

# 日本人大学生の英語語彙サイズ

野 中 辰 也

English Vocabulary Size of Japanese University Students

NONAKA Tatsuya

## 1. はじめに

本研究は、日本人学習者の英語習得研究の基礎データ収集を目的として、大学生の英語語彙サイズについて考察する。

英語母語話者の語彙サイズ（どれくらいの語彙を知っているか）については、これまでにいくつかの報告がなされている。Aitchison（2003）は、教養ある成人英語母語話者は、辞書の見出し語換算で最低50,000語を理解し、潜在的に使えるとしている。Goulden et al.（1990）は、大学生英語母語話者は、派生語をまとめてカウントするワードファミリー換算で17,200語を理解できるとしている。

こうした、外国人学習者には習得不可能に思える膨大な語彙サイズの報告がある一方で、日本人学習者にも達成可能と思える語彙サイズデータもある。Nation（1990）は、大型辞書の見出し語総数128,000のうち、通常のテキストで87%を占める高頻出語2,000に加えて同8%を占める大学レベル語彙800を知っていれば、通常のテキストは満足いく程度に理解できるとしている。また、Hirsh & Nation（1992）は、「読みを楽しむ（reading for pleasure）」ためには5,000語が必要としている。いずれの語彙サイズもワードファミリー換算であるが、この程度であれば、日本人学習者にも習得可能であるように思える。では、実際に日本人学習者の語彙サイズはどれくらいなのであろうか。

日本人学習者の語彙サイズについては、八島（2002）の調査がある。同調査では、5つの都立・県立高校に通う高校2年生643名の平均語彙サイズとしてワードファミリー換算で3,355語という結果が報告されている。（ただし、学校間の生徒の英語力の違いにより、その平均語彙数は2,788～3,793語まで大きな幅があったとしている。）八島（2002）では、さらに、2年間に渡る調査結果として、自らの勤務先の都立高校の1～3年生の平均語彙サイズが示されており、それによると、1年生が2,900～3,100語、2年生が3,200～3,500語、3年生が3,700～3,800語という結果が得られている。

本研究では、上記の八島（2002）の測定方法に準じて、大学生英語学習者の平均語彙サイズを調べることにする。

## 2. 研究の目的

本研究では、日本人大学生英語学習者について、その語彙サイズを測定し、以下の観点から分析・

考察を行なうこととする。

- (1) 被験者全体の語彙サイズの分布と平均語彙サイズ
- (2) 語彙レベル別の正答数
- (3) 語彙テストの項目別正答率

本研究では、「語彙サイズ」の定義として「ワードファミリー換算での受容語彙の推定総量」を、「語彙レベル」の定義として「望月（1998）の語彙サイズテストの各レベル」を、それぞれ採っている。これは、先行研究の八島（2002）に準じたものである。

### 3. 研究方法

#### 3.1 被験者

2年間で2つの大学の1年生172名を被験者としてデータ収集を行なった。内訳は以下のとおりである。

2002年 9月	国立大学工学部 1年生	55名
	私立大学福祉心理学科 1年生	41名
2003年 9月	国立大学工学部 1年生	57名
	私立大学福祉心理学科 1年生	19名

被験者の英語レベルについては、入学時（各調査の半年前）に行った実用英語技能検定準2級の過去問題を使ったテスト（75点満点）で、全体平均が36.4点（最高61点、最低14点、標準偏差10.1）となっている<sup>1)</sup>。英語を専門としない大学生としては、ごく普通のレベル（intermediate から lower intermediate）と言えよう。なお、入学から、語彙サイズ測定までの半年間に、国立大学ではリーディング中心、私立大学ではリスニング中心の授業（12～14週）を行なっている。

#### 3.2 語彙サイズ測定

被験者の語彙サイズ測定には、望月（1998）の語彙サイズテストを使用した。語彙テストとしてはNation（1990）の Vocabulary Levels Test もあるが、望月のテストの方は日本人学習者を対象とした配慮がなされているという点と、語彙レベルの設定が Nation よりも細かいという点から、これを採用することとした。望月の語彙サイズテストでは、1,000語～7,000語まで1,000語レベル毎の測定が可能であるが、先行研究（八島2002）の例に合わせ、1,000語～6,000語レベルの6つのテストフォームを利用した。（Appendix 参照）

#### 3.3 研究手順とデータ分析の方法

2002年および2003年 9月、両大学の英語授業時間内の約60分を使用して、語彙サイズ測定を行なった。

語彙サイズ測定では、各レベル30問の語彙サイズテストを1レベル最大10分という制限時間内で解答させ、6つのレベルを合わせて得た総語彙サイズを最初の分析データとした。語彙サイズは次の計算式により求めたものである。

$$\text{語彙サイズ} = \frac{\text{語彙サイズテストの総得点}}{\text{語彙サイズテストの問題数 (180)}} \times 6,000$$

総語彙サイズに加えて、語彙レベル別正答数の平均値を合わせて算出した。その後、一元配置の分散分析により、レベル別正答数の平均値に有意差があるかを調べた。

最後に、語彙テストの項目別正答率を算出し、語彙レベル別に正答率の高い・低い語彙について、その原因を考察した。

## 4. 結果と考察

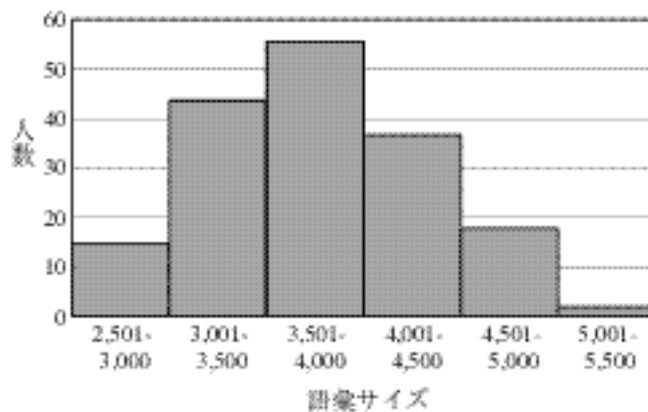
### 4.1 被験者全体の語彙サイズの分布と平均語彙サイズ

172名の被験者の平均語彙サイズは3,772.9語、最高は5,166.7語、最低は2,533.3語、標準偏差は547.45であった。500語毎に区切った語彙サイズレベル毎の度数分布を以下に示す。

表1：語彙サイズ分布 (N=172)

語彙サイズ	人数
2,501-3,000	15
3,001-3,500	44
3,501-4,000	56
4,001-4,500	37
4,501-5,000	18
5,001-5,500	2

図1：語彙サイズ分布 (N=172)



この結果は、八島(2002)で報告されている高校3年生の平均語彙サイズとほぼ同じである。八島が示したデータの被験者は、都立高校のうちでも進学校と分類される高校に通う生徒のため、高校生としては高い数値を示したと考えられる。それに比較して、今回得られた数値は、英語を専門としない学部・学科に通う大学1年生 (intermediate から lower intermediate level) としては、妥当なものと考えられる。

今回得られた平均語彙サイズ3,772.9語は、Nation(1990)が「通常のテキストを満足いく程度に理解できる」とする2,800語を超えてはいるが、Hirsh & Nation(1992)が「読みを楽しむ (reading for pleasure)」ために必要とする5,000語には到達していない。こうした先行研究の示唆は経験則からみると、納得のいくものであると感じられる。実際、語彙レベル3,000語に満たない被験者は授業での読解作業にかなり困難を感じているようであるし、語彙レベル4,500語を超える被験者は、読解作業に大きな困難を感じることはないように思われる。

#### 4.2 語彙レベル別の正答数の平均値

1,000語～6,000語の語彙レベル別の正答数平均値と標準偏差は、表2のとおりである。1,000語レベルはほぼ満点の状態であったが、以降徐々に正答数が減り、5,000語レベルでは正答率5割を切っている。

表2：語彙レベル別の正答数平均値と標準偏差（N=172）

	1,000語	2,000語	3,000語	4,000語	5,000語	6,000語
Mean	29.40	22.04	18.34	16.18	14.12	13.10
SD	0.84	4.06	3.61	4.06	4.46	3.29

なお、レベル別の平均点を一元配置の分散分析により比較したところ、危険率1%未満の水準で有意差が見られた（ $F(5, 855)=3.38, p<.01$ ）。この結果を受けて行った多重比較検定（SNK test）では、危険率1%未満の水準ですべてのレベル間に有意差があることがわかった。

この結果、1,000語レベルの語彙についてはほぼ定着しているが、2,000語レベルでは7割弱の定着度、3,000語レベルでは6割程度の定着度というように下降していきることがわかった。1,000語レベルの定着率の高さは、中学高校6年間の英語学習の後では当然とも言えよう。しかし、2,000語レベルでは定着率が一気に3割近く落ちている。そこで2,000語レベルのテストフォームに含まれる語彙を見ると、"kiss" "simple" といった比較的馴染みのある単語がある一方で、"seize" "urgent" など、ネイティブ・スピーカーには基本語彙であるが、日本人学習者には大学受験レベルと考えられる語彙も入ってきており、こうした単語が推薦入学ほか英語を必要とする試験なしで入学してきた被験者には難しかったと推察される。同様に3,000語レベル以降についても馴染みのある単語の比率に合わせて正答率が下がっていくように見える。さらに、5,000語レベルで正答率が半数を切っているレベルについては、当て推量での正答もかなりの数あるのではないかと考えられる。

また、今回得られた結果を八島（2002）で示されている高校2年生の語彙データ（表3）と比較してみると、高校から大学にかけての語彙習得の進み方についての示唆が得られる。

表3：高校2年生の語彙レベル別の正答数平均値と標準偏差（N=643）

	1,000語	2,000語	3,000語	4,000語	5,000語	6,000語
Mean	29.01	19.75	16.85	13.64	11.40	10.01
SD	1.29	4.50	4.08	4.67	4.78	4.66

八島（2002）より作成

表2、表3の数値を比較してみると、いずれの語彙レベルでも大学生の数値が上回っているが、特に、2,000語レベル以上では、いずれのレベルでも同様（1.49-3.09）に平均正答数が上がっていることがわかる。八島（2002）と今回得られたデータのみで一般化できるものではないが、高校生・大学生学習者にとっては、2,000語レベル以上については特に習得しやすい（習得が速い）語彙レベルがあるわけではなく、いろいろなレベルの語彙を同程度のスピードで習得していくと言えるかも知れない。この点については、さらに多くのデータを収集しての疑問の解決が期待される。

#### 4.3 語彙テストの項目別正答率

1,000語～6,000語の各レベルについて、それぞれの項目の正答率を算出し、正答率順に並べたのが、表3～表8である。それぞれのレベルについて、項目正答率から推測できる部分について考察を加えていくこととする。

表 4 : 1,000語レベル項目別正答率 (%)

	正答	正答率		正答	正答率		正答	正答率
1	face	100.0	11	large	99.4	21	grape	98.3
2	forty	100.0	12	man	99.4	22	onion	98.3
3	hundred	100.0	13	must	99.4	23	cookie	97.7
4	listen	100.0	14	sofa	99.4	24	great	97.7
5	my	100.0	15	town	99.4	25	piece	97.7
6	party	100.0	16	exam	98.8	26	have	97.1
7	sun	100.0	17	nice	98.8	27	meal	97.1
8	early	99.4	18	point	98.8	28	towel	94.8
9	elephant	99.4	19	bridge	98.3	29	mystery	93.0
10	hers	99.4	20	continue	98.3	30	pot	80.2

1,000語レベル(平均正答数29.40、標準偏差0.84)では、ほとんどの項目で高い正答率を示している。いずれも文字通り基本的な語彙項目であり、中学・高校6年間の学習を経て、このレベルの語彙はかなり定着率が高いと言える。

なお、このレベルでは "pot" という項目の正答率が極端に低い値になっている。ここでは、「丸い入れ物」という定義に対して "pot" がわからず、代わりに "bath" や "lamp" といった語を選択した被験者が少なくなかったためである。"pot" に対して、英語の「つぼ、料理に使われる鍋」のイメージはなく、日本語の「魔法瓶(ポット)」のイメージしか浮かばず、「丸い」というイメージから、"bath" あるいは "lamp" を選択したと推察される。

表 5 : 2,000語レベル項目別正答率 (%)

	正答	正答率		正答	正答率		正答	正答率
1	kiss	99.4	11	strength	90.7	21	flood	62.8
2	flag	98.8	12	discover	88.4	22	bonus	61.6
3	limit	98.8	13	view	88.4	23	contain	61.0
4	simple	98.8	14	repair	85.5	24	curious	58.1
5	victory	98.3	15	save	83.7	25	equipment	43.6
6	coach	97.1	16	hate	79.1	26	urgent	39.0
7	automatic	95.9	17	extra	77.9	27	passage	38.4
8	chess	95.3	18	separate	75.0	28	discipline	33.1
9	cabbage	91.9	19	raw	64.0	29	seize	24.4
10	expensive	91.3	20	coast	63.4	30	mend	20.3

2,000語レベル(平均正答数22.04、標準偏差4.06)では、平均正答率が7割程度まで下がり、散らばりも広がってきている。正答率の高い語彙はカタカナで日本語に入っている語も多く、いずれも納得のいく正答率であろう。一方、"mend" "seize" "discipline" ほか正答率の低い語彙は、ネイティブ・スピーカーにとっては基本語彙であろうが、日本人学習者にとっては大学受験レベルと考えられるものと考えられる。被験者には、学部の性格上工業高校出身者が少なくないことに加えて、推薦入学や英語を含まない入試を経て入学してきた学生が多いため、そうした大学受験レベルの語彙の定着度が低かったと推察される。

このレベルで正答率の低い語彙について、ある傾向が見られた。例えば「つかみ取る、奪う」という定義に対して正答 "seize" を選ばなかった被験者は、"appoint" "forgive" "spread" といった誤答を選択

する傾向があった。このことから、被験者は、未知語については、意味ははっきりと覚えていないものの見た記憶のある (familiar) 単語を選ぶ傾向があるのではないかと察せられる。

表 6 : 3,000語レベル項目別正答率 (%)

	正答	正答率		正答	正答率		正答	正答率
1	police	100.0	11	democracy	82.6	21	fell	39.0
2	curl	97.7	12	oven	81.4	22	anyhow	37.8
3	beam	97.1	13	actually	78.5	23	invest	36.0
4	ton	96.5	14	emergency	78.5	24	dine	34.9
5	mechanical	95.9	15	aware	58.7	25	approval	28.5
6	tourist	94.2	16	upright	52.3	26	flesh	26.7
7	psychology	93.6	17	potential	51.2	27	annual	24.4
8	mist	91.9	18	campaign	47.7	28	decay	15.7
9	explanation	90.7	19	revise	47.1	29	identical	14.0
10	ceremony	89.0	20	vessel	39.5	30	urge	13.4

3,000語レベル (平均正答数18.34、標準偏差3.61) では、平均正答率が6割程度となっており、散らばりは2,000語レベルより若干狭くなっている。2,000語レベル同様、正答率の高い語彙にはカタカナで日本語に入っている語が多い一方で、正答率の低い語彙には大学受験レベルの語彙が多いと考えられる。さらに、「しきりに勧める (urge)」に対して "celebrate" "construct" を選ぶなど、未知語について、意味ははっきりと覚えていないものの見た記憶のある単語を選ぶという傾向も引き続き確認された。

表 7 : 4,000語レベル項目別正答率 (%)

	正答	正答率		正答	正答率		正答	正答率
1	telescope	93.0	11	bump	70.9	21	gradually	36.6
2	microscope	90.7	12	resume	69.8	22	bid	34.9
3	chemist	89.0	13	linguistic	64.0	23	reconcile	33.7
4	gallery	86.0	14	singular	64.0	24	purchase	32.6
5	symphony	79.1	15	neutral	57.0	25	solely	27.9
6	widen	77.9	16	craft	54.7	26	plead	23.3
7	logical	73.8	17	transition	54.1	27	clue	21.5
8	packet	73.8	18	equator	47.1	28	hinder	17.4
9	admission	73.3	19	portion	40.1	29	prompt	14.5
10	consumer	72.7	20	flock	38.4	30	feast	6.4

4,000語レベル (平均正答数16.18、標準偏差4.06) では、正答率が5割少々となり、散らばりは3,000語レベルより若干広く、2,000語レベルと同様となっている。このレベルでは、3,000語レベル同様、半数近くの項目について正答率5割を切っているが、"reconcile" "hinder" "feast" など大学受験レベルを超えた語彙が含まれてきているのが見てとれる。高校までの英語学習でこのレベルの語彙習得を充実させるのはなかなか難しいのではないかと推察される。

表 8 : 5,000語レベル項目別正答率 (%)

	正答	正答率		正答	正答率		正答	正答率
1	excellence	80.2	11	genetic	57.6	21	mischievous	32.0
2	inspector	77.9	12	integration	55.2	22	profitable	28.5
3	restoration	74.4	13	organism	54.7	23	repel	27.3
4	pow	72.7	14	petroleum	54.1	24	whirl	24.4
5	identification	68.6	15	stupidity	52.3	25	swarm	23.8
6	fellowship	68.0	16	frantic	51.7	26	fiscal	21.5
7	liquor	67.4	17	niece	50.0	27	mortal	18.0
8	carpenter	66.3	18	constitutional	43.6	28	wring	18.0
9	injection	61.6	19	resemblance	41.9	29	correction	14.5
10	graceful	58.1	20	statistical	33.1	30	roam	14.0

5,000語レベル(平均正答数14.12、標準偏差4.46)では、平均正答率が5割を切っている。正答率5割を切っている項目の割合は、3,000語~4,000語レベルと大差ないように見える一方で、正答率8割を超える語彙項目は1つだけとなった。このレベルについては、大学・社会人レベルの語彙を多く含んでいるため、当て推量による解答が多くなってきているのではないかと推察される。

表 9 : 6,000語レベル項目別正答率 (%)

	正答	正答率		正答	正答率		正答	正答率
1	microwave	91.3	11	envelop	46.5	21	disrupt	27.3
2	economically	84.9	12	flush	41.9	22	congestion	26.7
3	mosaic	82.6	13	reservoir	39.5	23	deduction	25.0
4	cock	78.5	14	paralysis	36.6	24	rivalry	24.4
5	volcanic	78.5	15	pneumonia	33.1	25	scrutiny	21.5
6	sidestep	76.7	16	relevance	32.0	26	feeble	20.9
7	deepen	76.2	17	prosecute	31.4	27	heredity	18.0
8	researcher	69.2	18	void	30.2	28	inflammation	16.9
9	ceramic	64.5	19	tidal	29.7	29	bewilder	14.0
10	institutional	49.4	20	fundamentally	29.1	30	deliberation	14.0

6,000語レベル(平均正答数13.10、標準偏差3.29)では、5,000語レベル同様平均正答率5割を切っている。4,000語レベルでは正答率8割を超える語彙項目が1つであったのに対して、このレベルでは"microwave" "economically" "mosaic" の3項目で正答率8割を超えた。これらは、いずれもカタカナとして日本語に入っていたり("microwave" "mosaic")、定着度が高いと考えられる語("economy")の派生語("economically")であったりするためではないかと考えられる。また、このレベルでは、正答率5割を切る項目が全体の7割に及んでいるのが特筆すべき点である。大学受験レベルの語彙は非常に限られ、大学・社会人レベルの語彙が多くなったため、当て推量や見覚えのある単語を選ぶ解答が多くなってきていると推察される。

## 5 . おわりに

本研究では、高校生を対象とした八島（2002）の研究に準ずる形で、望月（1998）の語彙サイズテストを使用しての語彙サイズ測定を行なった。その結果、英語を専門としない intermediate から lower intermediate レベルの大学生の語彙サイズとして3,772.9語というデータを得た。これは、高校生を対象として同じ測定方法を使った先行研究の結果からみて妥当な数値であることが推察された。

また、1,000語～6,000語レベル別の正答数の平均から、1,000語レベルの語彙はほぼ定着しているが、2,000語レベルの語彙は7割程度、3,000語レベルの語彙は6割程度と定着度が下がっていき、5,000語レベルでは5割を切ることが確認された。

さらに、語彙レベル毎に項目別正答率を算出し、その内容を検討したところ、2,000語～3,000語レベルの語彙については高校レベルで十分対応可能である一方で、4,000語レベルを超える語彙については高校レベルでの習得は徐々に難しくなるということが推察された。さらに、被験者は正答を思いつかない場合には、見覚えのある単語を選ぶ傾向があるのではないかと考えられた。

本研究を通じて、望月の語彙サイズテストは、日本人学習者用によく練られているという印象を受けた。もちろん、本研究だけで日本人英語学習者の語彙サイズがはっきりしたわけではないが、今後、さまざまなレベルの被験者について同じ道具立てでの研究が行なわれ、語彙習得研究についての基礎データが十分に蓄積されることを期待して本稿を締めくくることとする。

### 謝辞

本研究の統計処理について、新潟青陵大学の木村哲夫教授から貴重なご助言をいただいた。

### 注

- 1) ここで行なった試験成績結果について4グループ間に有意差がないことが分散分析により認められた。
- 2) 八島（2002）では、勤務先の高校のレベルを、大手予備校が提示する偏差値で58、「3年生になると7割以上の生徒が4年制大学・短大への進学を希望している」と記している。

### 参考文献

- Aitchison, J. (2003). *Words in the Mind: An Introduction to the Mental Lexicon*. (3rd ed.) Oxford: Blackwell.
- Coady, J. & Huckin, T. (Eds.) (1997). *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press.
- Goulden, R, Nation, P., & Read, J. (1990). "How large can a receptive vocabulary be?" *Applied Linguistics*, 11, 4, pp.341-63.
- 林洋和 (2002). 『英語の語彙指導』 溪水社.
- Hirsh, D. & Nation, I.S.P. (1992). "What vocabulary size is needed to read unsimplified texts for pleasure." *Reading in a Foreign Language*, 8, pp.689-696.
- 門田修平編著 (2003). 『英語のメンタルレキシコン』 松柏社.
- 門田修平・野呂忠司編著 (2001). 『英語リーディングの認知メカニズム』 くろしお出版.
- Laufer, B. (1992). "How much lexis is necessary for reading comprehension?" In Arnaud, P.J.L. and Bejoint, H. (Eds.), *Vocabulary and Applied Linguistics*. (pp. 126-132) London: Macmillan.
- 望月正道 (1998). 「日本人学習者のための語彙サイズテスト」『語学教育研究所紀要』第12号, 27-53.



- 望月正道・相澤一美・投野由紀夫. (2003). 『英語語彙の指導マニュアル』 大修館書店.
- Nation, I.S.P. (1990). Teaching & Learning Vocabulary. Boston, MA: Heinle & Heinle.
- Nation, I.S.P. (2001). Learning Vocabulary in Another Language. Cambridge: Cambridge University Press.
- Nation, I.S.P. & Warning, R. (1997). "Vocabulary size, text coverage and word lists." In Schmitt, N. & McCarthy, M. (Eds.), Vocabulary: Description, Acquisition and Pedagogy. (pp. 6-19) Cambridge: Cambridge University Press.
- 投野由紀夫編著. (1997). 『英語語彙習得論』 河源社.
- 八島等. (2002). 「日本人高校生の語彙サイズ」『関東甲信越英語教育学会研究紀要』第16号, 29-42.

## Appendix: 望月 (1998) 語彙サイズ測定テスト

## 語彙サイズ測定テスト vst11

日本語の意味を表す英語を (1) ~ (6) の中から選び、その番号を解答欄に書き入れなさい。

1. 小麦粉を焼いた菓子 (1) birthday (2) cookie (3) fork	2. 集まり、会 (4) party (5) star (6) sweater
3. 玉ねぎ (1) grape (2) lettuce (3) onion	4. ぶどう (4) pear (5) rose (6) tree
5. 丸い入れ物 (1) bath (2) lamp (3) phone	6. クッションのある長椅子 (4) pot (5) sofa (6) stove
7. 40 (1) forty (2) hundred (3) month	8. 100 (4) six (5) twelve (6) year
9. 町 (1) bridge (2) garage (3) place	10. 橋 (4) scene (5) square (6) town
11. 食事 (1) air (2) meal (3) piece	12. 1つ、1個、1片 (4) sign (5) sound (6) white
13. 男の人 (1) change (2) elephant (3) man	14. 象 (4) rabbit (5) wolf (6) woman
15. 顔 (1) face (2) finger (3) hair	16. 手ぬぐい (4) leg (5) shoe (6) towel
17. 不可解なこと、不思議なこと (1) act (2) butterfly (3) exam	18. 試験 (4) mystery (5) tennis (6) trouble
19. 点、地点 (1) dam (2) magazine (3) pajamas	20. 太陽 (4) point (5) sun (6) war
21. 持っている (1) do (2) get (3) give	22. しなければならない (4) have (5) must (6) raise
23. 開く (1) add (2) continue (3) die	24. 続ける (4) listen (5) mean (6) understand
25. すてきな、素晴らしい (1) cool (2) hot (3) large	26. 大きい (4) least (5) nice (6) quiet
27. 偉大な (1) complete (2) early (3) great	28. 早く (4) most (5) never (6) usually
29. 彼女のもの (1) below (2) hers (3) my	30. 私の (4) past (5) which (6) whom