

英和辞書における語彙重要度指定の妥当性の検証

石川 慎一郎

Reliability of Word Grading in English-Japanese Dictionaries

Shin'ichiro Ishikawa

1 はじめに

英語教育における語彙学習の重要性はしばしば指摘されるが、一方で、どの語彙をどの段階で学習すべきか、という具体的な点については、いまだ明確な立場が存在しない。指導要領も、中高で学習する語数の目安については提示するものの、いわゆる中学必修語を除き、その内容に関しては一切触れていないのである。こうした状況の中で、中高の教育現場が、語彙指導に何らかの指針を求めることは当然と言える。こうした指針としては、教科書・受験用単語集・英和辞書などが挙げられるわけであるが、このうち、単語集については、すでに石川慎一郎(1998)で検証したため、本稿では、新たに英和辞書における語彙重要度指定に注目し、その内容と妥当性について詳しく見てゆくこととする。

2 辞書における「中高重要語」

本稿では、高校生から一般社会人を対象にした10種類の辞書の最新版を検討する。辞書名と総収録語数、および本稿中での略記の一覧は下記の通りである(順不同)。

Cl 『新英和中辞典第6版』(研究社,1995) [9万語]; Pg 『小学館プログレッシブ英和中辞典第3版』(小学館,1998) [11.5万語]; Cp 『旺文社新英和中辞典』(旺文社,1999) [13.5万語]; G 『ジーニアス英和辞典改訂版』(大修館書店,1994) [7.6万語]; L 『ライトハウス英和辞典第3版』(研究社,1996) [6.5万語]; A 『スーパーアンカー英和辞典』(学研,1997) [6.5万語]; Pc 『ニュープロシード英和辞典』(ベネッセ,1994) [4.5万語]; F 『フェアリット英和辞典』(東京書籍,1996) [5万語]; S 『サンライズクエスト英和辞典』(旺文社,1999) [4.6万語]; Cn 『グランドセンチュリー英和辞典』(三省堂,2000) [4.2万語]

語彙重要度指定として、各辞書は、中学・高校・大学・一般などの各種の段階を独自に設定し、個々の語がどのレベルに属するかを示している。こうした重要度指定のうち、本稿では、中学校・高等学校の6年間で標準的に学習すべき、いわゆる「中高重要語」に限って検証を行ってゆくこととする。

最初にその総数を管見しておこう。なお、Pg, Pc, Fでは、中学・高校という段階がないため、

それに相当する段階を対象とする。また、Pg, S, Clなどの辞書で設定されている「高校上級」「高校次位重要語」「大学受験」「高校上級～大学教養」といった段階については、中高の標準的学習範囲を逸脱するものと判断し、本稿では対象から除外する。Cpについては、重要語の総数が明示されていないため、冒頭40ページに含まれる当該語の数を調査し、全体のページ数の割合をかけて試算した値(7400語)を示す。

	中 学	高 校	中高総語数
Cl	中学学習程度の基本語 (1000)	高校学習程度の基本語 (1000)	2000
Pg	最重要語 (2000)		2000
Cp	最重要語<中学程度> (?)	次位重要語<高校程度> (?)	7400(?)
G	中学基本語 (1100)	高校基本語 (4800)	5900
L	中学用基本語 (1000)	高校用基本語 (1000)	2000
A	高校での学習に必要と思われる語 (4500)		4500
Pc	現代英語のコミュニケーションに役立つ使用頻度の高い語 (5000)		5000
F	最重要語 (2000)		2000
S	中学高校重要語 (1300)	高校重要語 (2000)	3300
Cn	中学学習語 (900)	高校学習語 (2800)	3700
平均	1100語	2266語	3780語

上表から明らかなように、辞書によって「中高重要語」の数は大きな幅を見せている。少ないものとしては、10種のうち4種までが2000語を指定している。文部省の指導要領によれば、中高の標準的な学習課程で習得すべき語彙の総数は、現行版では2400語であり、2003年施行予定の新指導要領では、2200語となっているが、これは、それらを下回る数字である。一方、多いものでは、3000語台が2種、4000語台が1種、5000語台が2種、さらに、試算値ではあるが7000語台が1種となった。ここで、最小と最大の差は3倍を超えている。

指導要領の言う2200～2400語という語数を遵守した辞書が皆無であったことから窺えるように、辞書は、指導要領の語数指定をさほど重要視していないように思われる。もちろんそこには、高梨庸雄・高橋正夫(1990:52)も言うように、「有名大学合格をねらうのであれば5000～6000語程度の語彙範囲」が必要という大学入試の現状があるわけであるが、それでも、英語教育に大きな影響力を持つ各種の英和辞書の記述が、このように指導要領と乖離し、また、辞書間ですら大きく食い違っていることは、教育現場から見れば混乱の要因であると言わざるを得ない。

なお、本稿では、以下、具体的な検証のために、a-aeの範囲の「中高重要語」のみをサンプルとして議論を進めてゆく。サンプリング範囲の「中高重要語」総数は下記の通りである。

	Cl	Pg	Cp	G	L	A	Pc	F	S	Cn	AV
総語数	2000	2000	7400?	5900	2000	4500	5000	2000	3300	3700	3780
当該語数	31	21	117	95	24	58	89	27	43	56	56.1

すでに見た辞書間の開きはここでも再現しており、しかも、最小のPgと最大のCpの差は実に6倍に及んでいる。さて、10種の辞書が「中高重要語」に指定した語彙数は延べで137語に上るが、これらはどの程度の一貫度を見せているのであろうか。次の表は、特定語が何種類の辞書で同

時に指定されているかを示したものである。

辞書数	10	9	8	7	6	5	4	3	2	1
語数	15 (11%)	7 (5%)	2 (1%)	8 (6%)	8 (6%)	15 (11%)	7 (5%)	20 (15%)	10 (7%)	45 (33%)

ここから分かることは、10種の辞書で共通に指定されている語彙が、総語数のわずか11%に過ぎないという驚くべき事実である。逆に、全体の33%は、いずれか1種類の辞書にしか出現していない。このことは、「中高重要語」の総数に加えて、その内容においても、各辞書が大きな食い違いを見せていることの証左と言える。

興味深いことに、こうした内容上の食い違いは、同一出版社の異なる辞書間にすら存在する。末尾の表にあるように、たとえば、abandoned, abbeyなどは、Cpでは指定されているが、同じ出版社のSでは指定から外れている。また、advance について見ると、Lで指定されている一方、同じ出版社のCIでは指定から外されている。状況はまさに混沌としていると言わざるを得ない。

こうして見てきたように、「中高重要語」は、辞書間で、量的にも内容的にもかなりのズレを含んでいた。そこで以下では、その妥当性を辞書ごとに検証してゆくこととしたい。

3 妥当性調査の方法

語の重要度は、あくまでも、それが実際のコミュニケーションにおいてどの程度用いられているか、という観点から決定されるべきである。したがって、各辞書の「中高重要語」指定の妥当性を検証するためには、オーセンティックな資料に基づいたコーパス・データとの照合が不可欠となる。論者はすでに、石川 (2001)において、時事英語140万語、児童文学120万語、映画セリフ130万語の3種のコーパスを処理し、各コーパスの構成語彙リストを作成している。そこで本稿では、このリストとの照合によって辞書の重要語指定の妥当性を検証してゆく。なお、リストは、活用形の基本形への繰り込み、低頻度語(5%未満)の除去といったデータ上の加工処理を済ませたものである。

さて、石川 (1998)でも指摘したように、語彙重要度指定の妥当性を見るためには、カバー率・ヒット率の2つに注目する必要がある。カバー率とは、各コーパスの構成語彙のうち、いくつが「中高重要語」に組み入れられているかを示す数値である。仮に、時事英語コーパスの構成語彙124語のうち、ある辞書が60語を「中高重要語」に指定していたとすれば、そのカバー率は約50%となる。一方、ヒット率とは、各辞書が「中高重要語」に指定した語のうち、いくつが実際のコーパスに出現していたかを示す数値である。仮に、ある辞書が60語を「中高重要語」と指定し、そのうち30語だけがコーパスの構成語彙になっていたとすれば、そのヒット率は50%となるわけである。

大量に重要語指定を乱発すれば、カバー率は向上するが、ヒット率は低下する。また、重要語指定の数を減らせば、ヒット率は向上するが、カバー率は低下する。つまり、本質的に矛盾する2つの要素をうまく調和させ、高いカバー率とヒット率の両方を同時に達成することこそが、すぐれた語彙指定には望まれるのである。そこで本稿では、カバー率とヒット率を加えたC/H値を示し、語彙指定の妥当性を測る一つの目安としたい。

なお、多角的な角度から検証を行うため、以下では、コーパス別出現語・3コーパス共通出現語・3コーパス延べ出現語の3種のデータに対して照合調査を行う。

4 コーパス別出現語との対照

コーパスとの照合によって語彙資料の妥当性を検証する試みは、これまでも広く行われてきたが、多くの場合、コーパスの全体的分量は問題にされるものの、その構成内容について十分な注意が払われてきたとは言いがたい。だが、石川(2001)で明らかになったように、テキスト・ジャンルによってその構成語彙はかなりの影響を受けるのである。そこで、ここではまず、コーパスごとの出現語に対して照合調査を行う。下表は、3種のデータに対する、各辞書のカバー率とヒット率、および両者を加えたC/H値を一覧した表である。なお、表中のAV(Average)は10種の辞書のそれぞれの値を均した平均、VAR(Variance)は分散、SD(Standard Deviation)は標準偏差を示す。

	Cp	G	Pc	A	Cn	S	Cl	F	L	Pg	AV	VAR	SD
辞書別指定語数	117	95	89	58	56	43	31	27	24	21	56.1	1021.9	32.0
時事英語(全124語)													
ヒット語数	98	75	78	54	50	40	28	27	23	21	49.4	640.8	25.3
カバー率(%)	79.0	60.5	62.9	43.5	40.3	32.3	22.6	21.8	18.5	16.9	39.8	416.8	20.4
ヒット率(%)	83.8	78.9	87.6	93.1	89.3	93.0	90.3	100.0	95.8	100.0	91.2	40.6	6.4
C/H 値	162.8	139.4	150.5	136.7	129.6	125.3	112.9	121.8	114.4	116.9	131.0	241.7	15.5
C/H 順位	1	3	2	4	5	6	10	7	9	8	-	-	-
児童文学(全89語)													
ヒット語数	61	61	54	46	46	38	28	25	22	18	39.9	233.1	15.3
カバー率(%)	68.5	68.5	60.7	51.7	51.7	42.7	31.5	28.1	24.7	20.2	44.8	294.3	17.2
ヒット率(%)	52.1	64.2	60.7	79.3	82.1	88.4	90.3	92.6	91.7	85.7	78.7	189.2	13.8
C/H 値	120.7	132.7	121.3	131.0	133.8	131.1	121.8	120.7	116.4	105.9	123.5	68.8	8.3
C/H 順位	8	2	6	4	1	3	5	7	9	10	-	-	-
映画(全101語)													
ヒット語数	81	72	70	52	49	38	27	27	22	21	45.9	450.9	21.2
カバー率(%)	80.2	71.3	69.3	51.5	48.5	37.6	26.7	26.7	21.8	20.8	45.4	442.0	21.0
ヒット率(%)	69.2	75.8	78.7	89.7	87.5	88.4	87.1	100.0	91.7	100.0	86.8	88.0	9.4
C/H 値	149.4	147.1	148.	141.1	136.0	126.0	113.8	126.7	113.4	120.8	132.2	174.9	13.2
C/H 順位	1	3	2	4	5	7	9	6	10	8	-	-	-

以下、コーパスごとに簡単な概観を行っておく。

(1) 時事英語コーパス出現語との対照

各辞書が指定する「中高重要語」で、英字新聞や雑誌などがどの程度読みこなせるのか、というのは興味深い問題である。しかしながら、時事英語コーパスに対する辞書のカバー率は、平均で4割程度にとどまっている。とくに、Pgの16.9%をはじめ、4種類の辞書ではその値が25%を割っている。もちろん、リーディングの局面においては、スキーマの利用による未知語の推測が可能であるとはいえ、4割程度のカバー率では、実用的なレベルで新聞や雑誌を読みこなすことはきわめて困難であると言わなければならない。

一方、G、Pc、Cpなどではカバー率は6割を超えた。中でも8割に近いカバー率を達成した

Cpについては、おおむね実用に耐えるレベルの語彙指定が行われていると結論してよいであろう。Cpは「時代の推移の中で生まれた新語・新表現と時事語を、この規模の中辞典の紙面が許す限り採録する」こと、具体的には「海外の新聞・雑誌や小説を読む上で、また、国際ビジネスに従事する上で万全といえる語彙収録」を行うことを目標に掲げている(「はしがき」)が、その目標はかなり達成されているようである。

ところで、ヒット率についてであるが、カバー率の低いPg, L, Cl, Fがおおむね90%以上となっているのは当然として、カバー率のもっとも高かったCpも、約84%という高い値になっている。このことは、一見過剰にすら思えるCpの指定語彙の中に、無駄なデータの混入がきわめて少なかったことを示している。なお、カバー率に比べると、ヒット率は標準偏差・分散の値がかなり小さくなっており、辞書間のばらつきが少ないことが分かる。つまりは高率で揃う傾向にあることが分かるであろう。

最後に、C/H値については、Cpの162.8がとくに高く、以下、Pcの150.5、Aの136.7などが続いている。一方、Cl, Pg, Lは110台に留まった。

(2) 児童文学コーパス出現語との対照

一般に、文学作品には、古風な表現や、過剰に凝った表現が散見される。しかし、本調査では、コーパスを児童文学から採取し、また、石川(2001)における語彙表作成の段階で、構成率5%以下の低頻度語(ノイズワード)を削除しているため、当該出現語のデータは、狭義の文学に限らず、平易な読み物を読む際の一般的な語彙リストと見なすことができる。

さて、こうした児童文学コーパスの出現語に対するカバー率であるが、10種の平均値は44.8%となり、時事英語の場合より5ポイント程度上昇している。しかしながら、時事英語で高い値をあげたCpがここでは7割程度に留まるなど、突出した高い値はなくなっている。逆に、2割を割るものもない。標準偏差が時事英語の20.4から17.2に、分散が416.8から294.3に下がっていることから分かるように、データのばらつきは、時事英語に比べると小さく収束している。このことは、多くの英和辞書が、時事英語よりも、むしろ、児童文学的な読み物のカバーを念頭に置いていることの反映かもしれない。

しかしながら、ヒット率になると様相は一変する。平均値は、時事英語の91.2%から78.7%へと10ポイント以上低下している。また、標準偏差・分散の値も上昇し、データのばらつきはむしろ大きくなっているのである。事実、時事英語では、全ての辞書がヒット率においておおよそ8割のラインを超えていたが、一方で、児童文学では、3種の辞書が6割台かそれ以下となっている。この3種は、カバー率が高かったCp, G, Pcであるが、こうした結果は、現代英語における高頻度語の中に、児童文学ではほとんど出現しない語群が少なくないことを示すものと言ってよい。

最後にC/H値であるが、時事英語に比べると、総じて低いレベルに集中している。G, A, S, Cnなどが130台と比較的高い値となっているが、下位との差は決して大きくない。

(3) 映画セリフコーパス出現語との対照

石川(2001)では、映画のセリフが、語彙データとしてもっとも標準性が高いという結論を得た。つまり、映画コーパスは、ESP(English for Special Purpose)的な、ジャンルに限定された語が少なく、汎用性の高い語彙が大半を占めているのである。この結果に従えば、映画

コーパスは、単に口語のみならず、現代の標準的な英語の総体をかなり正確に反映したものと
 言えよう。

さて、まずカバー率についてであるが、平均値は児童文学の場合とほぼ同じ45.4で、時事英
 語よりは5ポイントほど高くなっている。とくに高いものはCpの80.2%で、Gの71.3%、Pcの
 69.3%がこれに続く。ここで、Cpが、全てのコーパスに対してもっとも高いカバー率を記録した
 ことは注目に値しよう。なお、標準偏差や分散の値から見ると、映画コーパスのカバー率は、
 時事英語以上に辞書間のばらつきが大きい。これは、書き言葉については一定の配慮を行って
 いるものの、口語データに関しては、いまだ十分に対応していない辞書が多いからではないか
 と推測できる。

一方、低率のグループに関しては、Pgの20.8%をはじめ、あわせて4種が2割台に留まった。
 これは、ヒアリングの問題を別としても、映画のセリフ内容が、わずかその程度しか把握できな
 いことを意味する。すでに述べたように、映画コーパスは、他では用いられないような特殊語
 彙の混入が少なく、現代英語の標準に近いものである。だとすれば、映画コーパスに対するわ
 ずか2割というカバー率は、実践的コミュニケーション能力の涵養という英語教育の目標から
 見て、きわめて憂慮すべきものと言わざるを得ない。

ヒット率は、平均86.8%で、これは時事英語の91.2%と児童文学の78.7%のおおよそ中間に位
 置する。また、データのばらつき具合も、おおよそ他の2つのコーパスの中間である。ヒット
 率がもっとも低かったのは、Cpの69.2%であるが、同辞書のカバー率の高さを考えれば、約3
 割のアンヒットは許容の範囲と判断したい。

なお、C/H値は、平均で時事英語コーパスに近い。データの分散具合は比較的大きく、Cp、G、
 Pcの3種が150に近接する一方で、2種類は110台に留まっている。

5 3 コーパス共通出現語との対照

頻度は語彙の重要度を定める重要な要素であるが、より正確に語彙の重要度を測定するため
 には、村田年(1997)も言うように、「頻度と分布度」の両方に配慮することが望まれる。こうし
 た「分布度」への注目は、「語のスペクトル」分析を行った竹蓋幸生(1981)の基本姿勢とも共通
 するものである。

そこで、ここでは、分布度の高い語彙、つまりは3種のコーパス全てに出現した共通出現語の
 リストに基づき、改めて照合調査を行いたい。石川(2001)で示したように、3コーパスを構成
 する延べ164語のうち、全コーパスで共通に出現する語彙はわずか56語であった。いわゆる核
 語彙(core vocabulary)として、とくに高い重要度を持つこれらの語との対照調査の結果は
 以下に示す通りである。

	Cp	G	Pc	A	Cn	S	Cl	F	L	Pg	AV	VAR	SD
辞書別指定語数	117	95	89	58	56	43	31	27	24	21	56.1	1021.9	32.0
ヒット語数	53	48	46	40	40	36	27	25	21	18	35.4	131.2	11.5
カバー率(%)	94.6	85.7	82.1	71.4	71.4	64.3	48.2	44.6	37.5	32.1	63.2	418.5	20.5
ヒット率(%)	45.3	50.5	51.7	69.0	71.4	83.7	87.1	92.6	87.5	85.7	72.5	281.2	16.8
C/H 値	139.9	136.	133.8	140.4	142.9	148.0	135.3	137.2	125.0	117.9	135.7	68.2	8.3
C/H 順位	4	6	8	3	2	1	7	5	9	10	-	-	-

分散や標準偏差が示すように、カバー率、ヒット率共に、データ間にはかなり大きなばらつきが生じている。とくにヒット率における281.2という分散値は、すでに見たコーパス別調査の場合と比較しても大きな値である。これは、元データの56語に対して、全体の半数に当たる5種の辞書が、それを超える数を重要語として指定していることによる。

まず、カバー率について見ることにしよう。Cpの94.6%を筆頭に、当該範囲に約90～120語を指定した3種の辞書では、カバー率はすべて8割を超えているが、一方、指定語数が20台の3種の辞書では、その値は3～4割台にとどまっている。ジャンルを超えた汎用性の高い語彙群に対するカバー率が、わずか5割に満たないというのでは、中高6年間の英語教育修了時の語彙力としては、やはり問題が残ると言わなければならない。

次にヒット率であるが、指定語数の多いCp、G、Pcの3種は、揃って5割前後にまで落ち込んでいる。一方で、Fの92.6%を筆頭に、指定語数の少ないものはおおむね9割前後の値を記録している。このように、指定語数とヒット率の間にはおおむね反比例の関係が成立するわけであるが、もっとも指定語数の少ないPgのヒット率が、最高値から数えて4番目となっているように、その関係は絶対ではない。

ところで、56語というきわめて小さいデータに対して、総合的な高成績を収めるためには、カバー率とヒット率のバランスがとくに重要となる。興味深いことに、C/H値に基づく順位を見ると、指定語の多いグループと少ないグループは共に低順位となっている。これは、前者の場合はヒット率が、後者の場合はカバー率が著しく低かったためと考えられるであろう。一方、C/H値で上位にきたのは、A、Cn、Sなど、指定語数が40～60語のもので、指定語数の点で中間に位置するグループであった。このことは、「中高重要語」の選定に当たって、必ずしも語数の多いものがよいとはばかりは言い切れないことを示すきわめて重要な結果である。

6 3 コーパス延べ出現語との対照

最後に、英語コミュニケーションの総体をもっとも網羅的に反映する資料として、3つのコーパスに出現した延べ164語のリストをふまえた照合調査を行うことにしたい。結果は下表の通りである。

	Cp	G	Pc	A	Cn	S	Cl	F	L	Pg	AV	VAR	SD
辞書別指定語数	117	95	89	58	56	43	31	27	24	21	56.1	1021.9	32.0
ヒット語数	107	89	87	58	54	40	28	27	23	21	53.4	878.6	29.6
カバー率(%)	65.2	54.3	53.0	35.4	32.9	24.4	17.1	16.5	14.0	12.8	32.6	326.7	18.1
ヒット率(%)	91.5	93.7	97.8	100.0	96.4	93.0	90.3	100.0	95.8	100.0	95.8	11.8	3.4
C/H 値	156.7	148.0	150.8	135.4	129.4	117.4	107.4	116.5	109.9	112.8	128.4	302.6	17.4
C/H 順位	1	3	2	4	5	6	10	7	9	8	-	-	-

まずはカバー率から見ていくことにしよう。最初に気付くことは、その平均が32.6%というきわめて低い値に留まっていることである。このことは、中高の標準的な教育課程を修了した学習者が、たとえば100語の文章を読むとして、実にその7割が未知語であることを意味する。いかにスキーマを利用しようとも、3割の知識で残りの7割を推測することはきわめて非現実的であり、結果的に、現状の「中高重要語」指定では、オーセンティックな英語資料がほとんど理解できないということになる。もっとも、10種の中には、水準を上回るものも少なくはない。G

とPcはカバー率が5割を越え、Cpは7割に迫っている。先ほどとは逆に、7割の知識で3割の未知語を推測する形となる点で、Cpの語彙指定は、かなり実践的な段階に達していると言える。

次に、ヒット率について管見したい。ヒット率は、指定語数の少ないものでは高く、多いものでは低くなる傾向があるわけであるが、興味深いことに、もっとも指定語数の多いCpでもヒット率は9割を超えており、予想されたほどのばらつきは出ていない。これは、カバー率の分散が326.7であるのに対して、ヒット率の分散が11.8の幅で収束していることにも明らかである。ヒット率は、総じて、高い水準で一致していると言える。

最後にC/H値を管見するわけであるが、ヒット率で差が出ないために、C/H値は、おおむねカバー率の差をそのまま反映することになる。理論上の最高値200に対して、平均は128、とくに指定語数の少ない辞書ではその値は100近くまで下がっている。

7 必要な「中高重要語」のサイズ

今回調査した10種の辞書は、当該範囲における指定語の数から見て、Cl, F, L, Pgという少数(約20~30語)グループと、A, Cn, Sの中程度(約40~60語)グループ、それにCp, G, Pcの多数(約90~120語)グループに別れる。これらを比較すると、多数を指定したグループが、カバー率のみならず、全般的に良い結果を収めているようである。たとえば、3コーパスの延べ出現語に対する前章の検証結果においても、C/H値を概観すれば、多数グループが1~3位、中程度グループが4~6位、少数グループが7~10位をそれぞれ占めていたことに気が付く。

本研究は、限られたサンプルの調査に過ぎず、望ましい「中高重要語」の全体について論じることは目的ではない。しかしながら、サンプル結果から「中高重要語」の望ましいサイズを推測することは無益ではなからう。a-aeの範囲における3コーパスの延べ出現語は164語であったが、これに、辞書のa-ae間ページ数と総ページ数のおおよその比1:7をかけると、コーパスに出現する異語の総数はおおよそ11500語と推定できる。ここから、どの程度の語数を指定すれば、どの程度のカバー率をあげられるかの見当が付く。

指定語数	2000	3000	4000	5000	6000	7000	8000	9000	10000	11000	12000
予想カバー率(%)	17	26	35	43	52	61	70	78	87	96	104

前章の調査結果では2000語を指定した4種の辞書のカバー率は13~17%であり、一方、7400語(試算値)のCpでは65%、5900語のGでは54%となっていた。これらはおおむね、上表の目安と一致するものである。ところで、上表によれば、5割のカバー率を望むならば、必要な語彙数は6000、8割のカバー率を望むならば語彙数は9000ということになる。中高6年間の英語教育の目標をどこに置くかは議論が別れるところであるが、少なくとも実用英語を標榜する以上、文部省が主張する2200~2400という語数が、あまりにコミュニケーションの実態から乖離したものであることはもはや明白である。

8 辞書の「中高重要語」指定における偏向性

考察を終える前に、個々の語彙について、3コーパスにおける合計頻度と、それを「中高重要語」に指定した辞書の数とを比較しておくことにしたい。以下は、3コーパスの延べ出現語164語を降順にソートし、合計出現頻度118,050の0.1%(118回)ラインを超える高頻度語と、0.01%(11回)ライン以下の低頻度に分け、それぞれの語について、「中高重要語」指定を行う

HIGH FREQ.WORDS	3 Cps Ttl	Dic (/ 10)
a	91209	10
about	10541	10
across	1139	10
able	873	10
add	846	10
act	649	10
actually	565	10
according	541	5
above	533	10
action	529	10
account	487	9
address	406	10
accident	334	9
accept	334	10
advance	330	7
administration	285	6
access	273	3
admit	258	10
activity	238	10
advantage	223	9
adventure	222	7
actor	200	4
absolutely	194	5
advice	193	10
addition	191	6
ability	181	9
ad	164	3
acquire	157	5
adopt	154	6
additional	154	3
adult	136	5
abuse	130	3
achieve	125	8
aboard	123	4
adviser / or	122	3
acknowledge	121	4
accompany	120	5
admire	118	7

LOW FREQ.WORDS	3 Cps Ttl	Dic (/ 10)
adrenalin	11	0
accountant	11	1
accession	11	0
ab	11	0
admirable	10	2
absently	10	0
aback	9	0
admiring	8	0
adm	8	0
accessory	8	3
abracadabra	8	0
adjourn	7	0
aching	7	0
abominable	7	0
aesthetic	6	1
admonish	6	0
adjoin	6	0
acidosis	6	0
abstract	6	4
absorbing	6	0
abomination	6	0
abashed	6	0
aeroplane	0	5
advert	0	1
adverse	0	1
adverb	0	4
adolescent	0	1
adjective	0	4
addiction	0	1
acquainted	0	1
acoustic	0	1
accidental	0	1
abundant	0	1
abundance	0	1
absurdity	0	1
abrupt	0	1
abnormal	0	1
abdomen	0	1
abbey	0	1
a(alphabet)	0	1

た辞書の冊数を示したものである。

最初に高頻度語について見ておこう。高頻度語でありながら、指定辞書数が過半数の5以下となったのは下記の語群である。括弧内の左側の数字はコーパス頻度を右側の数字は指定辞書数を示している。

according (541-5), access (273-3), actor (200-4), absolutely (194-5), ad (164-3),
acquire (157-5), additional (154-3), adult (136-5), abuse (130-3), adviser (122-3),
acknowledge (121-4), accompany (120-5)

ここでは、ad, abuse, adviserなど、現代社会において広範囲に使われる語群が目立つ。これは、辞書の語彙指定がこうした社会の動きをいまだ十分に反映していないことをうかがわせる。中でも興味深いのは、広告を表すadである。コーパス出現頻度はadvertisementが26回であるのに対し、その省略語であるadは、7倍近い164回となっている。つまり、実際のコミュニケーションにおいては、adはもはや非公式な短縮語ではなく、独立した一般語彙となっているのである。しかしながら、指定辞書数はadvertisementの5冊に対して、adは3冊に過ぎない。短縮形より基本形を学ばせたいとする辞書編集者の教育的な意図は分からぬではないが、出現頻度にこれほど大きな差があるという現状をふまえれば、むしろadのみを指定するべきであると言える。辞書の基本形偏重は、また、absolutelyにも窺える。absoluteのコーパス頻度が45回であるのに対し、副詞形のabsolutelyの頻度は194回で4倍を超えている。これは、コミュニケーションにおけるabsolutelyの圧倒的優勢を示すデータであるが、にもかかわらず、どちらの語も指定辞書数は5冊ずつなのである。基本形を重視するあまり、実際にコミュニケーションに用いられる語の形態を軽視する姿勢が、そこに透けて見えるであろう。

また、今ひとつ注目すべきは、acquire / accessという高頻度語が、辞書ではしばしば指定から外されることである。名詞や動詞として幅広く使用可能なaccessは、いわゆるカタカナ英語としても、すでに日本人になじみの深い語彙であるが、にもかかわらずそれを指定した辞書はわずか3種にとどまっている。辞書がaccessやacquireを「中高重要語」として避けた判断には疑問が残る。

次に、コーパスでは低頻度語でありながら、4種以上の辞書が「中高重要語」に指定した語彙について見ておきたい。

accessory (8-3), abstract (6-4), aeroplane (0-5), adverb (0-4), adjective (0-4)

かつての日本の英語教育の問題点としては、文法・訳読・英国語に対する過度の偏重が挙げられる。わずか5語ではあるが、辞書が必要以上に重視する語彙を見ていると、そうした古いタイプの教育の名残がいまだ払底できていない状況が浮かび上がってくる。

まず、400万語に近いデータの中で1度も出現しなかった文法用語が、ここに2つも入っていることに注目したい。現在の中高の英語教育では、文法項目の原語名を指導することは行われておらず、実際、辞書自身が品詞名を日本語表記している。だとすれば、これらを今なお「中高重要語」とする必然性を理解することは困難である。次にabstractであるが、これも、過度に抽象的な文章を訳読形式で読ませる古いタイプの教育を連想させる語である。最後に、aeroplane

であるが、英国データも含むコーパスでの出現が0回であったにも関わらず、それが「中高重要語」として過半数の辞書で指定されていることは、英国語重視、というよりむしろ、実用的な米語の軽視の反映と言えるものである。

9 おわりに

以上で我々は、3つコーパス・データに基づき、日本の英和辞書の「中高重要語」指定の妥当性を検証してきたが、その結果、現行の語彙指定にはなお改善の余地が大きいことが明らかとなった。

冒頭でも述べたように、文部省の指導要領が、中高6年間で学ばべき語彙の内容についてほとんど触れていない現状の中で、英和辞書は、英語教育の現場にきわめて大きな影響力を持つものであり、その責任は重大である。今後の辞書改訂作業においては、オーセンティックなデータをふまえた上で、従来の「中高重要語」指定を抜本的に見直すことが強く求められよう。

注記

本稿は、2000年6月10日、静岡大学において開催された大学英語教育学会（JACET）中部支部2000年度大会における研究発表原稿を全面的に改訂したものである。

参考文献

- Fries, Charles. C.& A. Aileen.Traver (1950) *English Word Lists*, (Ann Arbor: The George Wahr Publishing Company) [増山節夫訳注 (1958) 『英語の制限単語表』(東京:大修館書店)].
- Hatch, Evelyn & Cheryl Brown (1995) *Vocabulary, Semantics, and Language Education* (Cambridge: Cambridge University Press).
- 稲村松雄(編)(1970)『講座英語教授法第7巻 語い・連語の指導』(東京:研究社).
- 石川慎一郎(1998)『英語コミュニケーションと語彙:大学入試用単語集の有効性の検討』『言語文化学会論集』11号(言語文化学会) pp.43-62.
- _____ (2001)『テキスト・ジャンルと構成語彙』『KELT』16号(神戸英語教育学会) pp.3-17.
- 伊藤健三・島岡丘・村田勇三郎(1990²)『英語学大系第12巻 英語学と英語教育』(東京:大修館書店).
- 高梨庸雄・高橋正夫(1990)『英語教育学概論』(東京:金星堂).
- 竹蓋幸生(1981)『コンピュータの見た現代英語 ポキャブラリーの科学』(東京:エデュカ株式会社).

[2000年10月26日 受理]

付表 (表中の1は指定ありを示す)

	Cl	Pg	Cp	G	L	A	Pc	F	S	Cn	Ttl		Cl	Pg	Cp	G	L	A	Pc	F	S	Cn	Ttl
a(alphabet)				1							1	acknowledge			1	1		1	1				4
a (冠詞)	1	1	1	1	1	1	1	1	1	1	10	acorn							1				1
abandon			1	1		1	1		1		5	acoustic			1								1
abandoned			1								1	acquaint				1		1	1				3
abbey			1								1	acquaintance			1	1							2
ABC			1								1	acquainted			1								1
abdomen							1				1	acquire			1	1		1	1		1		5
ability	1	1	1	1		1	1	1	1	1	9	acquisition			1				1				2
able	1	1	1	1	1	1	1	1	1	1	10	acre			1	1			1		1		4
abnormal							1				1	across	1	1	1	1	1	1	1	1	1	1	10
aboard			1	1			1	1			4	act	1	1	1	1	1	1	1	1	1	1	10
abolish			1								1	acting			1								1
abortion			1								1	action	1	1	1	1	1	1	1	1	1	1	1
about	1	1	1	1	1	1	1	1	1	1	10	active	1	1	1	1	1	1	1	1	1	1	10
above	1	1	1	1	1	1	1	1	1	1	10	actively							1				1
abroad	1		1	1	1	1	1		1	1	8	activist			1								1
abrupt			1								1	activity	1	1	1	1	1	1	1	1	1	1	10
absence	1		1	1		1	1		1	1	7	actor			1	1		1	1				4
absent	1		1	1	1	1	1	1	1	1	9	actress			1	1		1	1				4
absolute			1	1		1	1		1		5	actual	1		1	1		1	1		1	1	7
absolutely		1	1	1			1		1		5	actually	1	1	1	1	1	1	1	1	1	1	10
absorb			1	1		1	1				4	acute			1	1			1				3
abstract			1	1		1	1				4	ad			1			1	1				3
absurd			1	1			1				3	adapt			1	1		1	1				4
absurdity			1								1	add	1	1	1	1	1	1	1	1	1	1	10
abundance			1								1	addict			1								1
abundant			1								1	addiction			1								1
abuse			1	1			1				3	addition			1	1		1	1	1	1		6
academic			1	1		1	1				4	additional			1	1			1				3
academy			1	1							2	address	1	1	1	1	1	1	1	1	1	1	10
accelerate			1				1				2	adequate			1	1		1	1				4
accent	1		1	1		1	1		1	1	7	adjective	1			1					1	1	4
accept	1	1	1	1	1	1	1	1	1	1	10	adjust			1	1		1	1				4
acceptable			1	1			1				3	adjustment			1	1			1				3
acceptance			1	1			1				3	administer			1	1			1				3
accepted			1								1	administrati			1	1		1	1	1	1		6
access			1	1			1				3	administrati			1	1			1				3
accessory			1	1			1				3	administrato			1				1				2
accident	1		1	1	1	1	1	1	1	1	9	admirable			1	1							2
accidental			1								1	admiral			1								1
accidentally							1				1	admiration			1	1							2
acclaim			1								1	admire	1		1	1		1	1		1	1	7
accommodate			1	1		1	1				4	admission			1	1		1	1				4
accommodation			1	1			1				3	admit	1	1	1	1	1	1	1	1	1	1	10
accompany			1	1		1	1		1		5	adolescent			1								1
accomplish			1	1		1	1		1		5	adopt			1	1		1	1	1	1		6