would be a level of communication rarely found in nature. If dolphins can identify themselves and address friends with just a few squeaks, it’s easy to imagine what else they’re saying. However, as the authors of the study point out, all we can do right now is still imagine.

指定科目考試（⑥）から引用した。80分で約3700語、51問であり総語数は日本のセンター試験や韓国と大差はない。会話、語彙選択、内容選択問題も含まれる。主語が短く読みやすい反面、修飾部分は長く語彙難度も高い。内容も多岐に渡る。韓国と同じで200～300語の異なる分野の問題を多数出題するため、異なり語率が高い。

5. 考察

コーパス分析による語彙と文法調査項目では柏木（2018）、柏木（2019）と類似

した変化が見られ、平均語長、平均文長の増加と 限定詞、不定詞、助動詞の減少、及び日本語からの干渉の減少が示唆された。これは物語文的な要素から説明文の要素への転換を含み、人からモノ中心の説明へ客観的データに基づく主張と一般化、説明文形式の長い文章題への変化を意味するものと思われる。更にこの変化の裏には、文学的情緒的な叙述から客観的説明文への変化が透視されると同時に、会話文においても対話から説明へというコミュニケーションの質の変化が見て取れる。

さてここで、英語教育の本質的テストとしての議論の必要性があり資格試験か実力試験かという問題が浮かび上がってくる。異なり語彙率 (Type token ratio)、リーダビリティ、平均語彙難度は、韓国の CSAT や台湾の指定科目考試に比較すると明らかに低く、どちらかと言うと高校 2 年までの実力を試す台湾の学科能力測驗の性格を帯びていると言えよう。

問題の質として、2000 年以降の問題に対話形式であるにも関わらず、口語体と文語体の入り混じった設問が増え、コーパス作成の際に長文問題として分類して良いかどうか迷う問題が増えている点が挙げられる。更に、文を補充させる問題や、段落そのものを補充させる問題など内容に関する設問形式が多様化している上に、2017 年ごろから、出典のついたグラフや表の内容を英語で問う出題も増加している。

2021 年度に始まった大学入学共通テストでは、単問としての発音、アクセント問題及び文法・語法の正しさを問う問題が完全になくなった。当然、スピーキングテストなしでどうやって受験生の音声面でのアウトプット技能を測るのかという疑問が浮かび上がってくる。民間試験利用制度が頓挫した今となつては、「正確な」英語音声の産出はどうでもいいのであろうか。以前は、非文法的な文を正しく直す正誤問題が出題されていた。例えば、no car was not seen (否定語の重複)、two fishes (複数形の誤り)、play piano better (the の欠落)、like him more / better (選択)、how many...? (動詞なしの文)、here is two apple (複数形と動詞の誤用)、there is + 固有名詞など多くの間違えやすい表現形態が問題として出されていた。また発音記号を使用した音の違いを問う問題 (lock, luck, lack; played, wicked, stepped の ed 部分の発音; dogs, cats, horses の s, es 部分の発音)、アクセントの位置を問う問題 (agriculture, antenna, amateur)、語用の正しさを問う問題 (be interested by, people are many)、スペリングの正しさを問う問題 (nonsense, discipline, Wednesday, February, whose) など出題されていた。確かにこういった形式を問う問題が多くなりすぎ

てしまい、意志を発信する力や文意を読み取る力を測る試験問題が余り多くなかった時代もあった。しかしながら、高校入試では問われることはあっても大学入試共通テストでは完全に消えてしまったのである。

再度今回の研究の結果をまとめてみよう。日本の入試問題の20年の縦断的变化として以下の点が挙げられよう。形式的には、構文解析的な倒置、省略、否定辞、代名詞、不定詞、助動詞が多い文から、構造的には平易であるが、高難度語を増やし、文長をのばした(句や節の多い)テキストに変化している。また、ジャンルの的には、抽象的、哲学的、人間の思想や行動中心の文学的な問題から、政治、経済、生活、統計を中心とした経験主義を基準にした内容と、科学、環境、医学、コンピュータを中心としたモノを基軸とした内容への変化が見られる。

横断的比較としては、以下の点が挙げられる。韓国では日常から民族問題が中心だが、否定表現が多い。台湾は読みやすく話題も最も多岐にわたる。両国とも語彙レベルが高く、各長文問題の語彙量は日本に比べ少ないが、試験全体での語彙量は大きな差はない反面、話題の種類が多いので文脈を取るための速読力も要求される。

今後はセンター試験型の問題難度を更に上げることは可能かという問題が浮かび上がってくる。そうなると共通テストの資格試験的な意味合いはますます減り、実力を測る試験に特化することは間違いない。また韓国も台湾もほとんど出題していない発音・アクセント問題、文法・語法問題を再度取り入れるかという点で再度議論が必要になるであろう。日本語は確かに音素が非常に少なく、日本人英語学習者の多くが英語の聞き取りを苦手としていることは否定できない。しかし音素の違いやアクセントの違いを知識として持っていれば、聞き取り練習への内発的動機付けにつながる可能性もある。また、文法・語法問題はスピーキングやライティングを行えばある程度肩代わりとなるだろうが、複雑な構造の英文を読む力は、実際にそのような英文を問題として出題し、内容選択以外の方法でしか試せない。

入試問題を通して何を試すのか。内容を早く解読する速読力なのか、正確に解読する読解力なのか、それとも母語と比較し対照的に捉える能力なのか。出題する側は、大きな岐路に立たされていることは間違いない。しかしながら最後に、正式な(正確な)英語を話せない、書けない学生を増やさないために、英語という言語を対照言語学的に正確に捉え、そのレトリックを味わうために芸術としての英語を学ぶ必要があると提案したい。そのためには、不定詞、限定詞 (more, less, others,

any, either, neither) は必要であるし、人文科学 (文学、哲学、倫理学等) は消えてはならないのである。

参考文献

- 石川 慎一郎. (2005). 「日韓の中高英語教育における目標語彙水準の経年的変化：1994～2005 年度大学入試英語問題コーパスに基づく計量的調査」. 『神戸大学国際コミュニケーションセンター論集』、第 2 号、69-82.
- 柏木 哲也. (2015). 「大学入試英語長文問題の縦断的研究」. 『北九州市立大学基盤教育センター紀要』. 22. 51-68.
- 柏木 哲也 (2017). 「『日本人英語』考察 — その特徴と問題点」『北九州市立大学基盤教育センター紀要』第 25 号, 69-85 頁.
- 柏木 哲也 (2018). 「TOEFL, センター試験, 大学入試英語長文問題の文法・語彙比較研究」 Information Communication Technology Theory & Practice The Final (JACET - ICT 調査研究特別委員会最終報告書) 大学英語教育学会 JACET - ICT 調査研究特別委員会. 1-11 頁.
- 柏木 哲也 (2019). 「日本・韓国・台湾における大学入試英語長文問題の語彙・文法比較研究」 Studies in English Teaching and Learning in East Asia, No. 7, pp. 91-108.
- 文部科学省. (2018). 「各資格・検定試験と CEFR との対照表」 (令和元年 9 月 16 日検索).
http://www.mext.go.jp/b_menu/houdou/30/03/_icsFiles/afiedfile/2019/01/15/1402610_1.pdf
- JC 教育研究所. (2021). 『全国大学入試問題データベース Xam 2018』. (DVD-ROM)、JC 教育研究所.

参考資料

1. TOEFL コーパスを作成した問題集

TOEFLコーパスに使用した文献一覧				
	書名	発行年	出版社	著者
1	TOEFL ITPテストリーディング教本 (TOEFLテスト教本シリーズ)	2013	テイエス企画	山田 広之
2	iBT対応 TOEFLテスト完全攻略リーディング (TOEFLテスト完全攻略シリーズ)	2009	アルク	ワーデン, ポール・ヒルキ, ロバート・霜村 和久
3	TOEFLテストITP(団体受験)リーディング完全攻略	2006	ICU TOEFL問題研究会	田中 真紀子
4	毎日ミニ模試TOEFLテストiBTリーディング	2014	テイエス企画	林 美由紀・ 山田 広之・ 高木義人
5	これならわかるTOEFL ITPテスト本番模試 (TOEFL(R)大戦略)	2012	旺文社	田中 真紀子
6	TOEFL TEST対策iBTリーディング	2006	テイエス企画	宮前 一廣
7	TOEFL Test iBTリーディング 実践編	2010	南雲堂	クヌーセン,ジ ム・生井 健 一
8	TOEFL ITP(R)テスト 公式テスト問題 & 学習ガイド	2012	研究社	田地野 彰・ 金丸 敏幸
9	TOEFLテストリーディング問題 270 4訂版 (TOEFL(R)大戦略)	2014	旺文社	田中 真紀子

2. CLAWS の基本的分類基準

付録 2.1 CLAWS の分類項目 1	
APPGE	possessive pronoun, pre-nominal (e.g. my, your, our)
AT	article (e.g. the, no)
AT1	singular article (e.g. a, an, every)
BCL	before-clause marker (e.g. in order (that), in order (to))
CC	coordinating conjunction (e.g. and, or)
CCB	adversative coordinating conjunction (but)
CS	subordinating conjunction (e.g. if, because, unless, so, for)
CSA	as (as conjunction)
CSN	than (as conjunction)
CST	that (as conjunction)
CSW	whether (as conjunction)
DA	after-determiner or post-determiner capable of pronominal function (e.g. such, former, same)
DA1	singular after-determiner (e.g. little, much)
DA2	plural after-determiner (e.g. few, several, many)
DAR	comparative after-determiner (e.g. more, less, fewer)
DAT	superlative after-determiner (e.g. most, least, fewest)
DB	before determiner or pre-determiner capable of pronominal function (all, half)
DB2	plural before-determiner (both)
DD	determiner (capable of pronominal function) (e.g. any, some)
DD1	singular determiner (e.g. this, that, another)
DD2	plural determiner (these, those)
DDQ	wh-determiner (which, what)
DDQGE	wh-determiner, genitive (whose)
DDQV	wh-ever determiner, (whichever, whatever)
EX	existential there
FO	formula
FU	unclassified word
FW	foreign word
GE	germanic genitive marker - (' or's)
IF	for (as preposition)
II	general preposition
IO	of (as preposition)
IW	with, without (as prepositions)
JJ	general adjective
JJR	general comparative adjective (e.g. older, better, stronger)
JJT	general superlative adjective (e.g. oldest, best, strongest)
JK	catenative adjective (able in be able to, willing in be willing to)
MC	cardinal number, neutral for number (two, three..)
MC1	singular cardinal number (one)
MC2	plural cardinal number (e.g. sixes, sevens)
MCGE	genitive cardinal number, neutral for number (two's, 100's)
MCMC	hyphenated number (40-50, 1770-1827)
MD	ordinal number (e.g. first, second, next, last)
MF	fraction, neutral for number (e.g. quarters, two-thirds)

付録 2. 2 CLAWSの分類項目 2	
ND1	singular noun of direction (e.g. north, southeast)
NN	common noun, neutral for number (e.g. sheep, cod, headquarters)
NN1	singular common noun (e.g. book, girl)
NN2	plural common noun (e.g. books, girls)
NNA	following noun of title (e.g. M.A.)
NNB	preceding noun of title (e.g. Mr., Prof.)
NNL1	singular locative noun (e.g. Island, Street)
NNL2	plural locative noun (e.g. Islands, Streets)
NNO	numeral noun, neutral for number (e.g. dozen, hundred)
NNO2	numeral noun, plural (e.g. hundreds, thousands)
NNT1	temporal noun, singular (e.g. day, week, year)
NNT2	temporal noun, plural (e.g. days, weeks, years)
NNU	unit of measurement, neutral for number (e.g. in, cc)
NNU1	singular unit of measurement (e.g. inch, centimetre)
NNU2	plural unit of measurement (e.g. ins., feet)
NP	proper noun, neutral for number (e.g. IBM, Andes)
NP1	singular proper noun (e.g. London, Jane, Frederick)
NP2	plural proper noun (e.g. Browns, Reagans, Koreans)
NPD1	singular weekday noun (e.g. Sunday)
NPD2	plural weekday noun (e.g. Sundays)
NPM1	singular month noun (e.g. October)
NPM2	plural month noun (e.g. Octobers)
PN	indefinite pronoun, neutral for number (none)
PN1	indefinite pronoun, singular (e.g. anyone, everything, nobody, one)
PNQO	objective wh-pronoun (whom)
PNQS	subjective wh-pronoun (who)
PNQV	wh-ever pronoun (whoever)
PNX1	reflexive indefinite pronoun (oneself)
PPGE	nominal possessive personal pronoun (e.g. mine, yours)
PPH1	3rd person sing. neuter personal pronoun (it)
PPHO1	3rd person sing. objective personal pronoun (him, her)
PPHO2	3rd person plural objective personal pronoun (them)
PPHS1	3rd person sing. subjective personal pronoun (he, she)
PPHS2	3rd person plural subjective personal pronoun (they)
PPIO1	1st person sing. objective personal pronoun (me)
PPIO2	1st person plural objective personal pronoun (us)
PPIS1	1st person sing. subjective personal pronoun (I)
PPIS2	1st person plural subjective personal pronoun (we)
PPX1	singular reflexive personal pronoun (e.g. yourself, itself)
PPX2	plural reflexive personal pronoun (e.g. yourselves, themselves)
PPY	2nd person personal pronoun (you)
RA	adverb, after nominal head (e.g. else, galore)
REX	adverb introducing appositional constructions (namely, e.g.)

大学入試センター試験 32 年のコーパス分析による
通時的变化と近隣のアジア諸国との共時的比較研究

付録 2.3 CLAWSの分類項目 3	
RG	degree adverb (very, so, too)
RGQ	wh- degree adverb (how)
RGQV	wh-ever degree adverb (however)
RGR	comparative degree adverb (more, less)
RGT	superlative degree adverb (most, least)
RL	locative adverb (e.g. alongside, forward)
RP	prep. adverb, particle (e.g about, in)
RPK	prep. adv., catenative (about in be about to)
RR	general adverb
RRQ	wh- general adverb (where, when, why, how)
RRQV	wh-ever general adverb (wherever, whenever)
RRR	comparative general adverb (e.g. better, longer)
RRT	superlative general adverb (e.g. best, longest)
RT	quasi-nominal adverb of time (e.g. now, tomorrow)
TO	infinitive marker (to)
UH	interjection (e.g. oh, yes, um)
VB0	be, base form (finite i.e. imperative, subjunctive)
VBDR	were
VBDZ	was
VBG	being
VBI	be, infinitive (To be or not... It will be ..)
VBM	am
VBN	been
VBR	are
VBZ	is
VD0	do, base form (finite)
VDD	did
VDG	doing
VDI	do, infinitive (I may do... To do...)
VDN	done
VDZ	does
VH0	have, base form (finite)
VHD	had (past tense)
VHG	having
VHI	have, infinitive
VHN	had (past participle)
VHZ	has
VM	modal auxiliary (can, will, would, etc.)
VMK	modal catenative (ought, used)
VV0	base form of lexical verb (e.g. give, work)
VVD	past tense of lexical verb (e.g. gave, worked)
VVG	-ing participle of lexical verb (e.g. giving, working)
VVGK	-ing participle catenative (going in be going to)
VVI	infinitive (e.g. to give... It will work...)
VVN	past participle of lexical verb (e.g. given, worked)
VVNK	past participle catenative (e.g. bound in be bound to)
VVZ	-s form of lexical verb (e.g. gives, works)
XX	not, n't
ZZ1	singular letter of the alphabet (e.g. A,b)
ZZ2	plural letter of the alphabet (e.g. A's, b's)

