

## TOEFL 語彙の語彙レベル, 真正性, 特徴の検証

水本 篤 (甲南女子大学 非常勤講師)

### Word Levels, Authenticity, and Characteristics Verification of the TOEFL Vocabulary

Atsushi, MIZUMOTO

#### 1. はじめに

TOEFL (Test Of English as a Foreign Language) は, 北アメリカにおける大学などの教育機関に入学, または交換留学するときに必要な英語力の基準とされるテストであるが, 文部科学省が 2003 年度から 5 年間に取り組む施策をまとめた行動計画における, 「英語が使える日本人」の育成を重点とし, 「すべての英語教員が英検準 1 級か TOEFL550 点, または TOEIC730 点以上の英語力がつくよう国内外の研修を強化する」という目標にも挙げられているように, 一般的な英語力を測るテストとしても TOEFL は日本で広く用いられている。しかし, そのテストにおける語彙は, 「普通の英語よりも難しく専門的なものが多い」というはっきりとしない印象しか持たれていないのが実状であるといえる。TOEFL テストを利用するすべての関係者にとってみれば, そのテストにおける語彙がどのような難しさであり, また実際の言語使用を反映しているものであるのか, などという点は明らかにされるべきものであるだろう。ゆえに本研究では, TOEFL コーパスと, その他のコーパスや語彙リストとの比較を行い, いわゆる「TOEFL 語彙」とはどのような特徴があるのかを考察する。

#### 2. 先行研究

Chujo and Nishigaki (2003) は, 中学, 高校までに学ぶテキストの中で使われている English for General Purposes (EGP) の語彙と, サンプルとして用いた TOEFL テスト 20 回分の中で使用されている語彙 (Total Token 145,391 語) との語彙レベルの比較を行い, 2 つの間には大きなギャップがあることを証明した。そして, そのギャップを埋めるために, 1023 語の”Specialized TOEFL Vocabulary List”を作成し, 一般的な読解に必要な語彙カバー率とされる 95 パーセントという数値を, TOEFL においても, EGP 語彙とそのリストの 2 つを学習することで達成することができるようになったと述べている。また, TOEFL を対象としたものではないが, 石川 (2002) の研究では, TOEIC テストにおいて使用されている語彙の特徴とレベルが明らかにされている。

### 3. 研究の目的

前述の Chujo and Nishigaki (2003) の先行研究では、カバー率を中心にして行った研究であり、石川 (2002) のように、それぞれの語彙が全体のレベルとして見た時にどのような特徴があるかまでは述べられてはおらず、このような点を更に調査する余地が残っている。ゆえに本研究では、これらの先行研究のデザインを基にして、より大きな語彙数を含んだ TOEFL コーパスを作成し、そのコーパスにおける語彙の特徴を、その他の広く利用されている語彙リストなどとの情報と比較することにより、TOEFL で使用されている語彙にはどのような特徴があるのかを、難易度と真正性の観点からも検証することを目的とする。

### 4. 研究の手順

#### 4.1. TOEFL コーパスの構築

TOEFL において使用されている語彙を科学的に検証するために、様々な市販教材(資料)から自作の「TOEFL コーパス」を作成した。特に TOEFL テスト開発を行っている ETS (Educational Testing Service) の教材は、過去に実施された TOEFL のマテリアルも含まれており、妥当性が高いものであるといえる。また、その他の市販 TOEFL 教材も、実際の TOEFL に近い形であるものを、全ての問題に目を通した上で選択した。TOEFL コーパスに収録した英文は問題文を含む全てのもので(セクションの始めの指示文は除く)、収録教材のうち CD-R が付属されているものは、その中の問題の文字情報も含んだ。最終的に TOEFL コーパスは総語数 1,054,851 語となり、100 万語を超えるコーパスとなった。この総語数は、語彙単位での分析を行うにおいて、全体像をつかむにはある程度十分な数であるといえるであろう。

#### 4.2. TOEFL 頻出語彙リストの作成

語彙の研究においては、何を 1 語と見なすかがよく問題となるので(例えば、look は名詞と動詞の用法なら 2 語と見なすのかなど)、本研究では JACET8000 (2003) に収録されている v8an というプログラムを使ってレマ化と頻度付与を行った。このプログラムは JACET8000 に基づいて、テキストの中で使われている語をレマ化し、1~8 までのランク付けとテキストカバー率を算出するものである。JACET8000 を基にしたランクの分類は表 1 のようになる。JACET8000 のリストに含まれていない語(例えば flora, octopus などのようなもの)は OVER8 となり、その他、省略形(Cont. forms)や固有名詞(Proper nouns)、そして数字など(Non-words)がこのプログラムによって、自動的に分類された。

表 2 は、本研究で作成した TOEFL コーパスを v8an で分析した結果である。後の分析で用いた「TOEFL コーパス語彙リスト」は、以下の表におけるレベルと、その単語の頻度が

付与された形のものである。このリストと他の語彙リストとの比較においては、以下の表のレベル分けのうち Over 8, Cont. forms, Proper nouns, Non-words は、語彙の難易度（レベル）の対象にはならないので除外して分析が行われた。

表 1 v8an による語彙レベルの分類

分類（レベル）	説明
Level 1	ランクが 1 位～1000 位までの語
Level 2	ランクが 1001～2000 までの語
Level 3	ランクが 2001～3000 までの語
Level 4	ランクが 3001～4000 までの語
Level 5	ランクが 4001～5000 までの語
Level 6	ランクが 5001～6000 までの語
Level 7	ランクが 6001～7000 までの語
Level 8	ランクが 7001～8000 までの語
Over 8	ランク外の語で、すべて小文字表記のもの（表 2 では OV8）
Cont. forms	シングルクォート（'）を含む語（表 2 では Cont.）
Proper nouns	大文字を含む語（表 2 では Prop.）
Non-words	アルファベット以外の文字を含む文字列（表 2 では Non.）

表 2 TOEFL コーパスの v8an による分析結果

Category	Indexes	%	Tokens	%
L 1	1123	4.08	760325	72.08
L 2	992	3.60	74489	7.06
L 3	957	3.47	34782	3.30
L 4	923	3.35	27959	2.65
L 5	836	3.04	17816	1.69
L 6	811	2.94	11340	1.08
L 7	733	2.66	9289	0.88
L 8	675	2.45	7102	0.67
OV8	11910	43.24	41407	3.93
Cont.	259	0.94	10449	0.99
Prop.	2875	10.44	30687	2.91
Non.	5453	19.80	29206	2.77
Total	27547	100.00	1054851	100.00

(注) Indexes は JACET8000 のリストにおける見出し語であり、Tokens は総頻度を示す。

L1 の Indexes が 1000 語を超えているのは、L1 には“plus 250”という 250 語の追加語が含まれているからである。

### 4.3. TOEFL コーパスの比較の対象としたコーパスと語彙リスト

このセクションでは、TOEFL コーパスから作成した、頻度つきの語彙リストとの比較を行った他のコーパスや、語彙リストがどのようなものであったかを説明する。

#### 4.3.1. アカデミック・コーパス

TOEFL は一般的に English for Academic Purposes (以下 EAP) のテストであるといわれている。この事実から、一般的な EAP のテキストと TOEFL コーパスを比較することが必要であると考えられたため、ICAME Collection of English Language Corpora (1999) の EAP テキストにあたる部分を抜き出して「アカデミック・コーパス」として使用した。Academic Corpus として使用したテキストは以下の表 3 である。また、このアカデミック・コーパスを前述のプログラム v8an を使ってレマ化し、JACET8000 のレベル付けを行ったものが表 4 である。

表 3 アカデミック・コーパスに収録したコーパス名とカテゴリー

コーパス	カテゴリー	テキスト数	語数 (約)
Brown Corpus	J Learned and scientific writings	80	160,000
LOB Corpus	J Learned and scientific writings	80	160,000
Frown Corpus	J Learned and scientific writings	80	160,000
FLOB Corpus	J Learned and scientific writings	80	160,000
Wellington Corpus	J Learned and scientific writings	80	160,000

表 4 アカデミック・コーパスの v8an による分析結果

Category	Indexes	%	Tokens	%
L 1	1123	3.05	541732	68.16
L 2	993	2.70	63028	7.93
L 3	952	2.59	27770	3.49
L 4	970	2.64	42161	5.31
L 5	937	2.55	15669	1.97
L 6	872	2.37	9714	1.22
L 7	838	2.28	6371	0.80
L 8	751	2.04	4480	0.56
OV8	17613	47.86	44580	5.61
Cont.	1605	4.36	5475	0.69
Prop.	198	0.54	6750	0.85
Non.	9953	27.04	27078	3.41
Total	36805	100.00	794808	100.00

#### 4.3.2. JACET8000

これまでの TOEFL コーパスと Academic Corpus の分析を行った v8an のレベル付けの基となっている語彙リストが、この JACET8000 である。JACET8000 は 1 億語の British National Corpus (以下 BNC) と日本人英語学習者の環境を踏まえて独自に作成されたサブコーパスに準拠し、かつ中高の教育現場の状況にも配慮して、全面的にコンピュータを活用し、妥当性の高い統計的手法に基づいて作成された「日本人英語学習者のための科学的教育語彙表」である (大学英語教育学会基本語改定委員会, 2003)。語彙選定過程において、リストの順位の変更が日本の教育現場に合わせて行われているが、その順位調整は一貫して統計的手法に基づくものであり、客観性が重視された“Corpus-based”の語彙リストとなっている。本研究では、この JACET8000 における順位 (レベル) を語彙の難易度の基準として用いた。

#### 4.3.3. Academic Word List

Academic Word List (Coxhead, 2000, 以下 AWL) は、学術誌と arts, commerce, law, natural science の 4 分野のサブコーパスを中心とする、大学のテキストの 350 万語から構成されているコーパスを基にして作られたものである。また AWL には、General Service List (West, 1953) の英語における最頻出 2000 語は含まれていない。また、語彙の選定には頻度 (frequency) とともに、「様々なアカデミック・テキストにどれだけ出てくるか」という範囲 (range) の考え方も用いられており、まさにアカデミックな英文を読んだり書いたりするときに必須となる語彙を集めた語彙リストであるといえる。収録語彙は 570 語であるが、単語の派生形や屈折形も含んだワード・ファミリー換算なので、見出し語化せずに全ての形を合わせると、3100 語ぐらいになるであろうと考えられる。

#### 4.3.4. BNC Lemmatized Frequency List

1 億語からなる British National Corpus における高頻度の語彙のうち、頻度が 800 以上のものをレマ化してまとめたものがこの BNC Lemmatized Frequency List (Kilgariff, 1996) である。アメリカ英語や日常語、また spoken の資料が不足しているとはいえ、BNC は現代を代表する「標準英語」のコーパスであるといえる。ゆえに本研究では、EAP と比較を行うための English for General Purposes (以下 EGP) のサンプルとして、この BNC 高頻度リストを使用した。この BNC 高頻度リストでは同じ単語であっても品詞が違えば、別々に違う単語として見出し語扱いされているので (例えば、help は動詞と名詞が別の語として収録されている)、本研究でレマ化した基準 (JACET8000 の v8an) に基づいて変更を加えた。また v8an を使えば、イギリス式スペルであるものもアメリカ式スペルに自動で変換してくれるので、アメリカ式のスペリングで統一したリストを作成した。そして、最終的に修正を加えたリストの収録語異数は 5441 語となった。

#### 4.3.5. TOEIC コーパス

最後に、EAP と EGP とは違うジャンルの語彙を使用していると思われるテキストを収録しているコーパスとして、水本（2004）で作成された TOEIC コーパスを比較の対象として使用した。TOEIC は国際ビジネスの場において、どれだけ英語の運用力があるかを測定するテストといわれているので、English for Specific Purposes (ESP) の考え方に近いテストであるといえる。この TOEIC コーパスを v8an で分析したものが表 5 である。

表 5 TOEIC コーパスの v8an による分析結果

Category	Indexes	%	Tokens	%
L 1	1082	11.21	79561	69.53
L 2	840	8.70	10283	8.99
L 3	649	6.73	3776	3.30
L 4	718	7.44	6229	5.44
L 5	501	5.19	1997	1.75
L 6	380	3.94	1199	1.05
L 7	271	2.81	723	0.63
L 8	241	2.50	609	0.53
OV8	1761	18.25	2729	2.39
Cont.	82	0.85	903	0.79
Prop.	853	8.84	2468	2.16
Non.	2273	23.55	3950	3.45
Total	9651	100.00	114427	100.00

## 5. 結果と考察

### 5.1. TOEFL 語彙の難易度

本研究で使用した 3 種類のコーパス (TOEFL コーパス, アカデミック・コーパス, TOEIC コーパス) の語彙の難易度を, v8an を使い, JACET8000 におけるレベル分けをしたものを比較して, TOEFL 語彙の難易度が他と違うのかを調べた結果が図 1 である。これは, 表 2, 4, 5 で示された総語数に対する JACET8000 のそれぞれのレベルでのカバー率を累積する形でまとめたものであるが, この図を見る限り, TOEFL コーパスの語彙が他のコーパスで使用されている語彙と比べて, 特に難しいというわけではないということが明らかになった。むしろ, TOEFL コーパスに対する JACET8000 の 4000 語 (L4) までのレベルでのカバー率は, 他の 2 つのコーパスよりも高いことから, TOEFL には基礎的な語彙が多く出現するということが考えられる。また, 3 つのコーパスのそれぞれのレベル (Others を除く L1 から OV8 までの 9 つ) での累積カバー率を使って, クラスカル・ウォリスの検

定を行った結果、3つのコーパスにおける平均カバー率には有意差はなかった。また、多重比較においても、どのコーパス間の組み合わせにも差がないということが確認された。

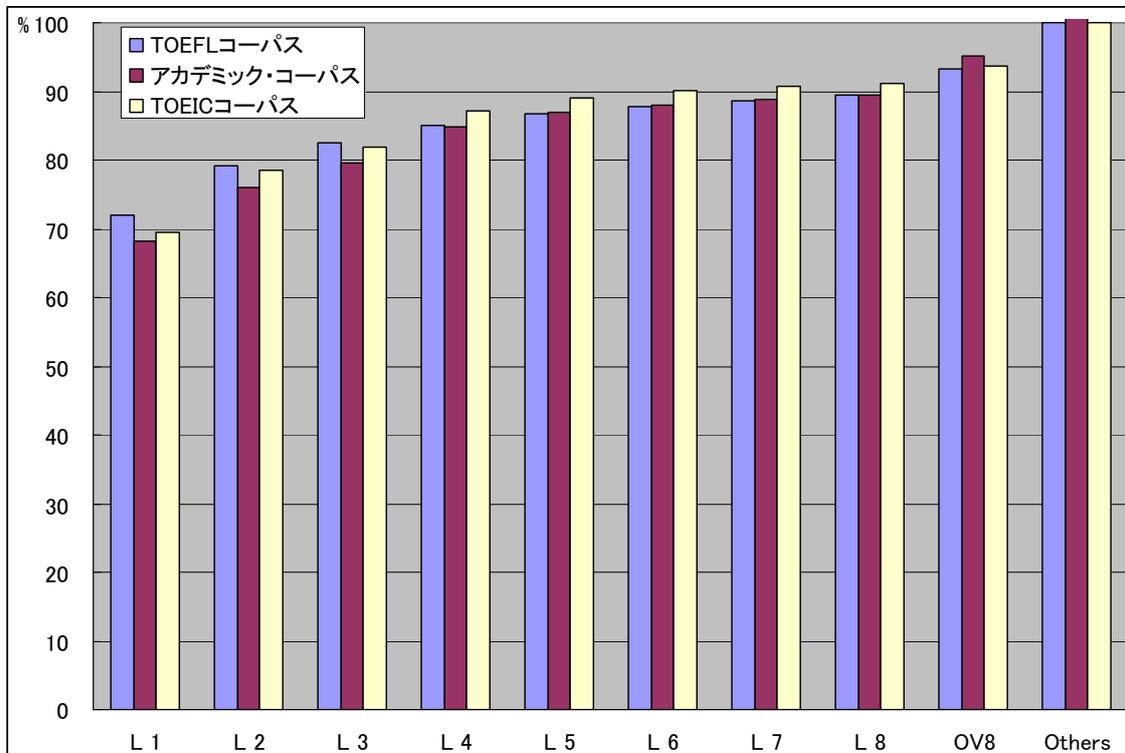


図1 3つのコーパスにおける JACET8000 のカバー率

(注) Others は省略形 (Cont. forms), 固有名詞 (Proper nouns), そして数字など (Non-words)をまとめたもの。

## 5.2. TOEFL 語彙の真正性

### 5.2.1. Academic Word List (AWL) との比較

AWLの580語のうち、本研究で作られたTOEFLコーパスの語彙リストにおいては、562語が含まれていた(96.9%)。AWLは4.3.3.で説明した作成過程を考えてみても、アカデミックな語彙リストであるといえるので、TOEFLはEAPの語彙が多く含まれていることが証明された。

### 5.2.2. クラスタ分析

TOEFLコーパスから得られた語彙リストは、他のどのコーパス、語彙リストと近いものであるのかを比較するために、頻度を基にしてクラスタ分析(平方ユークリッド距離, ウォード法)を行った。クラスタ分析はある集団内で似ているグループを探すときに使われる手法である(Yamamori et al, 2004)。クラスタ分析はSPSS14.0Jを用いて行った。

語彙リストに出現している全ての語を比較するには数が多すぎるので、EGP(English for General Purposes)の対象として用いたBNCの5441語リストを基として、そのリストの

語が TOEFL コーパス, アカデミック・コーパス, TOEIC コーパスのそれぞれの語彙リストにおいてどれだけの頻度であったかを参照することによって行った (図 2)。分析の際には, BNC は 1 億語のコーパスから構成されているものであり, 頻度が他のコーパスと比較できないので, 100 分の 1 にしたものを比較とした。また, TOEIC コーパスが他の 2 つのコーパスと比べて総語数が 10 分の 1 であったので, TOEIC コーパスの総語数は 10 倍にして比較を行った。同様にアカデミック・コーパスも総語数が 100 万語となるように頻度に調整が加えられた。これにより, すべての語の頻度が 100 万語のコーパスにおいて出現するものである仮定に基づいて分析が進められた。

word	bnc	academic	toefl	toeic	var	var
1 the	61873	58423	85625	73500		
2 be	42396	33467	42433	41220		
3 of	30934	35523	32386	31870		
4 and	26879	21032	20470	20740		
5 a	21864	16394	27175	21560		
6 in	19243	19717	23367	19810		
7 to	16209	20036	30085	35720		
8 have	13756	7232	9472	12600		
9 it	10902	5528	9422	5400		
10 for	8879	7761	6996	15400		
11 i	8846	1019	5754	3120		
12 that	7604	8980	14072	10140		
13 you	6955	245	5055	10520		
14 he	6813	1697	3155	1440		
15 on	6807	4791	6293	8340		
16 with	6750	5356	5557	6860		
17 do	5596	1993	6319	2860		
18 at	5342	3512	3944	7220		
19 by	5172	5652	4855	6030		
20 not	4555	4229	6654	5310		
21 this	4619	4935	9153	6790		
22 but	4596	2533	2479	1880		
23 from	4345	3745	4571	5030		
24 they	4334	2011	4310	1740		
25 his	4269	1774	2436	1690		
26 she	3803	332	2211	1020		
27 or	3738	4063	3736	4450		
28 which	3720	4015	3130	2010		
29 as	3642	6553	5776	5530		
30 we	3580	2281	1848	6190		
31 an	3431	3352	5438	4410		
32 say	3335	599	886	830		

図 2 BNC リスト, TOEFL コーパス, アカデミック・コーパス, TOEIC コーパスにおける頻度を SPSS にまとめたもの

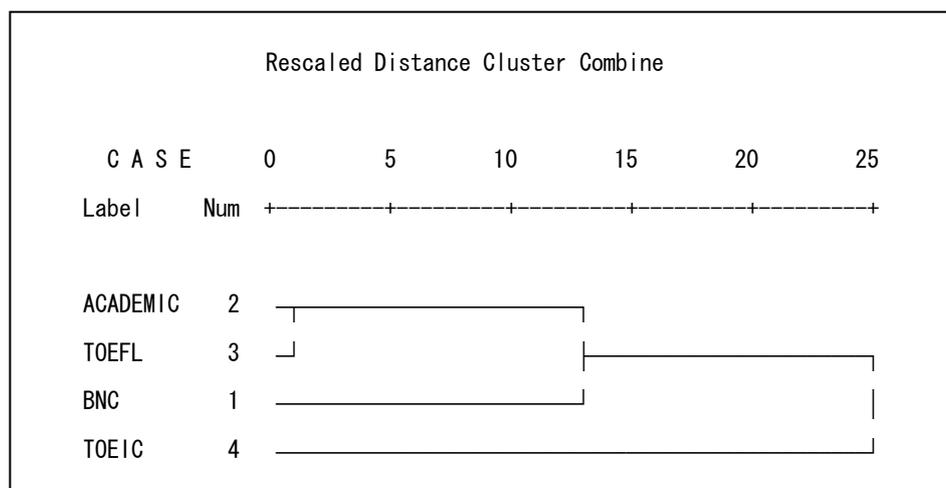


図 3 クラスタ分析の結果 (デンドログラム)

図 3 はクラスター分析の結果をデンドログラムで示したものである。この図からアカデミック・コーパスと TOEFL コーパスが語の使用においては一番近いものであるとわかる。またこの 2 つのコーパスと、基本とした BNC とが近くなっており、TOEIC は、この 2 つとは一番は距離が離れているので、TOEIC とアカデミックなものは対極に位置しているということが明らかになった。

AWL との比較でも立証されたとおり、TOEFL で使用されている語彙はアカデミックなものが多いという裏付けが、このクラスター分析でもなされたといえる。これらの結果から、TOEFL 語彙の真正性が科学的な方法により確認された。

### 5.3. TOEFL 語彙の特徴

ここでは、BNC リストをもとにして、どのような語彙がいわゆる「TOEFL らしい」ものであるかという TOEFL 語彙の特徴を検証してみる。比較の方法としては、BNC の頻度と TOEFL コーパスの頻度を比べるというものであるが、そのままであると 1 億語の BNC コーパスと 100 万語の TOEFL コーパスの比較はできないので、log-likelihood ratio (対数尤度比) というサイズが異なるコーパス間の出現頻度の比較に用いられる手法を用いた。こうして得られた数値を比較して、その語彙が BNC にも現れる「一般的な語彙」なのか、TOEFL に現れやすい「アカデミック語彙」なのかの区別ができるようにした。ちなみに、本研究で用いた BNC リストの 5441 語のうち、BNC にしか現れなかった語が 365 語あり、その語は比較の対象からは除外した。

表 6, 7, 8 は、BNC リストを基にする TOEFL コーパスの語彙リストに、log-likelihood によって得られた数値を付与した表である。log-likelihood (表中では LL の列) の値がマイナスになればそれは BNC に出現頻度の多いものであり、プラスの値ならば TOEFL により多く見られる語彙である。また、絶対値が大きくなればなるほどそれぞれに特有な語彙であることを示している。5441 語のリスト全てを掲載することは原稿スペースの関係上難しいので、TOEFL に対しての log-likelihood の値が一番大きな 20 単語 (表 6)、プラスとマイナスの境界線にある 20 単語 (表 7)、BNC に対しての log-likelihood の値が一番大きな 20 単語 (表 8) のみを掲載した。また、log-likelihood の計算は JACET8000 (2003) に付属のスプレッドシートを用いた。

表 6 で提示されている 20 語は、log-likelihood のプラスの値が一番大きな語彙であり、BNC と比べて TOEFL に出てくる確率の高い語である。これらの語は、問題文や文法説明の中でよく用いられている語も多いことがわかる。また the や state や united が使われやすいという事実は、TOEFL はアメリカ (the United States) のテストであり、それが使われている語彙にも色濃く反映されているということを表しているのは興味深い。また、この 20 単語以外にも music, plant, temperature, jazz などの TOEFL のリーディング・パッセージ特有である語彙も log-likelihood が高い値を示していたことがわかった。

表 6 TOEFL に出てくる確率の高い語 (上位 20 語)

Word	BNC 頻度	TOEFL 頻度	LL
answer	12596	6150	32655.13
sentence	8645	4272	22764.77
first	4772	3005	17143.15
paragraph	3401	2710	16405.50
passage	5116	2920	16218.53
to	1620850	30085	7923.95
professor	2168	1122	6060.71
the	6187267	85625	5710.92
good	100652	4151	5047.93
contain	17732	1752	4576.23
state	41351	2520	4569.98
woman	63087	3032	4376.09
that	760399	14072	3670.78
following	13592	1339	3490.79
therefore	23218	1621	3295.67
united	5081	870	3093.15
colony	2078	666	3079.18
change	37884	1918	2920.66
known	2744	662	2743.20
lecture	2867	632	2521.46

表 7 BNC と TOEFL の両方に出てくる確率の高い語 (20 語)

Word	BNC 頻度	TOEFL 頻度	LL
small	51626	545	0.00
upper	5391	57	0.00
adaptation	1132	12	0.00
factor	14960	158	0.00
classical	3215	34	0.00
continuous	2649	28	0.00
disc	2365	25	0.00
investigator	1040	11	0.00
superb	2085	22	0.00
beast	1232	13	0.00
canvas	1139	12	0.00
production	15837	167	0.00
pursuit	1614	17	0.00
potential	6456	68	0.00
neglect	1429	15	0.00
supermarket	1621	17	0.00
publication	5706	60	0.00
soar	956	10	0.00
sometimes	20517	216	0.00
burial	1052	11	0.00

表 7 は log-likelihood の値がプラスとマイナスの値のちょうど境界線にある語彙であり、比較に用いた BNC, そして TOEFL コーパスの 2 つのどちらのジャンルで使われてもおかしくないといえる語彙である。比較的、日常生活の中で使われることの多い語が含まれているのが特徴であるといえよう。

表 8 BNC に出てくる確率の高い語 (上位 20 語)

Word	BNC 頻度	TOEFL 頻度	LL
any	123655	529	-591.10
case	63577	147	-596.77
will	297281	1844	-619.96
thing	77612	213	-633.12
for	887877	6996	-651.49
know	185534	851	-788.23
you	695498	5055	-789.75
she	380257	2211	-957.96
shall	57056	18	-1035.43
his	426896	2436	-1130.90
if	261089	1178	-1141.88
we	358039	1848	-1205.48
but	459622	2479	-1400.83
I	884599	5754	-1576.31
him	165014	347	-1656.42
may	113024	47	-1975.00
have	1375636	9472	-1978.77
and	2687863	20470	-2405.95
say	333518	886	-2800.10
he	681255	3155	-2844.55

表 8 の 20 語は log-likelihood のマイナスの値が大きな語であり、TOEFL に比べて BNC に出てくる確率が高い語である。表 6 や表 7 と比べてみても、明らかに代名詞や接続詞などの機能語で一般的な使用が多い語が含まれていることがわかる。そして、EAP で特に多く使われるというような特徴のある語は少ないということがいえるであろう。

## 6. おわりに

本研究では、約 100 万語の TOEFL コーパスを基に、さまざまなコーパスや語彙リストとの比較を行うことによって TOEFL 語彙の語彙レベル、真正性、そして特徴を明らかにした。本研究で得られた結果を以下にまとめておく。

- (1) 一般的に TOEFL の語彙は難しいとされているが、他のジャンルのテキストで使われている語彙と比べて、それほど難易度が高いわけではない。
- (2) TOEFL で使用されている語彙はアカデミックなものが多く、アカデミックな語彙リ

スト (AWL) やアカデミック・コーパスと比較すると、とても使用される語彙が近いという結果から、TOEFL で使用される語彙には真正性がある。

(3) BNC という大規模な汎用英語コーパスとの比較では、いわゆる「TOEFL らしい」と考えられている TOEFL の特徴語が存在する。

これらの結果から、TOEFL の語彙を学習、指導する際には、Nation (2001) などでも提案されている、基礎的な語彙を完璧に学習した上で、AWL などアカデミックな語彙を学習するという学習方法がふさわしいということを裏づけており、初級レベルの学習者には、まず頻度の高い基本語をしっかりとマスターさせることが重要であり、いきなり AWL のような語彙を学習させても、学習効果は低いであろうということが示唆されている。また、AWL に収録されているようなアカデミックな語彙や TOEFL に特有の語彙は、一般的な英語テキストではそれほど頻度が高くないために、日本のような EFL 環境においては、基礎語をマスターした後に意識的に学習していく必要があるということも示唆されている。

今回の研究の結果は、TOEFL の語彙が実際にはどのようなものであるのかということをも明らかにする上で大変有益であった一方で、実際に学習者がそれらの語彙をどのように学習すれば良いかということまでは明らかにできていない。また、学習においては単語単位での理解だけではなく、コロケーションなどのより大きな範囲での語彙知識も重要であることはいうまでもないだろう。ゆえに、今後の研究においては、今回の研究結果を語彙学習に効果的に取り入れる方法を模索し、対象を単語レベル以上に広げて研究を進めていくつもりである。

(注)

本研究の一部は、2004 年度大学英語教育学会 (JACET) 関西支部秋季研究大会において、"Characteristics, Word Levels, and Authenticity Verification of the TOEFL Vocabulary" (妻鳥千鶴子氏との共同発表) というタイトルで発表された。

## 参考文献

- Chujo, K. & Nishigaki, C. (2003). Bridging the vocabulary gap: from EGP to EAP. *JACET Bulletin*, 37, 73-84.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34 (2), 213-238.
- 大学英語教育学会基本語改定委員会(編) (2003). 「大学英語教育学会基本語リスト: JACET List of 8000 Basic Words」東京: 大学英語教育学会
- 石川慎一郎. (2002). 「TOEIC の語彙 - その特徴とレベル -」『言語文化学会論集』19. 201-214.
- Kilgarriff, A. (1996). *BNC database and word frequency lists*. Retrieved October 8, 2004, from <http://www.itri.bton.ac.uk/~Adam.Kilgarriff/bnc-readme.html>.
- 水本篤. (2004). 「JACET8000 と TOEIC の相関性: TOEIC ミニコーパスを使っの数量的

検討」『JACET8000 活用事例集』. 60-61. 東京: 大学英語教育学会.

Nation, P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

West, M. (1953). *A General Service List of English Words*. London: Longmans, Green and Co.

Yamamori, K., Isoda, T., Hiromori, T., & Oxford, R. (2003). Using cluster analysis to uncover L2 learner differences in strategy use, will to learn, and achievement over time. *International Review of Applied Linguistics in Language Teaching*, 41, 381-409.

**資料** 本研究の TOEFL コーパスで使用した教材名と出版社

教材名	出版社
TOEFL Practice Tests, Volume 1	ETS
TOEFL Practice Tests, Volume 2	ETS
TOEFL Test Preparation Kit, 2nd Edition	ETS
TOEFL Sample Test, 6th Edition	ETS
POWERPREP Software: Preparation for the Computer-based TOEFL Test	ETS
UNDERSTANDING TOEFL	ETS
MASTER THE TOEFL CBT	ARCO/THOMSON
Delta's Key to the TOEFL Test	Delta
Cambridge Preparation for the TOEFL Test	Cambridge University Press
CBT TOEFL INSIDER	Linga Forum
Cracking the TOEFL CBT 2004	Princeton Review
Longman Prepare for the TOEFL Test: Computer Test Overview	LONGMAN
Longman Complete Course for the TOEFL Test	LONGMAN
Heinle and Heinles Complete Guide to TOEFL Test	Thomson
Kaplan TOEFL CBT	Kaplan
In-A-Flash READING TOEFL	Petersons
アルク 190 点完全攻略直前模試	アルク
アルク 250 点完全攻略直前模試	アルク
TOEFL TEST 対策	トフルゼミナール
Perfect Preparation Book	旺文社
はじめての TOEFL TEST	語研
TOEFL テストパーフェクト模試	東進ブックス
TOEFL テスト実戦問題集	語研